SMR/ICHMI/INL/TR-2014/03
INL/EXT- 14-32639
Revision 0

Effects of Levels of Automation for Advanced Small Modular Reactors: Impacts on Performance, Workload, and Situation Awareness

Johanna Oxstrand
Katya L. Le Blanc

July 2014

The INL is a U.S. Department of Energy National Laboratory
operated by Battelle Energy Alliance

SMR/ICHMI/INL/TR-2013/03
INL/EXT- 14-32639
Revision 0

Effects of Levels of Automation for Advanced Small Modular Reactors: Impacts on Performance, Workload, and Situation Awareness

Johanna Oxstrand
Katya L. Le Blanc

July 2014

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Office of Nuclear Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

EXECUTIVE SUMMARY

The Human-Automation Collaboration (HAC) research effort is a part of the Department of Energy (DOE) sponsored Advanced Small Modular Reactor (AdvSMR) program conducted at Idaho National Laboratory (INL). The DOE AdvSMR program focuses on plant design and management, reduction of capital costs as well as plant operations and maintenance costs (O&M), and factory production costs benefits.

It is expected that AdvSMRs will employ technology that is significantly more advanced than the analog systems in the existing light water reactor fleet as well as utilizing automation to a greater extent. Moving towards more advanced technology and more automation does not necessary imply more efficient and safer operation of the plant. Instead, a number of concerns about how these technologies will affect human performance and the overall safety of the plant need to be addressed. More specifically, it is important to investigate how the operator and the automation work as a team to ensure effective and safe plant operation, also known as the human-automation collaboration. There is a distinction between the HAC and human-system interaction (also known as human-computer interaction) where the focus of HAC is on understanding the effects of various characteristics of automation (such as its reliability, processes, and modes) on an operator's use and their awareness of plant conditions.

The HAC research project investigates how to best design the collaboration between the operators and the automated systems in a manner that has the greatest positive impact on overall plant performance and reliability. The overall goal of the HAC research effort is to develop design guidance that supports optimal interaction between humans and automated systems. This guidance will be incorporated into an interactive tool for design and evaluation of HAC. The tool is intended to be used by engineers in their process of designing AdvSMR systems.

This report addresses the milestone M3SR-14IN1301057: Conclude experimental studies using simulated process systems. The milestone marks the completion of a first experimental study using simulated process systems investigating knowledge gaps identified during FY13's studies related to Human-Automation Collaboration for complex and highly automated systems.

Previous activities in the HAC research aimed to identify HAC related knowledge gaps needing to be investigated in order to develop the design guidance. The researchers studied the current state-of-practice of HAC in the human factors and automation research fields, in the AdvSMR industry, and in similar industries. Based on the results from these activities a model of human-automation collaboration was developed. The HAC model identified several important features of automation that influence human-automation interaction including; level of automation, cognitive function automation replaces (e.g., detecting anomalies or taking control actions), adaptability, reliability, process of automation, and automation modes. Using the model, the researchers were also able to identify several important human performance and HAC performance outcomes. These outcomes should be used to evaluate HAC design. Examples of outcomes identified via the model are; system performance, situation awareness, workload, trust in automation, and team Situation Awareness.

The process by which the team plans to develop the guidance can be summarized as: 1) Identify HAC research needs, 2) Identify functions and tasks for AdvSMR, 3) Prioritize research needs based on functions and tasks identified, 4) Conduct HAC research to fulfill the needs , and 5) Synthesize results from research to develop design guidance for HAC in AdvSMR. The investigation of the current state-of-practice of HAC in the human factors and automation research fields, in the AdvSMR industry, and in similar industries as well as the development of the HAC Model supported the identification of high level research needs. The identified needs are:

- Models of Team Work
- Initiators and Triggering Conditions for Adaptive Automation.
- HAC Performance Measures
- Impact of Highly Automated Advanced Small Modular Reactors on Operator Awareness
- Regaining/Reacquisition of Operator Awareness
- Effects of HAC Characteristics on Operator's Use of Automation

The high level research needs and the defined research question were used to identify and prioritize the research gaps to address first. Three analytical studies were identified and conducted: 1) Model of Teamwork, 2) Standardized HAC Performance Measurement Battery, and 3) Initiators and Triggering Conditions for Adaptive Automation. For more information about the analytical studies and their results, see Oxstrand et al, 2013b.

Some knowledge gaps are better suited to be explored by experimental studies rather than analytical. The high level knowledge gaps identified earlier which could be investigated by experimental studies are: Impact of Highly Automated Advanced Small Modular Reactors on Operator Awareness, Regaining/Reacquisition of Operator Awareness, and Effects of HAC Characteristics on Operator's Use of Automation.

The analytical study on initiators and triggering conditions for adaptive automation identified a research path that had potential to address some of the general research questions. Specifically, "How do we get the most benefits out of using automation while keeping the operators engaged in the process at the same time?" Based on its potential to cover multiple knowledge gaps it was decided to further investigate adaptive automation through an experimental study.

Through the literature review conducted during the identification of the current state-of-practice activity, it was found that many research studies have demonstrated the effectiveness of adaptive automation in enhancing performance compared to fully manual and a few studies have also demonstrated that automation fault recovery with adaptive automation is superior to fully automated systems. In addition, many studies have shown that situation awareness is better with adaptive automation than with fully automated systems. Essentially, the promise of adaptive automation is that it provides system performance similar to fully automated designs, but without the negative human performance consequences. However, there are still many unanswered questions regarding the applicability and superiority of adaptive automation. For example, few studies actually compare adaptive automation to intermediate levels, instead most studies focused on comparing the extremes (fully manual and fully automatic). Thus, the question still remains; does adaptive automation manage human performance and system performance tradeoffs more effectively than static intermediate levels?

Previous research efforts in the HAC project identified adaptive automation based on a task delegation interface as a good candidate for enhancing performance in the AdvSMR domain. Researchers conducted an experimental study to compare performance on a simplified process control simulation using the adaptive automation design to varying levels of automation under normal conditions and automation failure conditions.

A simplified process control simulation was developed to test human-automation interaction. The process control simulation was developed to enable the researchers to investigate the interaction and collaboration in a controlled way and that could also be generalized to the context of AdvSMRs. The process was designed to be simple enough for undergraduate psychology students to be able to operate with minimal training, but complex enough to allow for varying the level of automation. The process also needed to be difficult enough for researchers to detect differences in performance across the automation conditions. The process was also designed with proposed concepts of operations of AdvSMRs in mind. For example, the process requires the management of two separate product streams with different operational requirements. One of these product streams has dynamic requirements. This design is mimicking integrated energy systems and load-following, which are potential operational concepts for AdvSMRs.

The process simulation is designed with four possible levels of automation:

1. Fully manual: The operator handles all tasks manually.
2. Intermediate: The automation monitors, generates responses, and presents possible actions to the operators, and the operator chooses the actions to be taken, the automation then carries out the chosen action.
3. Fully Automatic: Automation handles all tasks, Operator task is to monitor and take manual control if necessary.
4. Adaptable: The process was divided into four high level tasks. The operator can choose to delegate any (or all) of the four tasks to the automation or perform the tasks manually.

Participants were 101 undergraduate psychology students at Idaho State University. There were 35 males and 66 females. The average age was 23years. The participants were given course credit for their participation.

The main independent variable was the automation condition, which was manipulated between participants (each participant executed both scenarios using a single level of automation). The other independent variable was the scenario type (introduced automation fault and no fault). This was manipulated within participants where each participant did one of each scenario type. The automation fault was the same for each participant. During the fault scenario, one pump fails to turn on automatically when a valve is opened. The participant had to manually control the pump in order to recover from the fault and keep the parameter in range. The participants were trained on how to do so before starting the scenario. The fault was injected at the same time for each participant.

The researchers applied the results of the performance measure battery development (Oxstrand et al., 2013c) to design the specific metrics for the study. The researcher measured system performance by assessing the degree to which the parameters are kept within range. They adapted the discrepancy score metric reported in Ha and Seong (2009). In addition, the researchers measured the total time that any one of the parameters was out of range.

Situation awareness (SA) was assessed by a freeze probe. The SA questionnaire was based on the Situation Awareness Control Room Inventory technique (SACRI) (Hogg et al., 1995). The participants were asked about each of the parameters they were required to monitor. Workload was assessed by NASA Task Load Index (NASA TLX) (Hart & Staveland, 1988). The NASA TLX is a subjective workload questionnaire that is administered after each task.

Automation fault management was assessed by determining if the participant; 1) switched the pump to manual, and 2) returned the parameter associated with the fault to acceptable range before the scenario ended. If both conditions were satisfied, then it was counted as a success. If not then it was counted as a failure.

It was concluded that the results from the experimental study do not confirm the common claim that adaptable automation is an effective method to manage human performance and system performance tradeoffs associated with increasing automation, but do not necessarily refute it. The results indicate that performance using adaptable automation is similar to that of using intermediate automation. This does not support the claim that adaptable automation is a better solution than intermediate levels of automation, which is often stated by other researchers.

However, it is also important to note that there were other results from the study that were not consistent with previous literature. This inconsistency could be due to some of the limitations present in the experimental study. Some of the identified limitations were:

- The study was conducted with University students who received minimal training.
- The process control simulation was not sufficiently complex to design a truly hierarchical abstraction scheme needed to detect differences between adaptable automation and intermediate levels of automation.
- The definition for intermediate level of automation used in the study does not represent static automation.

In conclusion, the results of the experimental study indicates that although many researchers have suggested adaptive automation may enable higher levels of automation without introducing human performance costs, further research is needed to determine if it is truly superior to intermediate levels of automation. Experimental research efforts in the HAC project will investigate the effects of a modified process control simulation (with modification to the automation schemes to more accurately reflect the schemes used in other research) on expert performers. The researchers will extensively train a group of participants on operation of the simulation in order to eliminate the issues related to novice participants controlling the process.

Empirical and analytical research will be conducted as needed to investigate the affect of automation on performance, efficiency, and safety. The design of human-system interaction to best support operation of automated systems and multiple product streams will be investigated through process simulator studies. As additional studies are planned and conducted the SA measures will be refined and the automation design will be revised based on insight from the first experimental study.

The results from the empirical and analytical studies will be used to refine the HAC model, which was initially developed by the researchers 2013. The refined model will be used to identify design requirements needed to ensure safe and effective operation. Based on the requirements a guidance document for development and evaluation of HAC design will be developed. In order to make the guidance document more practical to use, the researchers will develop a software tool for design and evaluation of Human-Automation Collaboration.

ACKNOWLEDMENTS

The authors would like to express special gratitude to the following people for their collaboration and support of this research effort. Dr. Roger Lew at University of Idaho and Aaron Bly at the Idaho National Laboratory for the process simulation and its user interface development. Dr. Maria Wong, professor in Developmental Psychology at Idaho State University for her support of the experimental study and Jennifer McDonald at Idaho State University for running participants on the INL researchers' behalf. Gordon Bower, statistician at the Idaho National Laboratory, for conducting the data analysis. Dr. John O'Hara at Brookhaven National Laboratory for providing helpful comments on draft versions of the report.

The research team also wants to extend its gratitude to all the participants and others involved in the successful execution of the empirical study.

CONTENTS

FIGURES

TABLES

ACRONYMS

AdvSMR	advanced small modular reactor
ANOVA	analysis of variance
DOE	Department of Energy
HAC	human-automation collaboration
HSI	human-system interface
LOA	level of automation
NASA	The National Aeronautics and Space Administration
NASA TLX	NASA Task Load Index
NRC	U.S. Nuclear Regulatory Commission
SA	situational awareness
SACRI	situation awareness control room inventory

Effects of Levels of Automation for Advanced Small Modular Reactors: Impacts on Performance, Workload, and Situation Awareness

1. INTRODUCTION

The Human-Automation Collaboration (HAC) research effort is a part of the Department of Energy (DOE) sponsored Advanced Small Modular Reactor (AdvSMR) program conducted at Idaho National Laboratory (INL). The DOE AdvSMR program focuses on plant design and management, reduction of capital costs as well as plant operations and maintenance costs, and factory production costs benefits.

It is assumed that AdvSMR designs will employ some level of advanced digital instrumentation, controls, and human-machine interfaces to optimize the use of instruments and controls, increase plant availability, and reduce staffing requirements. In other words, AdvSMRs will employ technology that is significantly more advanced than the analog systems in the existing light water reactor fleet. It is expected that AdvSMRs will rely on automation to a greater extent than the current nuclear power plant fleet. As always when moving towards more advanced technology, there are a number of concerns about how these technologies will affect human performance and the overall safety of the plant.

The HAC research project investigates how to best design the collaboration between the operators and the automated systems in a manner that has the greatest positive impact on overall plant performance and reliability. There is a distinction between the HAC and human-system interaction (also known as human-computer interaction). HAC is defined as how the operator and the automation work as a team to ensure effective and safe plant operation. At this level, the focus is to understand the effects of various characteristics of automation (such as its reliability, processes, and modes) on an operator's use and their awareness of plant conditions. The overall goal of the HAC research effort is to develop design guidance that supports optimal interaction between humans and automated systems. This guidance will be incorporated into an interactive tool for design and evaluation of HAC. The tool is intended to be used by engineers in the design process of AdvSMR systems.

The first step to develop design guidance, and thus reach the goal of the research effort, is to identify and characterize the current state-of-practice of HAC. A literature review aimed to identify human factors and automation factors that influence HAC was conducted. The team identified and analyzed key contributing factors, such as levels of automation, reliability, cognitive functions that the automation is responsible for, and how those aspects of automation affect operator performance and system performance. The majority of the human factors' literature defines performance problems associated with certain HAC configurations, but does not necessarily describe circumstances that lead to successful collaboration. Hence, there are knowledge gaps needed to be addressed to provide adequate knowledge and understanding of the specific HAC needs for AdvSMRs and other highly complex and highly automated systems.

1.1 Background

Previously conducted activities in the HAC research project aimed to identify HAC related knowledge gaps that needed to be investigated in order to develop the design guidance. The researchers studied the current state-of-practice of HAC in the human factors and automation research fields, in the AdvSMR industry, and in similar industries. Based on the results from these activities a model of human-automation collaboration was developed. The HAC model identified several important features of automation that influence human-automation interaction including; level of automation, cognitive function automation replaces (e.g., detecting anomalies or taking control actions), adaptability, reliability, process of automation, and automation modes. Using the model, the researchers were also able to identify several important human performance and HAC performance outcomes, such as system performance, situation awareness, workload, trust in automation, and team situation awareness. These outcomes should be used to evaluate HAC design.

The process by which the researchers plan to develop the HAC design guidance can be summarized as: 1) Identify HAC research needs, 2) Identify functions and tasks for AdvSMR, 3) Prioritize research needs based on functions and tasks identified, 4) Conduct HAC research to fulfill the needs, and 5) Synthesize results from research to develop design guidance for HAC in AdvSMR. The investigation of the current state-of-practice of HAC in the human factors and automation research fields, in the AdvSMR industry, and in similar industries as well as the development of the HAC Model supported the identification of high level research needs. The areas where additional research is needed to close knowledge gaps are:

- Models of team work
- Initiators and triggering conditions for adaptive automation
- HAC performance measures
- Impact of highly automated advanced small modular reactors on operator awareness
- Regaining/reacquisition of operator awareness
- Effects of HAC characteristics on operator's use of automation

The high level research needs were used to identify and prioritize the research gaps to address first. Three analytical studies were identified and conducted: 1) Model of teamwork, 2) Standardized HAC performance measurement battery, and 3) Initiators and triggering conditions for adaptive automation. For more detailed description of the previously conducted research and the path towards identifying the current state-of-practice and knowledge gaps, see Oxstrand et al., 2013a, 2013b, and 2013c.

The analytical studies, described in Oxstrand et al. 2013c, addressed some of the research gaps, but many important research issues remain to be addressed before guidance on human automation collaboration can be developed. For example, the analytical study on initiators and triggering conditions for adaptive automation identified a potentially fruitful research path that would also help to answer some of the general research questions. Specifically, "How do we get the most benefits out of using automation while keeping the operators engaged in the process at the same time?"

The question mentioned above is one of the questions the researchers aimed to address in the experimental study they conducted in 2014. This report describes the research activities leading up to the study, the method used in the study, and the result of the study. Hence, the report addresses the milestone "M3SR-14IN1301057: Conclude experimental studies using simulated process systems". The milestone marks the completion of a first experimental study using simulated process systems investigating knowledge gaps identified during FY13's studies related to Human-Automation Collaboration for complex and highly automated systems. The following sections provide a summary of the main areas covered in the literature review and the knowledge gaps identified. Some of these gaps were selected for further investigation in the experimental study, which is described in Section 2– Method and Section 3 – Results.

1.2 Current State-of-Knowledge

AdvSMRs are expected to use higher degrees of automation than the existing fleet of light water reactors in the United States. Consequently, the role of the human operator in these systems will shift from a direct controller, to the role of monitoring and supervising automated activities. Reliable automation typically enhances overall performance of the human-system; however it often also has the negative consequence of reducing human monitoring and situation awareness (Endsley, 1995, 1996, 1997; Endsley & Kaber, 1999; Kaber & Endsley, 1997, 2004; Parasuraman et al., 2000; Sheridan, 2002;

Wickens & Hollands, 2000; Wright & Kaber, 2005). In systems where the human operator has the ultimate responsibility to maintain a safe system state (such as in process control systems and nuclear power plants), the operator must maintain a level of situation awareness that will allow him to recover in case the automation fails. Therefore, the overall performance of the human-system needs to be characterized based on combination of system performance and human performance (including situation awareness (SA) and automation failure recovery).

In fulfilling their responsibilities, agents (automation or human) perform primary tasks (i.e., cognitive functions). These cognitive functions include activities such as monitoring plant parameters, executing procedures, starting pumps, and aligning valves. Cognitive functions have several common elements, whether the agent is automation or a human operator: monitoring and detection, situation assessment, response planning, and response implementation. When the agents are human, they also must perform interface management tasks such as navigating or accessing information at workstations and arranging various pieces of information on the screen. These secondary tasks are important to consider because they create workload and may divert attention away from primary tasks and make them difficult to perform (O'Hara & Brown, 2002). The proposed a set of cognitive functions that consider all of the factors by O'Hara et al. (2010) are:

Monitoring and detection refer to the activities involved in extracting information from the environment. Monitoring is checking the state of the plant to determine whether it is operating correctly, including checking parameters indicated on the control panels, monitoring those displayed on a computer screen, obtaining verbal reports from other personnel, and sending operators to areas of the plant to check on equipment. An alarm system is an example of automation applied to monitoring and detection.

Situation assessment is evaluating current conditions to assure their acceptability or determining the underlying causes of any abnormalities (e.g., diagnosis). An example of automation applied to a situation assessment is a disturbance analysis system and other computerized operator-support systems.

Response planning refers to deciding on or choosing a course of action to address the current situation. In a nuclear power plant, procedures usually aid response planning. An example of automation applied to response planning is a computer-based procedure system.

Response implementation is undertaking the actions specified by response planning. They include selecting a control, providing control input, and monitoring the responses of the system and process. An example of automation applied to implementing a response is an automatic safety system such as soft controls.

Interface management encompasses activities such as navigating or accessing information at workstations and arranging various pieces of information on the screen. An example of applying automation to interface management is automatic identification of a display appropriate to the ongoing situation (e.g., identification of an emergency-procedure display upon detecting any of the procedures entry conditions). In this context, human-system interface (HSI) notifies the operator of the availability of the display (i.e., by a blinking icon at the bottom of the screen), rather than disrupting the operator's ongoing activity by obtrusively showing the display.

Researchers exploring human-automation interaction typically describe automation based on the level and type of automation being employed. There are numerous taxonomies of levels of automation (LOA) (e.g., Parasuraman et al. 2000; Billings, 1991; Endsley and Kaber, 1999, and O'Hara et al., 2010), and each researcher defines level of automation slightly differently. Most researchers acknowledge that level of automation refers to the amount of automation that is employed in a particular task. This typically varies from fully manual to fully automatic with intermediate levels representing shared responsibility between the human and automation. Where researchers tend to disagree is whether the type of activity that is being automated should be included in characterizations of level of automation. Some researchers include cognitive function (i.e., the agent's primary function as discussed above) as an additional dimension in the characterizations of level of automation (Endsley & Kaber, 1999; Parasuraman et al., 2000). Like the functions listed by O'Hara (2010), Parasuramn et al (2000) describe a taxonomy that lists the cognitive functions of information acquisition, information analysis decision and action selection, and action implementation. Each of these functions could be delegated to the human or the automation, or some combination thereof.

Endsley and Kaber (1999) proposed an LOA taxonomy intended to be generic enough to have applicability to a wide range of cognitive and physical tasks that require real-time control in a number of industries. In developing this model, Endsley and Kaber assigned the monitoring, generating, selecting, and implementing cognitive functions to the human operator, automation, or a combination of the two to develop the 10 levels of automation shown in Table 1. Note that as the LOA increases, automation takes over progressively more of each cognitive function.

Table 1. Level of Automation Taxonomy (Endsley & Kaber, 1999).

	Roles			
Levels of Automation	Monitoring	Generating	Selecting	Implementing
(1) Manual control	Human	Human	Human	Human
(2) Action support	Human/ computer	Human	Human	Human/ computer
(3) Batch processing	Human/ computer	Human	Human	Computer
(4) Shared control	Human/ computer	Human/ computer	Human	Human/ computer
(5) Decision support	Human/ computer	Human/ computer	Human	Computer
(6) Blended decision making	Human/ computer	Human/ computer	Human/ computer	Computer
(7) Rigid system	Human/ computer	Computer	Human	Computer
(8) Automated decision making	Human/ computer	Human/ computer	Computer	Computer
(9) Supervisory control	Human/ computer	Computer	Computer	Computer
(10) Full automation	Computer	Computer	Computer	Computer

The majority of the taxonomies are developed in the context of other industries than nuclear power generation. In an effort to bridge this gap, O'Hara et al. (2010) adapted the existing taxonomies to account for the types of automation used in the nuclear industry. The Table 2 below describes the O'Hara et al. taxonomy.

Table 2. Preliminary Levels of Automation for Nuclear Power Plant Applications (O'Hara et al., 2010).

Level	Automation Functions	Human Functions
1. Manual Operation	No automation	Operators manually perform all functions and tasks
2. Shared Operation	Automatic performance of some functions or tasks	Manual performance of some functions/tasks
3. Operation by Consent	Automatic performance when directed by operators to do so,	Operators monitor closely, approve actions, and may intervene to provide

	under close monitoring and supervision	supervisory commands that automation follows
4. Operation by Exception	Essentially autonomous operation unless specific situations or circumstances are encountered	Operators must approve critical decisions and may intervene
5. Autonomous Operation	Fully autonomous operation; system or function cannot normally be disabled, but may be started manually	Operators monitor performance and perform backup if necessary, feasible, and permitted

There are many different definitions of cognitive functions depending on the researcher and the model of cognition used. The HAC research team studied how cognitive functions were incorporated in to the taxonomies for LOA. Even though the O'Hara et al. LOA taxonomy was developed for the nuclear industry it was developed for near-term applications in advanced plants. While those plants are likely to use more automation than an existing light water reactor plant, they may not use automation to the degree that is anticipated for AdvSMRs. Therefore, research is needed to identify potential revisions of existing taxonomies, including the O'Hara et al., to ensure a more accurate reflection of the AdvSMR context.

The process of assessing the current situation, i.e., to produce a mental representation which reflects the person's understanding of what is going on is generally referred to as "situation awareness" (SA) (Endsley, 1995). In other words, SA is the evaluation of current conditions to determine that they are acceptable or to determine the underlying cause of abnormalities when they occur. Operators actively try to construct a coherent and logical explanation to account for their observations. Thus, accurate SA is critical to taking proper human action.

There are several models of SA: sensemaking (Klein, Moon, & Hoffman, 2006; Klein et al., 2007), perceptual cycle theory of SA (Smith & Hancock, 1995; Adams, Tenney, & Pew, 1995), and the functional model of orienting activity (Bedny & Meister, 1999;; Bedny, Karwowski, & Jeng, 2004). The model that has received the most empirical investigation and support is Endsley's model, which was developed through work in the aviation industry and has been applied in numerous additional industries, such as air traffic control, military command and control, and power plant operations. With increased use of automation across many fields, researchers have observed persistent findings related to operator awareness of what is happening in the plant or process and awareness of what automation is doing. These findings show that automation does not necessarily improve operator performance (Endsley, 1996, 1997; Endsley & Kaber, 1999; Endsley & Kiris, 1995; Jou et al., 2009; Kaber & Endsley, 2004; Lin et al., 2009, 2010a, 2010b; van de Merwe et al., 2012).

Endsley's SA model is an information-processing model that documents the product of situation assessment in three levels. Level 1 involves perception of the status, attributes, dynamics, and other relevant aspects of elements in the environment, such as information and objects (Endsley, 1995). Level 1 SA simply involves perception of the relevant elements; higher-level comprehension does not occur until Level 2. Level 2 SA involves combining, integrating, and interpreting the information perceived in the Level 1 SA into an understanding of the current situation (Endsley, 1995, 2000). Level 3 SA involves projecting the current situation into the future to mentally forecast the future state of the situation given currently available information (Endsley, 1995, 2000), enabling the person to project and anticipate how the situation is going to evolve. Each level builds on the previous level to create understanding of the situation and errors made at an earlier level impair subsequent levels of awareness.

One of the most significant findings related to effects of automation on operator performance and SA is referred to as the out-of-the-loop phenomenon or out-of-the-loop performance problem (Endsley, 1995, 1996, 1997; Endsley & Kaber, 1999; Kaber & Endsley, 1997, 2004; Parasuraman et al., 2000; Sheridan, 2002; Wickens & Hollands, 2000; Wright & Kaber, 2005). When operators are out of the loop, they are not aware of the state of automation or the system parameters (Endsley, 1996). This contributes to operators failing or being slow to detect that a problem has occurred in the system that necessitates their intervention. Furthermore, the out-of-the-loop phenomenon means that once operators have detected a problem, they need additional time to determine and adequately understand the state of the system (in other words, restore their SA of the system and automation) in order to take appropriate action.

The literature points to many issues with high levels of automation for many t functions related to keeping the operator in the loop. However, AdvSMRs are likely to be designed to employ a higher level of automation than current nuclear power plants. Therefore, extensive research needs to be conducted to investigate how to enable higher levels of automations, while keeping the operator actively engaged in the operation of the plant.

Studies have shown a decrease in SA when the operator's workload is either too high or too low. Hence, operator workload seems to be a major consideration for designing automation. Many believe that automation can improve operator SA by reducing operator workload. However, some studies indicate that automation does not always decrease operator workload. Instead, automation may actually increase workload depending on how it is designed and the cognitive functions automation is responsible for (Lin et al., 2010a). In addition, automation may reduce operator workload at the expense of SA (van de Merwe et al., 2012, Miller & Parasuraman, 2007).

It is important to recognize that automation is not perfect. In fact, all engineered systems have less than perfect reliability. Automation's reliability is defined as how well automation accomplishes its task and reliability is one of the characteristics of properly functioning automation. For simple automated systems the measure of reliability is quite straight forward. However, when automation's tasks are complex it becomes more difficult to define measures of reliability. In addition, automation's reliability may differ across different context of use and/or modes of operation.

One factor that has a large impact on the reliability of the automation is degradation. Problems arising in the instrumentation and control infrastructure can lead to degradation or failure of any aspect of automation (O'Hara et al., 2010). In a degraded condition, automation will continue to operate, but the loss of functionality may lead to incorrect performance. In a failed condition, automation does not perform at all. Degradation of automation can lead to two types of problems for the operator: 1) automation does not do what it is supposed to do when it should do it, and 2) automation does something that it is not supposed to do at all. With each of these types of problems, operators must detect the degraded or failed automation, determine the proper actions to take (via assessing the situation and planning a response), and/or transition to back-up systems or operations. Each of these human actions is potentially subject to human performance issues.

Highly reliable automation is assumed to improve task performance, but not the detection of automation failures. In fact, the higher the reliability of automation, the less likely it is that the operator will recognize when it fails (Dixon & Wickens, 2006; Wickens et al., 2010). This is one of the ironies of automation identified by Bainbridge in 1983. When operators know the actual reliability of the system, they can make use of that knowledge to adjust their use of automation. Based on the reviewed literature, the HAC researchers hypothesized the relationship as follows; As automation becomes less reliable, its support for task performance becomes less and performance declines. At some reliability threshold, automation's lack of reliability draws operator attention away from the task to automation monitoring and task performance suffers. At an even lower threshold, operators abandon automation altogether and perform the task manually.

Many of the factors that can lead to SA problems can be traced directly to the way automation is designed. Therefore, it is essential to minimize these problems during system design and optimize the benefits of automation without sacrificing operator SA (Endsley, 1995). In addition, the human-system interface (HSI) can both affect objective performance and has an important impact on operator SA. Dehais, Causse, and Tremblay (2011) studied how "cognitive countermeasures" presented through the

HSI can help mitigate cognitive errors that the operator might commit when automation behaves in an unexpected manner. The study demonstrated that the dynamic presentation of visual cues in the HSI were effective at getting the operator's attention without causing over fixation on the visual cue and conveyed to the participant what aspects of the situation had changed and affected the collaboration task such that the operator was no longer surprised by the change in the automation's behavior. The results of this study reinforce the ideas that the HSI is an important mediator between humans and automation and that it is important that the HSI communicate the right information to the operator at the right time.

HSI is also key to whether operators detect the automation degradation or failure. Operator SA can minimize the "routine-failure tradeoff", i.e., failures are better handled when operators have good SA regarding the system and automation state (Wickens et al., 2010). However, even if operators do monitor automation, the design of the operator's interface with the automation may not support monitoring needs and, may be misleading. Willems and Heiney (2002) stated that "As errors involving automation tend to be more cataclysmic and costly, the human interface has become more important than ever" (p. 3). The HSIs typically provide insufficient information about automation's goals, current activities, and performance (Lee & See, 2004; Liu, Nakata & Furuta, 2004; Parasuraman & Riley, 1997; Rook & McDonnell, 1993; Roth et al., 2004).

Hence, one of the important factors that contribute to successful HAC is whether the HSI is successful in providing adequate information to keep the operator in the loop. In other words, the HSI must successfully inform the operators about what the automation is doing, what its reliability of the automation is, etc. The content and format of the HSI design impacts operator trust (Lee & See, 2004; Parasuraman & Riley, 1997). Trust in automation is closely related to the concept of perceived reliability of the automation. Parasuraman and Riley (1997) suggest that monitoring of automation is improved when its behavior can be determined easily using the HSIs, especially those that minimize attentional demands such as displays that integrate information and provide emergent features. They noted that there is evidence to indicate that automation failures were better detected with these types of displays. Conversely, operators are less likely to monitor automation when the HSI does not offer an easy means to do so. How to best design such HSI for highly complex systems, such as for AdvSMRs, to ensure efficient collaboration between the operators and the automation has not been adequately explored at this time. In addition, the increased complexity could affect operator use in more and more complex ways. Understanding how more sophisticated automation affects operator use and HAC system performance is important to the AdvSMR designer. Therefore, research needs to be conducted to identify the AdvSMR specific HAC characteristics and investigate their affect on the use of automation.

In summary, the investigation of the current state-of-knowledge related to HAC concludes that even though there is a lot of research conducted in the fields of human factors and automation that is applicable to the AdvSMR context there are still knowledge gaps that needs to be addressed. One of these gaps is how to best enable higher levels of automation, which is something believed to be of high importance to the field of AdvSMRs.

1.3 Adaptable Automation as a Way to Enable Higher Levels of Automation

Generally, higher LOAs result in better system performance. Unless there is an automation failure, that is (Wickens, et al. 2010; Smith & Jameison, 2012; Manzay, Reichenbach, & Onnasch, 2008; Onnasch et al., 2013). In the case of automation failure, higher LOAs often result in the human operator failing to recover due to a reduction in SA during the automated period. Lower LOAs result in better human monitoring of the system, but fail to take advantage of the enhanced performance that automation can provide during routine performance. Many researchers have described the relationship between system performance and fault recovery performance as a fundamental tradeoff with level of automation (Wickens, et al. 2010; Smith & Jameison, 2012; Manzay, Reichenbach, & Onnasch, 2008; Onnasch et al., 2013).

For tasks that require human-automation collaboration (i.e., automation is desirable, but the operator must be involved in monitoring and some manual control), researchers have recommended using intermediate levels of automation to manage these tradeoffs. Unfortunately, using intermediate levels of automation doesn't necessarily optimize performance of either the human or the automation (Onnasch et al. 2013). Intermediate levels of automation represent a "least-bad-of both worlds" approach to managing system performance and human monitoring of automation. In intermediate automation, system performance is typically worse than fully automated, but better than manual. Similarly, human monitoring and fault recovery under condition of intermediate automation are typically superior to fully automated conditions, but not as good as manual.

Several researchers have proposed that adaptive automation is a better way to manage and perhaps eliminate these tradeoffs. Adaptive automation can be simply described as dynamic allocation of functions. One potential benefit of adaptive automation is that it can result in more balanced workload because tasks can be assigned to automation based on the conditions encountered during operation. Another potential advantage is that tasks can be assigned to the automation when the human operator's performance or situation awareness declines, but that the human can remain engaged in the process by manually controlling (or partially controlling it) at other times.

Many studies have demonstrated the effectiveness of adaptive automation in enhancing performance compared to fully manual systems (Kaber, Wright, Prinzel & Clamann, 2005; Kaber, Perry, Segall, McClernon, & Prinzel, 2006; Kaber, & Riley, 1999). A few studies have also demonstrated that automation fault recovery with adaptive automation is superior to fully automated systems (Calhoun, Ruff, Spriggs, & Murray, 2012). Further, studies have shown that situation awareness is better with adaptive automation than with fully automated systems (Parasuraman, Cosenzo, & De Visser, 2009). Essentially, the promise of adaptive automation is that it provides system performance closer to fully automated designs, but without the negative human performance consequences.

There are still many unanswered questions regarding the applicability and superiority of adaptive automation. For example, most of the empirical studies on adaptive automation demonstrate that it enhances performance compared to static automation at the extremes (fully manual and fully automatic), few studies compare adaptive automation to intermediate levels (although a few do, e.g., Parasuraman, Cosenzo, and De Visser, 2009 and Sauer, Nickel, and Wastell, 2013). In addition most studies investigate adaptive automation's effect on system performance or the effect on human performance, but few look at both. Thus, the question still remains; does adaptive automation manage human performance and system performance tradeoffs more effectively than static intermediate levels?

Further, there is the question of how to initiate a shift in automation. As a part of one of the three analytical studies mentioned above, the HAC researchers conducted a literature review aimed at identifying the effectiveness of triggering conditions for adaptive automation. Empirical results of studies the effect of various triggering conditions on performance and situation awareness were reviewed. The researchers compared the following triggering conditions; operator-initiated adaptive automation, performance based triggering conditions, and model based triggering conditions, physiological-based triggering conditions, and critical event triggering conditions. It is important to point out that the researchers have made a distinction between *adaptive* automation and *adaptable* automation. Adaptive automation is typically described as system-initiated shifts in automation, while adaptable automation is typically described as operator-initiated shifts in automation. In both cases, the level of automation can be shifted to support operators. All of the triggering conditions listed besides operator initiated triggering are considered system initiated shifts.

The review revealed that there are many tradeoffs associated with the various triggering mechanisms. System-initiated shifts in automation may surprise operators or reduce their SA, but operator-initiated shifts may increase operator workload. Further operators may choose to use lower levels of automation when the system is adaptable rather than adaptive, resulting in reduced system performance. Finally, some adaptive automation designs, such as those based on physiological assessment of operator workload and SA may consistently enhance performance compared to static automation, but may not be feasible to implement in a control room environment (Le Blanc & Oxstrand 2104).

Based on the results of the analytical study on triggering conditions, the HAC researchers concluded that adaptable automation using a task delegation interface may be the best way to approach automation in the AdvSMR context. In this adaptive automation scheme, the tasks are broken down into a hierarchical abstraction scheme with higher level goals as the highest abstraction and component-level manipulation as the lowest level of abstraction. The automation can be shifted up and down the abstraction hierarchy based on the current conditions or operator's needs. Many studies have investigated the implementation of this type of adaptive using an operator-controlled task delegation interfaces (note that because the shifts in automation are controlled by the operator, it is adaptable rather than adaptive in this case). Adaptable task delegation interfaces tend to enhance performance compared to static automation (Parasuraman et al., 2005; Miller et al., 2011; Shaw et al., 2010), however it is still unclear whether this type of adaptable automation is superior to intermediate automation.

Researchers hypothesized that an adaptable task delegation interface could be used to enable high levels of automation (and hence system performance) without sacrificing human performance on a simplified process control simulation. Section 2 describes the experimental study in detail.

2. METHOD

In order to determine whether an adaptable task delegation interface enhances performance compared to intermediate levels of automation, the researchers conducted an experimental study to using simplified process control simulation. The researchers compared human performance and system performance using the adaptive automation design to varying levels of automation (manual, intermediate and fully automatic) under normal conditions and automation failure conditions. As discussed in the previous section, some of the human performance issues with automation are reduced SA and fault management. Fault detection and management were therefore incorporated in the experimental study. The researchers investigated the effect of automation condition on the participants' ability to detect and manage the failing automation; the researcher also measured SA, system performance, and subjective workload.

The goal of the study was to demonstrate that the adaptable automation is better than intermediate automation in managing the human performance costs of increased automation. If adaptable automation is indeed better than intermediate automation at managing the tradeoffs, then system performance using the adaptable automation should be better than that of manual and intermediate automation and closer to fully automatic. In addition human performance (i.e., SA, and fault management) using the adaptable automation should be better than fully automatic and intermediate, and closer to manual.

2.1 Simulation Development

With the intention to investigate human-automation interaction and collaboration in a controlled way that could also be generalized to the context of AdvSMRs, the research team developed a simplified process control simulation that could be used to test human-automation interaction.

The process was designed to be simple enough for undergraduate psychology students to be able to operate with minimal training, but complex enough to allow for varying the level of automation. The process also needed to be difficult enough for researchers to detect differences in performance across the automation conditions (i.e., it needed to be difficult enough that not all participants performed perfectly). In order to ensure that the task was difficult enough (i.e., the researchers wanted to avoid a ceiling effect in which all participants displayed error-free performance), the simulation was designed so that it required the participants to actively maintain the process at all times. In manual mode, the operators had to take many actions per minute in order to keep the process within the prescribed parameters. It was expected that in all conditions, the participants would not be able to perform perfectly.

The process was also designed with proposed concepts of operations of AdvSMRs in mind. For example, the process requires the management of two separate product streams (Mixing tanks A and B) with different operational requirements. This is analogous to the proposed concept of using the AdvSMR plant to supply both electricity generation and process heat. One of the product streams also has dynamic requirements, which is analogous to load-following, another potential operational concept for AdvSMRs. Even though the process control simulation incorporates these AdvSMR-like concepts of operation it is not intending to mimic an actual process found in a nuclear power plant.

2.1.1 Task Description

The goal of the process is to provide appropriate concentrations of a chemical to two different processes. The chemical must be mixed with heated water. The participant must manage 5 tanks and associated pumps, valve and heaters to control the process.

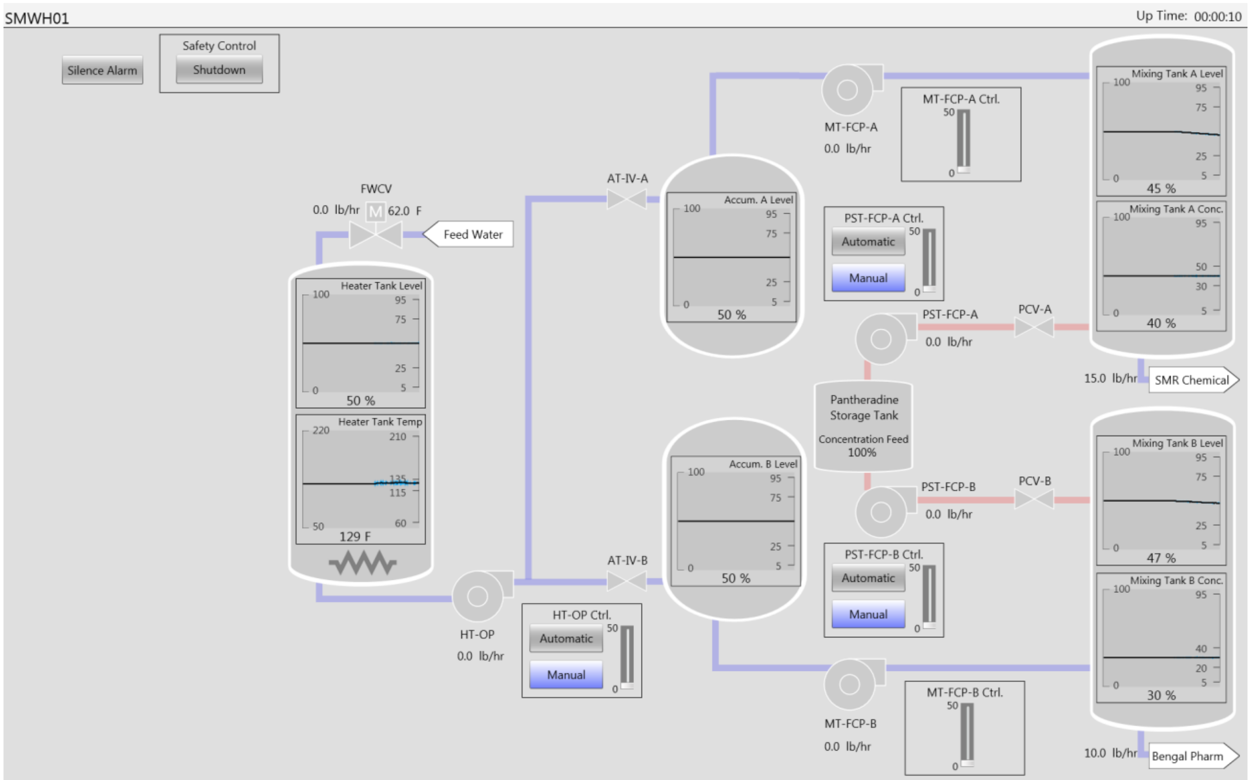

Figure 1. Screenshot From Simulation.

As indicated in Figure 1, the process starts with the heater tank. Feed water flows into the tank from the water treatment plant and the flow is controlled by the feed water control valve. Once in the tank, the temperature of the water is controlled by the heater. The water temperature needs to be kept within a certain range to ensure proper mixing later in the process. The heater tank also needs to be kept within a certain level range in order to support the rest of the process without risking overfilling the tank. The operator must monitor heater tank temperature and level and adjust the heater and valves as needed.

The heated water then flows into two accumulator tanks where it is stored until it is needed later in the process. Each tank has a control valve associated with it that controls the flow of heated water from the heater tank to the accumulator tank. The levels of both accumulator tanks need to be kept between within a certain range.

From the accumulator tanks, the water flows to the mixing tanks where it is mixed with a concentrated chemical. To manage the mixing tanks, the operator must control both the flow of chemical from the tank, and the flow of heated water from the accumulator tanks. The flow of chemical is controlled via a valve and a pump. The pump defaults to automatic control mode, which means that the pumps automatically turns on when you open the control valve and turns off when you close it. In manual mode, the pump needs to be turned on if you want water to flow from the tank when the valve is opened and vice versa. The flow of water is controlled via one pump that moves water from the accumulator tank to the mixing tank.

There are two mixing tanks. One feeds the main process, while the other feeds a different company's process. Each mixing tank needs to be kept within a certain level. In order to maintain the level, the operator must control the flow of both water and chemical into the mixing tank. The concentration of chemical needed in each tank depends on the manufacturing process the tank is supplying. The concentration in each mixing tank needs to be kept within a specific range in order to support the process. The mixing tanks have different demands depending on the manufacturing process they are feeding.

Each important process parameter is presented as a trend graph, see Figure 2. The margins that the process parameter needs to be kept in are displayed as tick marks on the right side of the trend display. The trend display turns bright yellow when the process parameter moves outside the required margins. An audible alert also sounds when the parameters goes outside of the require range. If the parameters get too far outside the required ranges, the trend display turns red. This means that the operator must take immediate action in order to maintain safety and the parameter back in range.

There are a total of ten process parameters, out of which eight are parameters that the participants were required to monitor and keep within range. These eight parameters are:

- Heater Tank Level
- Heater Tank Temperature
- Accumulator Tanks A and B Levels
- Mixing Tank A Level
- Mixing Tank A Concentration
- Mixing Tank B Level
- Mixing Tank B Concentration

The other two parameters are set in the scenario, but the participants are asked about them in the SA questionnaires. These two parameters are:

- Mixing Tank A outflow
- Mixing Tank B outflow

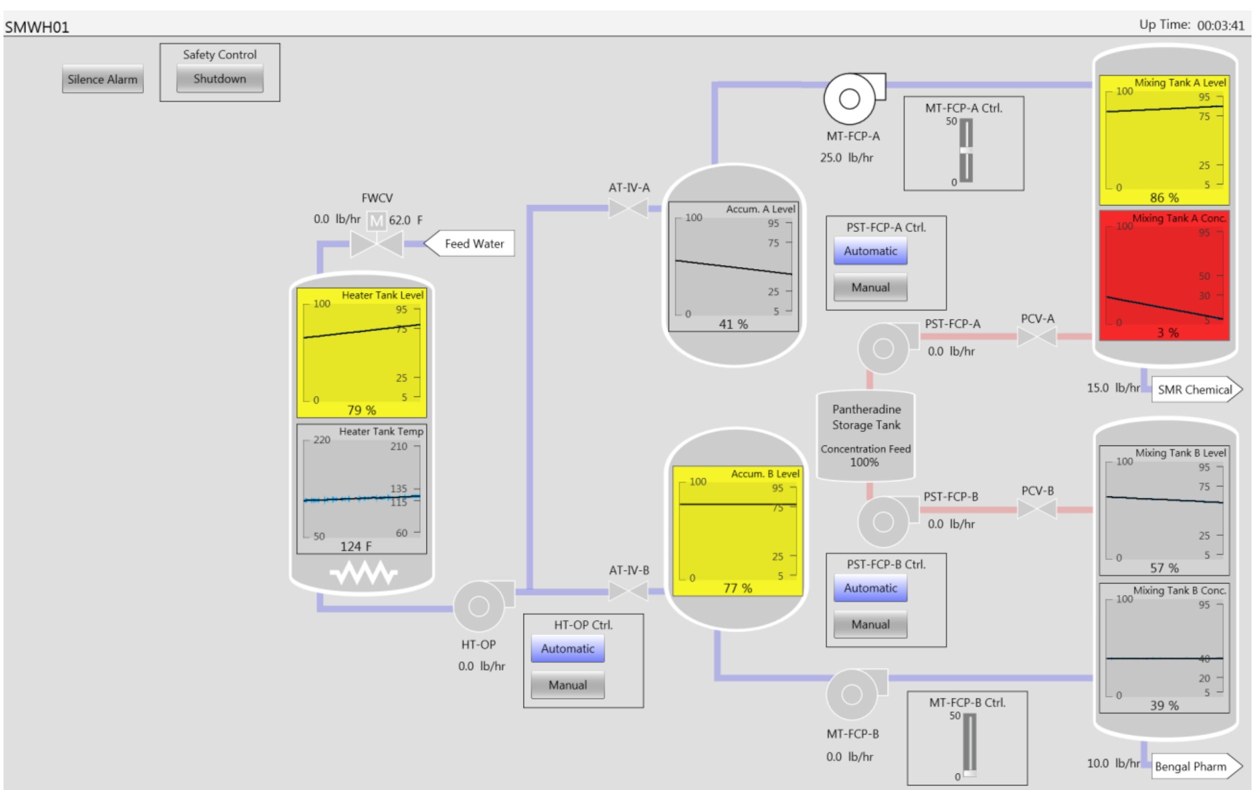

Figure 2. Screenshot of Yellow and Red Warnings.

2.1.2 Automation Descriptions

The process simulation is designed with four possible levels of automation: Fully Manual, Intermediate (decision support system), Adaptable (task delegation interface), and Fully Automatic. These levels of automation are described in the sections below.

2.1.2.1 Fully Manual

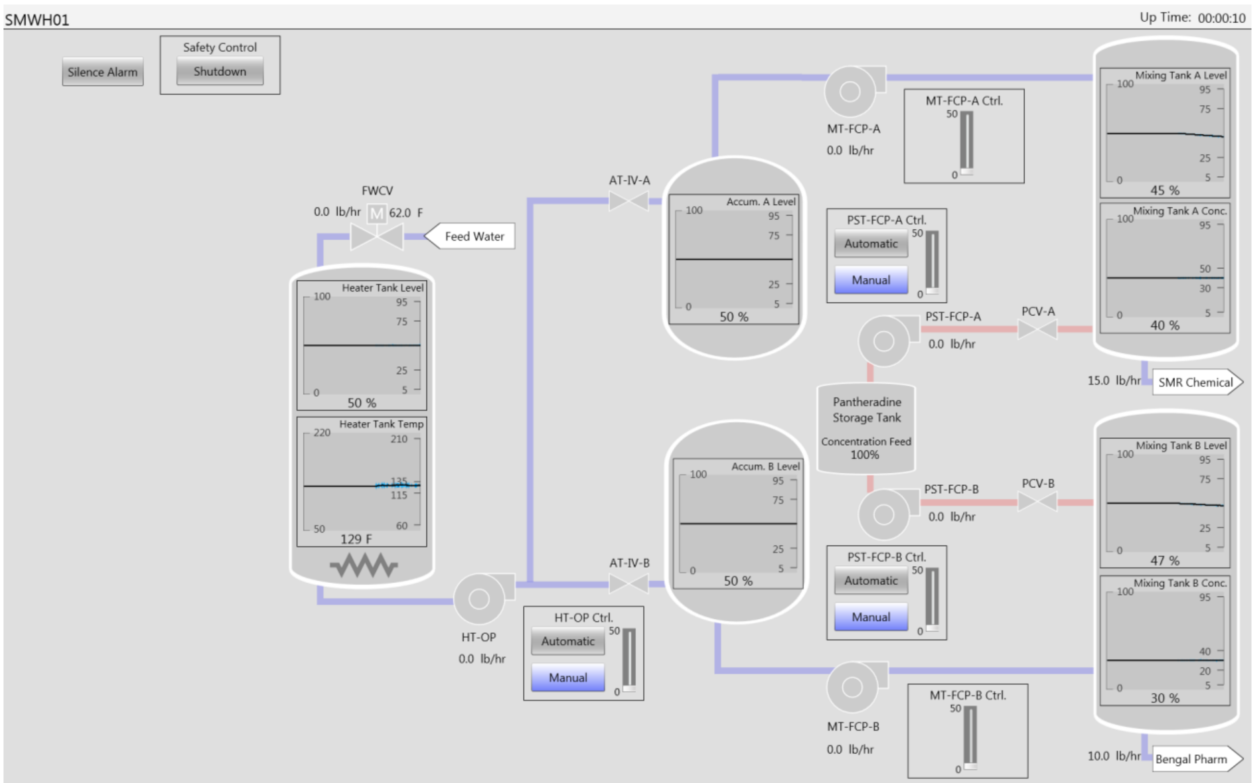

Figure 3. Screenshot From Simulation in the Manual Condition.

In the Fully Manual condition, the control of the entire process is carried out by the human. The human must monitor the levels, temperature and concentrations of all the tanks. The operator must open valves, turn on pumps, and turn on the heater, in order to main the appropriate process parameters. Figure 3 is a screenshot from the manual condition.

2.1.2.2 *Intermediate*

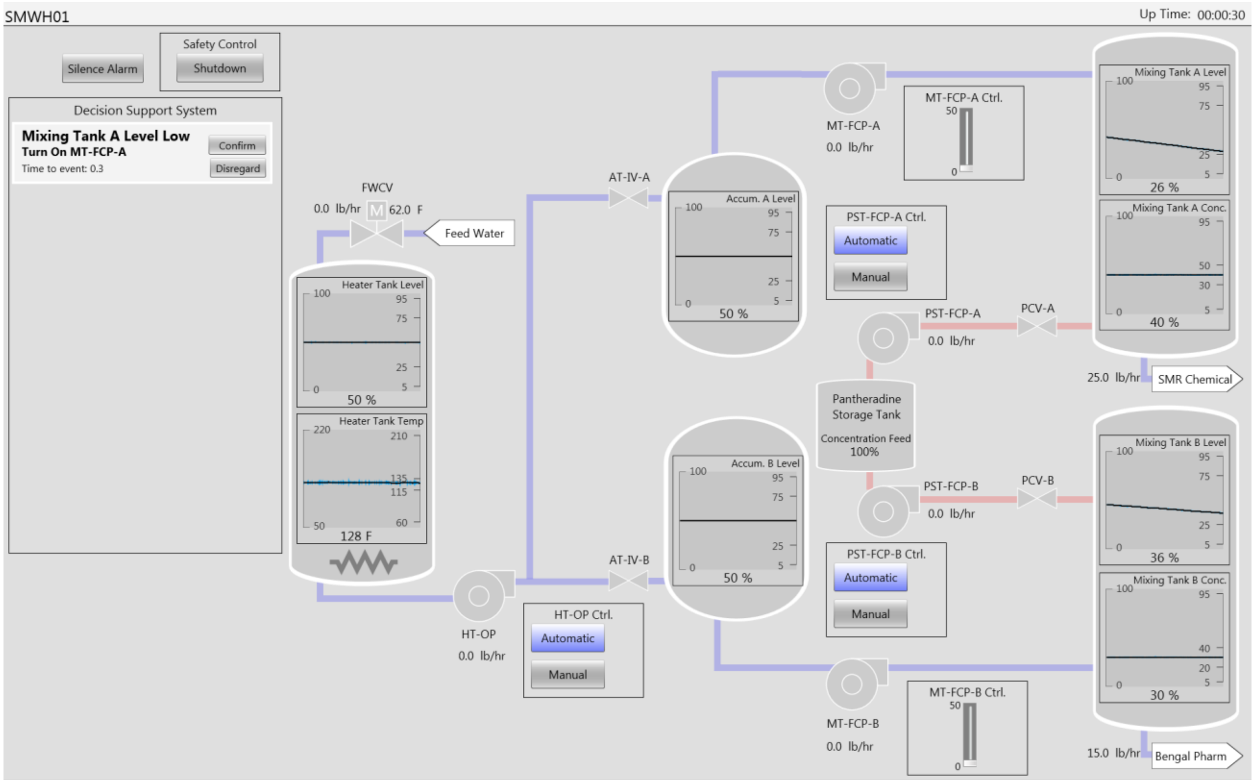

Figure 4. Screenshot From Simulation in the Intermediate Condition.

The definition of intermediate automation used in this study was taken directly from Endsley and Kaber's (1999) taxonomy of levels of automation, see Section 1.1 – Current State-of-Knowledge. The researchers used the fifth level of automation, Decision Support, which is described as:

"The computer generates a list of options that the human can select from; the operator may still generate his or her own options. The computer is responsible for implementing the chosen action. This LOA is common in many expert systems or decision support systems in which the operator may use or ignore the option guidance provided by the system."

In the Intermediate automation condition, the automation monitors the process and prompts the operator via the decision support system when a parameter is about to go out of range. The prompts appear in the panel on the left of the screen, see Figure 4. The prompt contains a description of the conditions, a suggested course of action, and an estimate of the time until the parameter is out of range. The operator can click a "confirm" button on the prompt, and the automation will carry out the suggested action. The operator can also dismiss the prompt and carry out the suggested action manually or ignore the prompt. The rest of the interface is identical to the Fully Manual condition.

2.1.2.3 *Adaptable*

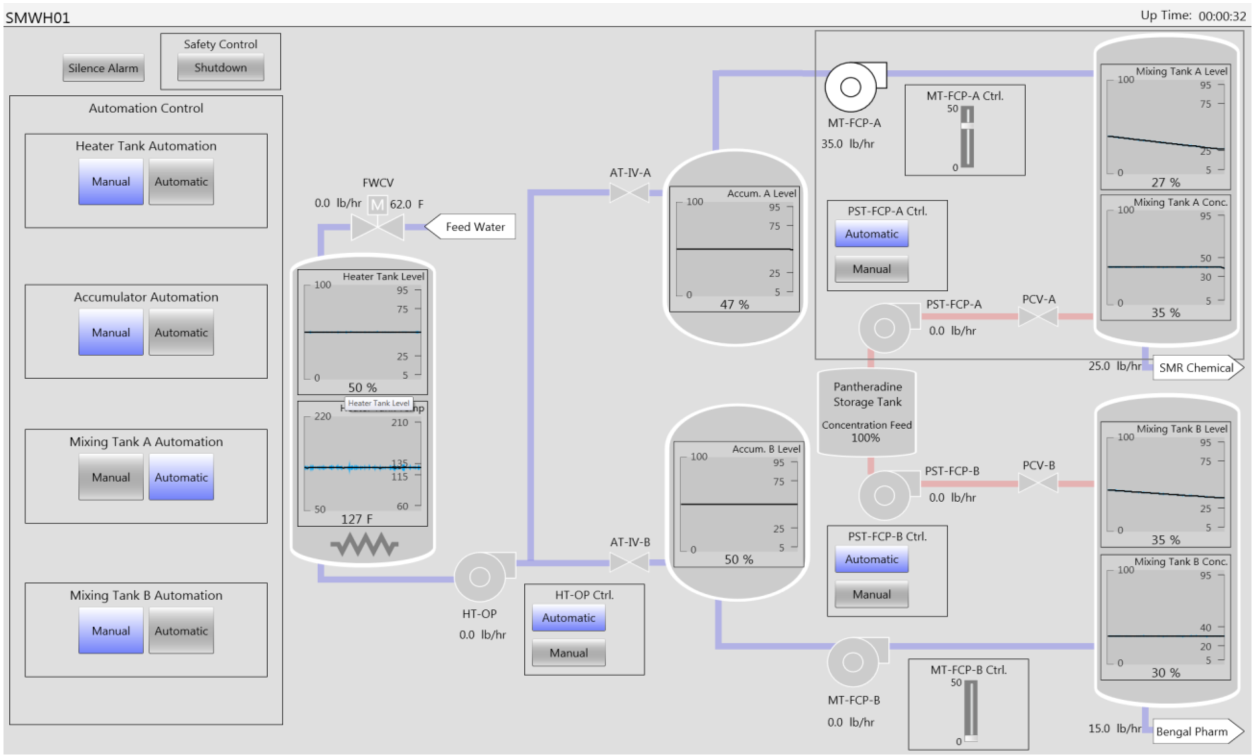

Figure 5. Screenshot From Simulation in the Adaptable Condition.

In the Adaptable automation condition, the process is divided into four high level functions: managing the heater tank, managing the accumulator tanks, managing mixing tank A and managing mixing tank B. Through the task delegation interface, see Figure 5, the operator can choose to delegate any or all of the high level functions to the automation or perform the functions manually. The operator uses the panel to the left to select whether the functions are automated (the default is manual). When a function is delegated to automation it is highlighted with a gray box. The rest of the interface is identical to the manual condition.

2.1.2.4 Fully Automatic

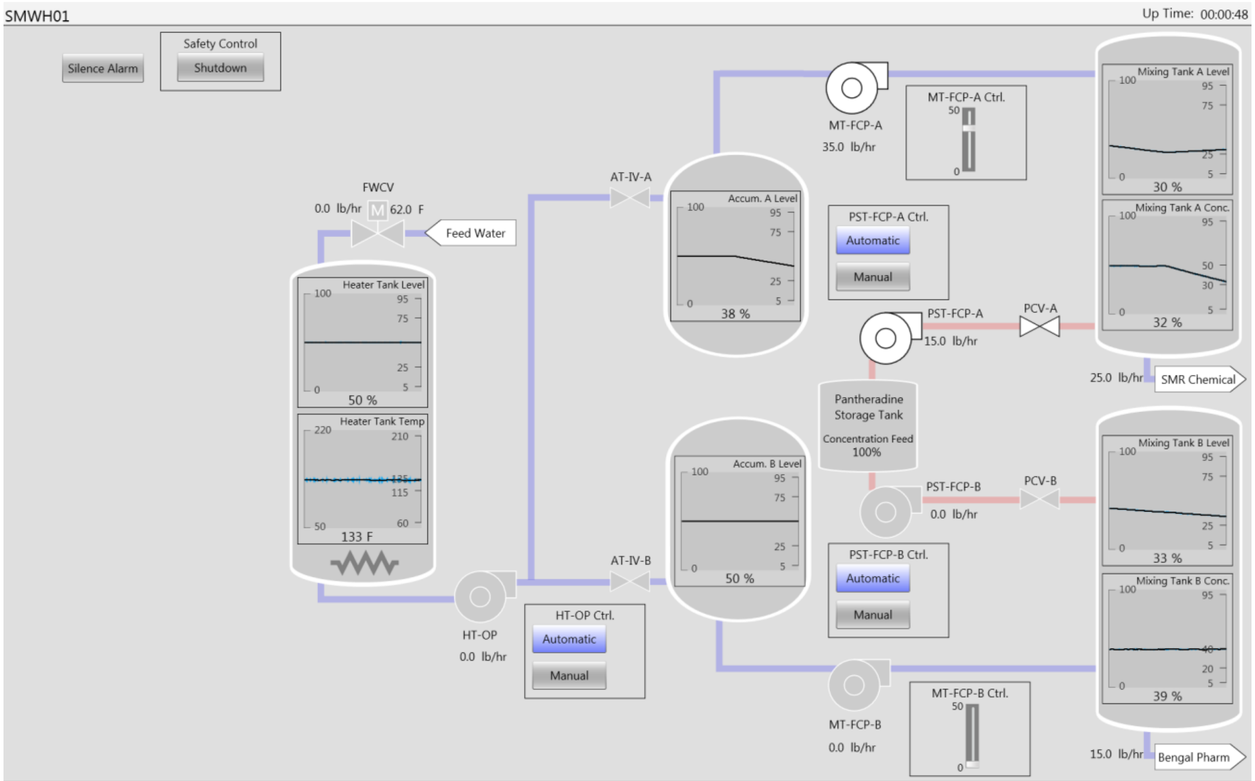

Figure 6. Screenshot From Simulation in the Fully Automatic Condition.

In the Fully Automatic condition, all of the tasks are carried out by the automation. The operator's task is to monitor the system to ensure the automation is working properly, and to take manual action if the automation fails to keep the process within the prescribed parameters. To take a manual action the operator has to select the "Manual" option in the component controller. When the manual action is taken the operator can switch the operating mode back to automatic by selecting the "Automatic" option in the component controller. Figure 6 shows the interface used in the Fully Automatic condition.

2.2 Participants

Participants were 107 undergraduate psychology students at Idaho State University. Six of the 107 participants had incomplete data sets and were dropped from the analysis, leaving 101 participants. There were 35 males and 66 females. The average age was 23 years. The participants were given course credit for their participation.

2.3 Experimental Design

2.3.1 Independent Variables

Automation condition is the main independent variable. As described in Section 2.1.2, there are four levels of automation: Fully Manual, Intermediate (decision support system), Adaptable (task delegation interface), and Fully Automatic. Automation condition was manipulated between participants (each participant executed both scenarios using a single level of automation). Below is a summary of the automation conditions.

1. Fully Manual: The operator handles all tasks manually.

2. Intermediate: The automation monitors, generates responses, and presents possible actions to the operators, and the operator chooses the actions to be taken, the automation then carries out the chosen action.
3. Adaptable: The process was divided into four high level tasks. The operator can choose to delegate any (or all) of the four tasks to the automation or perform the tasks manually.
4. Fully Automatic: Automation handles all tasks, Operator task is to monitor and take manual control if necessary.

The second independent variable is whether the scenario had a fault or had no fault. This was manipulated within participants, where each participant completed one of each scenario type. The fault and the no fault descriptions are below.

1. Fault: The automation fault was the same for each participant. In the simulated process, the chemical flows from the concentration tank to the mixing tank. The flow is controlled by a valve and a pump that, under normal circumstances, turns on automatically when the valve is opened (this was true in all four levels of automation). During the fault scenario, the pump fails to turn on automatically when the valve is opened. The participant must select manual control from the pump controller and turn on the pump manually in order to recover from the fault and keep the parameter in range (participants were trained on how to switch the pump from automatic to manual and how to manually control the pump during the tutorial). The fault was injected at the same time for each participant.
2. No Fault: In the no fault scenario, the participant needs to monitor and control the process parameters, and respond to changing demand in mixing tank B (required concentration and outflow changed during the scenario).

See Appendix B – Introduction Material for a complete description of the scenarios.

2.3.2 Dependent Variables

The researchers measured system performance by assessing the degree to which the 8 parameters were kept within range. They adapted the discrepancy score metric reported in Ha and Seong (2009). In addition, the researchers measured the total time that any one of the 8 parameters was out of range. The simulation software logged the state of the 8 parameters, the automation mode of the subsystems (automatic or manual), and the state of the subsystems (e.g., valves were logged as opened or closed). The data was logged approximately once every five seconds.

2.3.2.1 System Performance

The main performance measure for system performance was discrepancy scores. The discrepancy scores were calculated by comparing the actual value of a given parameter at the times sampled across the entire scenario to the acceptable range of that parameter. The formula for computing the discrepancy score is as follows:

- Discrepancy at time t, $D_i(t) = \begin{cases} \frac{X_i(t)-S_{Ui}}{S_{Ui}-S_{Li}}, if\ X_i(t) > S_{Ui} \\ 0,\ if\ S_{Li} \leq X_i(t) \leq S_{Ui} \\ \frac{X_i(t)-S_{Li}}{S_{Ui}-S_{Li}}, if\ X_i(t) < S_{Li} \end{cases}$
    - S_{Li} = lower bound of parameter i
    - S_{Ui} = upper bound of parameter i
    - $X_i(t)$= value of parameter i at time t

- Average discrepancy for parameter i = $\sum_{t=0}^{T} \frac{D_i(t)}{T}$
- Overall discrepancy score = Sum across all parameters

The difference between this measure and Ha and Seong's (2009) measure is the method by which the scores were normalized. Ha and Seong (2009) use the mean of the parameter at steady state to normalize the score. The process used in this study was designed such the operator must actively maintain the parameters, so it does not have what could described as a steady state. The distance between the upper and lower bounds were use, instead, to normalize the scores.

2.3.2.2 Human Performance

According to many researchers, adaptive automation has the advantages of enabling higher levels of automation by reducing the human performance costs of automation. Therefore it is important to assess the situation awareness, workload and fault management performance using each of the levels of automation.

Situation awareness (SA) was assessed by a freeze probe questionnaire (See Appendix C – Situation Awareness Questionnaire). The SA questionnaire was based on the Situation Awareness Control Room Inventory (SACRI) (Hogg, Follesø, Strand-Volden, & Torralba 1995). SACRI is typically administered by freezing a simulation scenario and quizzing participants about important parameters (identified by a subject matter expert). The questions ask about the current development of parameters, as well as the past and future development of parameters. The SA questionnaire used in this study was a slightly modified version of SACRI.

The participants were asked about each of the 10 parameters they were required to monitor. The questions asked what the current value of the parameter was and how it would develop in the next 10 seconds as well as how it had developed in the past 10 seconds. The 10 second timeframe might seem short, but the scenarios were designed to evolve rather rapidly. The participants were instructed to enter indicate (by entering and 'n' in the input field of the current state questions and selecting a '?' on the past and future trend questions) that they honestly did not know that the value of the parameter was. The formula for computing the SA discrepancy score is as follows:

- Average SA discrepancy for parameter i =
$$\begin{cases} \frac{1}{2}(S_{Ui} - S_{Li}), if\ the\ response\ was\ 'n' \\ \sum_{t=0}^{T} \frac{|X_i(t) - S_i(t)|}{S_{Ui} - S_{Li}}, if\ a\ numerical\ value\ was\ provided \end{cases}$$
    - $X_i(t)$ = Value of Parameter i at time t
    - $S_i(t)$ = Reported Value (on SA questionnaire) of parameter i at time t
    - S_{Li} = lower bound of parameter i
    - S_{Ui} = upper bound of parameter i
- Discrepancy for each freeze probe questionnaire is the sum average across all parameters
- Overall discrepancy score = Sum of the three freeze probe questionnaire scores

The responses to the past and future trend questions were scored in the following way:

- Did the parameter decrease at any time during the 10-second period? If so, "decrease" is considered correct.

- Did the parameter increase at any time during the 10-second period? If so, "increase" is considered correct.
- Was the parameter constant at any time during the 10-second period? Was the parameter's value at the end of the period within 1% or 1 degree of where it was at the beginning of the period? If either is so, "same" is considered correct.
- In some cases, there is more than one correct answer.
- If the subject's answer is among the list of correct answers, score = 0.
- If the subject answered "decreasing" or "increasing", but "same" was a correct answer, score = 0.5.
- If the subject answered "same", but "increasing" or "decreasing" was a correct answer, score = 0.5.
- If the subject said he didn't know, score = 0.5.
- If the subject answered "increasing" but the only correct answer was decreasing, or vice versa, score =1.

Workload was assessed by NASA Task Load Index (NASA TLX) (Hart & Staveland, 1988). The NASA TLX is a subjective workload questionnaire that is administered after each task (See Appendix A for the NASA TLX Questionnaire).

Automation fault management was assessed by determining if the operator, 1) switched the pump to manual, and 2) returned the parameter associated with the fault to acceptable range before the scenario ended. If both conditions were satisfied, then it was counted as a success. If not then it was counted as a failure.

2.4 Experimental Protocol

Participants first read and signed an informed consent form. During the consent process, the participants were informed that they have the opportunity to be included in a drawing for a $25 gift card at the end of the experiment. They were told that they will be eligible to be included based on how they perform on the task (criteria are keeping the process parameters within the required bounds at least 70% of the time). The participants in the Intermediate and Adaptable automation conditions were also encouraged to choose a combination of manual and automatic control that optimized their performance. The gift card drawing was intended to motivate participants to perform well. They then completed a short demographics survey. Participants were also randomly assigned to one of the four automation conditions (in groups of four to force equal sample sizes between automation conditions).

Following the demographic survey, the participants were given a brief overview and tutorial on how to use the simulator. The tutorial took approximately 10 minutes to complete. The tutorial described the system and instructed the participant how to operate the system in manual. The participant then completed a 5-minute practice session operating the simulation in manual, which was followed by another brief tutorial describing how to operate the system in the automation condition the participant had been assigned to. When the tutorials were completed, the participants completed another 5-minute practice scenario.

The participants completed two 10-minute experimental scenarios (one with a fault and one without, order was counterbalanced). Before each scenario, the participants were given a brief overview of the conditions they would encounter in the scenario including acceptable ranges of the eight parameters and the values of the two tank outflows. Participants were also given a brief explanation of the SA questionnaire. Participants in the intermediate and adaptable automation conditions were instructed to use the combination of manual control and automation that they felt provided the best performance. During the scenario, the simulation froze periodically (three times per scenario) and prompted the participants to answer a small set of situation awareness questions. Following each scenario, the participants filled out the NASA TLX survey. Once the last scenario was complete, the participant filled out a short debrief questionnaire. See Figure 7 for an illustration of the experimental protocol. Table 3 describes the number of participants in each condition.

Table 3. The Number of Participants in Each Condition.

	Fault/No fault
Auto	n = 25
Adaptable	n=26
Intermediate	n = 25
Manual	n = 25

Figure 7. A Visualization of the Experimental Protocol.

2.5 Hypotheses

Based on the existing literature in human automation interaction, the researcher s developed the hypotheses presented in Table 4. In addition to the hypothesis listed above, the research investigated the following research questions: How much do participants in the intermediate and adaptable automation conditions use automation?

Table 4. List of Hypotheses.

System Performance	
Discrepancy scores (lower scores represent better performance)	
	Fully Automated <= Adaptable <Intermediate < Manual
Human Performance	
Operator SA	
SA discrepancy scores	Manual <= Adaptable <Intermediate < Fully Automated
SA past and future trend scores	Manual <= Adaptable <Intermediate < Fully Automated
Fault Management	
Success/failure performance	Adaptable >= Manual >Intermediate> Fully Automated
Operator Workload	
NASA TLX Score	Manual >Adaptable >=Intermediate > Fully Automated
Number of manual control actions	Manual >Adaptable >=Intermediate > Fully Automated

3. RESULTS

3.1 System Performance

System performance was assessed by determining how well the human-automation team kept the 8 parameters within the required ranges during the 10-minute scenario. This was computed in two ways. First by discrepancy scores, which is an adaptation of the measure reported in Ha and Seong (2009), as described in Section 2.3.2 – Dependent Variables. The second is by computing the total time any of the 8 parameters were out of range.

3.1.1 Discrepancy Scores

The main performance measure for system performance was discrepancy scores. Discrepancy scores were calculated by comparing the actual value of a given parameter at the times sampled across the entire scenario to the acceptable range of that parameter. The formula for computing the discrepancy score is described in Section 2.3.2.1 – System Performance. Table 5 presents the means and standard deviations of the discrepancy score. It is important to note that the standard deviations are quite large relative the means, indicating that this data is non-normal. In each condition, there are many subjects who perform well and a few subjects who perform very poorly; the difference among groups is primarily a difference in how many subjects perform very poorly.

Table 5. Mean Discrepancy Score by condition. Lower discrepancy scores indicate better performance.

		Mean	Standard Deviation
Fault	Auto	0.19	0.33
	Intermediate	0.33	0.33
	Adaptable	0.36	0.35
	Manual	0.81	0.69
No Fault	Auto	0.09	0.42
	Intermediate	0.17	0.26
	Adaptable	0.17	0.27
	Manual	0.60	0.57

An ANOVA revealed a significant main effect of automation condition ($F(3,96)=17.4$, $p<.0001$) and significant main effect of fault versus no fault ($F(1, 96)= 7.9$, $p=.006$), but not a significant Automation condition X Fault condition interaction. Tukey's Studentized range test revealed significant pairwise comparisons between manual and the other three automation conditions, but there were no significant differences among the other three automation conditions. This indicates that based on discrepancy scores, performance was worst when the process was Fully Manual, but that performance was similar between Fully Automated, Intermediate and Adaptable conditions. Contrary to the hypothesis, there was not a significant difference in system performance between Intermediate and Adaptable automation. Figure 8 illustrates the mean discrepancy scores across the automation and fault conditions. If the fault had a differential effect on discrepancy scores based on the automation condition (i.e., if the fault itself caused one group to perform worse than the others) then we would expect there to be a significant interaction between fault type and automation condition.

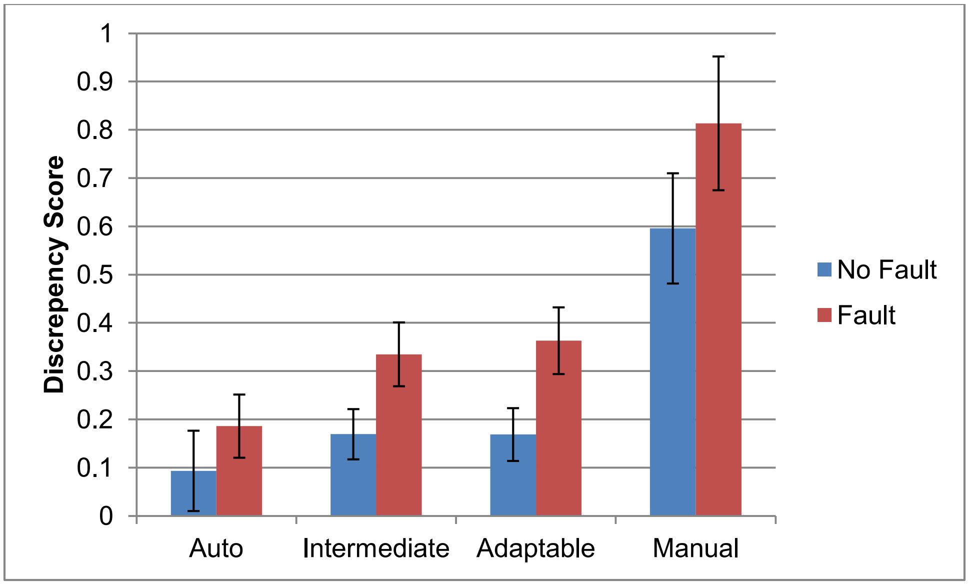

Figure 8. Average Discrepancy Scores. Error bars represent standard error of the mean.

3.1.2 Time Out of Range

Performance on the task was also assessed by computing the total time that any 1 of the 8 parameters was out range.

Table 6 presents the Means and Standard deviations of Time out of Range performance across Automation and Fault conditions. An ANOVA revealed a significant main effect of automation condition ($F(3, 96) = 54.9$, $p<.0001$) and significant main effect of fault versus no fault ($F(1, 96) = 7.9$, $p<=.006$)). All of the pairwise comparisons were significant except for the Intermediate and Adaptable automation conditions. This indicates that performance was best when the process was Fully Automated and worst when the process was Fully Manual. The Intermediate and Adaptable conditions produced better performance than Fully Manual, but not as good as Fully Automated. Contrary to the hypothesis, there was not a significant difference in system performance between Intermediate and Adaptable automation conditions. Figure 9 shows graphical representation of the time out of range.

Table 6. Mean Time out of Range by Automation condition and Fault Condition. Lower time indicates better performance.

		Mean	Standard Deviation
Fault	Auto	132	53
	Intermediate	235	144
	Adaptable	277	129
	Manual	458	113
No Fault	Auto	74	77
	Intermediate	188	166
	Adaptable	233	146
	Manual	400	162

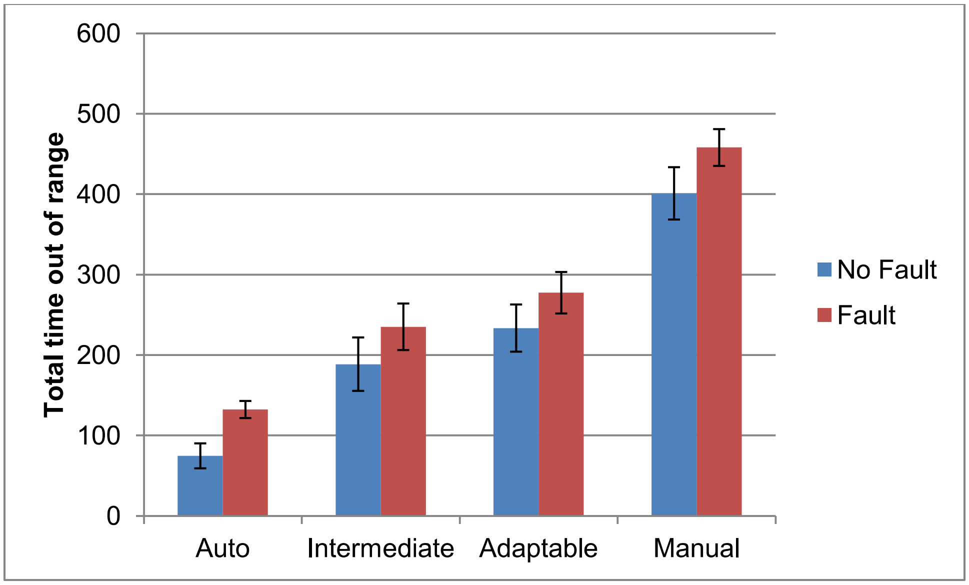

Figure 9. Time Out of Range. Error bars represent standard error of the mean.

3.2 Situation Awareness

Situation awareness was assessed by determining how well the participants kept track of the 8 parameters. This was done by assessing whether the participants knew the value of each parameter at the time of the questionnaire, and by assessing whether they were aware of how the parameters were changing in the past and how they would change in the future.

3.2.1 Situation Awareness Discrepancy Scores

The response to the reports values of the parameters on the SA questionnaire were used to compute SA discrepancy scores. See Section 2.3.2.2 – Human Performance for the formula used to do so. Table 7 presents the means and standard deviations for SA discrepancy score by automation condition and fault condition. In the table, lower time indicates better performance.

Table 7. Means and Standard Deviations for SA Discrepancy Score by Automation condition and Fault Condition. Lower time indicates better performance.

		Mean	Standard Deviation
Fault	Auto	9.97	3.49
	Intermediate	10.16	3.28
	Adaptable	10.30	3.84
	Manual	12.27	5.46
No Fault	Auto	10.97	4.13
	Intermediate	10.70	4.04
	Adaptable	11.98	4.46
	Manual	12.70	5.98

The sum of the SA discrepancy scores across the three SA questionnaires was analyzed using a 2 X 4 mixed ANOVA. As indicated in Figure 10, there was not a significant effect of automation condition or fault condition. This indicates that SA discrepancy scores did not differ depending on automation condition or the presence of a fault.

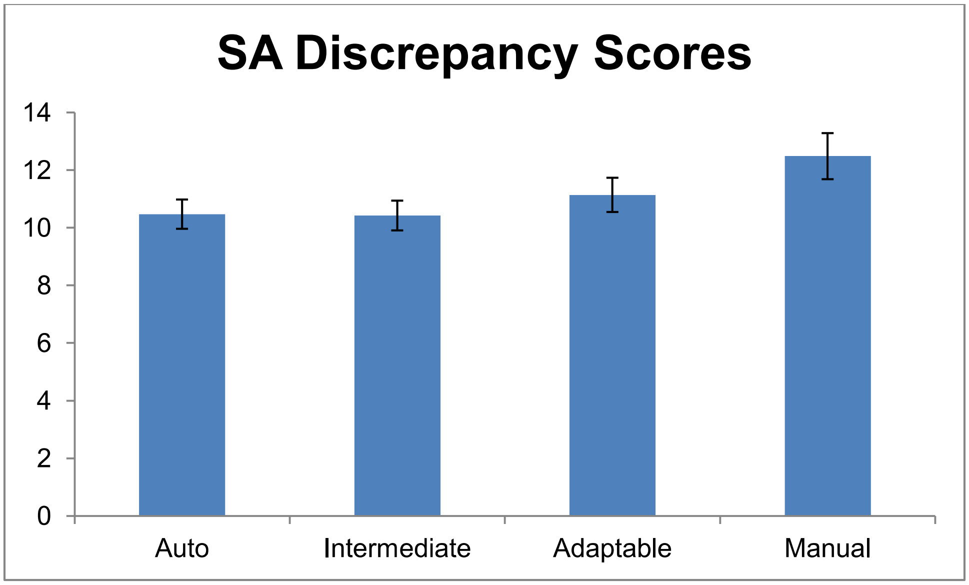

Figure 10. Situation Awareness Discrepancy Scores. Error bars represent standard error of the mean.

3.2.2 Situation Awareness Past and Future Trend Scores

The overall SA Past and Future Trend Score for a scenario is the sum across all parameters on the three freeze probe questionnaires. A total of 60 questions were asked; 0 is a perfect score, blind guessing or a full slate of "I don't know" results in a score just under 30.

An ANOVA revealed there was not a significant effect of automation condition or fault condition on SA Past and Future Trend Scores. This indicates that SA past and future trend scores did not differ depending on automation condition, see Figure 11.

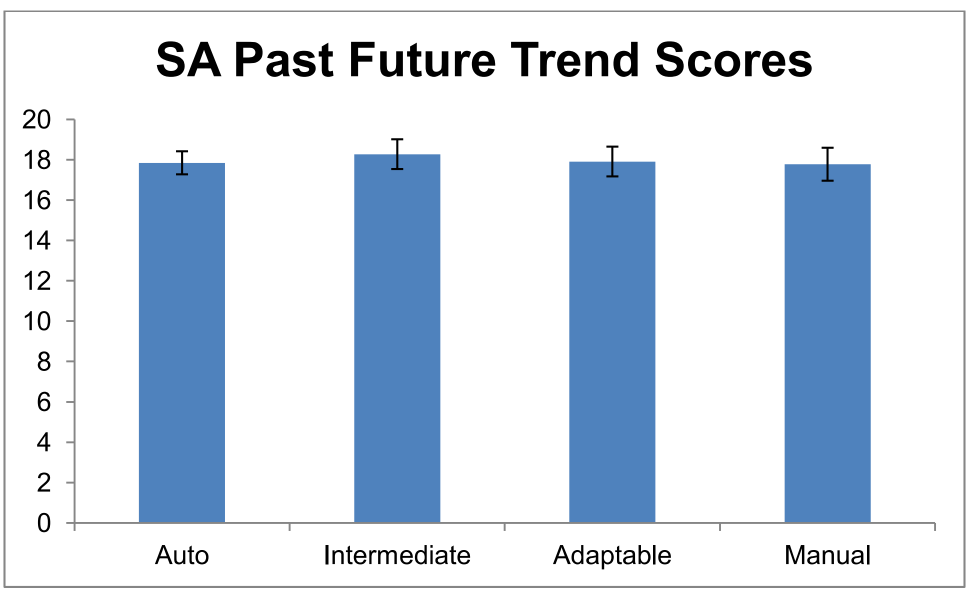

Figure 11. Situation Awareness Past and Future Trend Scores. Error bars represent standard error of the mean.

3.3 Fault Management

3.3.1 Success Rate

Fault performance was scored based on whether the participant took appropriate action and got the parameter back in range after the fault was injected (5 participants kept the parameter within range even after the fault was injected, they were counted as successes). A one way ANOVA revealed a significant effect of automation condition on successful fault management ($F(1, 97) = 6.3$, $p=.001$), and Tukey's Studentized Range Test confirms that the following pairwise comparisons are significant: Automation*Intermediate, Automation*Manual, Manual*Adaptable, and Intermediate*Adaptable. Figure 12 shows the Fault Management Success where the number of successes is based on the following group sizes: Automation: 25, Intermediate: 25, Adaptable: 26, and Manual: 25.

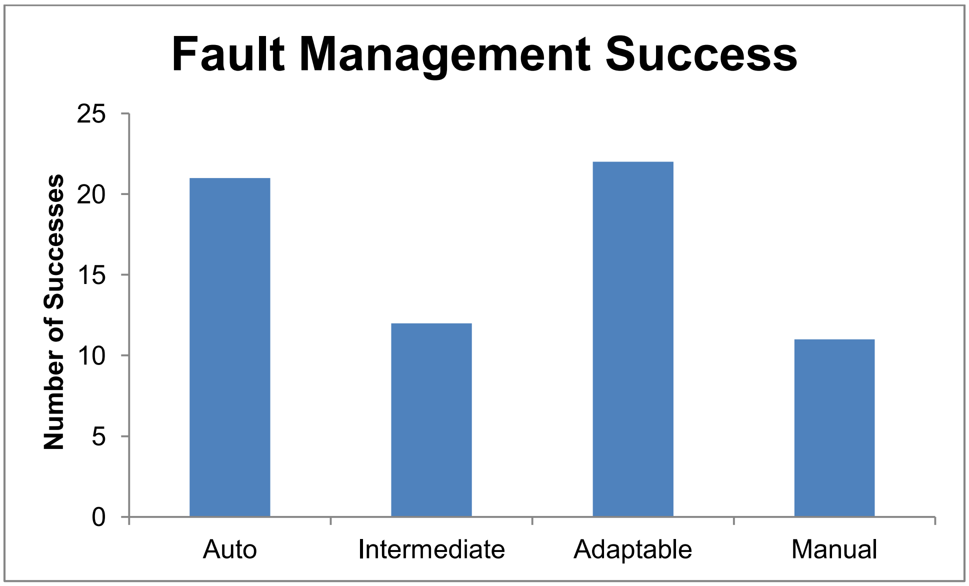

Figure 12. Fault Management Success. The number of successes is based on the following group sizes: Automation: 25, Intermediate: 25, Adaptable: 26, and Manual: 25.

3.4 Workload

3.4.1 Subjective Workload

Subjective workload was assessed by scores on the NASA TLX questionnaire. There were no significant effects among the workload scores. Figures 13 and 14 depict the mean NASA TLX scores grouped by sub score for all four automation conditions. In the figures the error bars represent standard error of the mean.

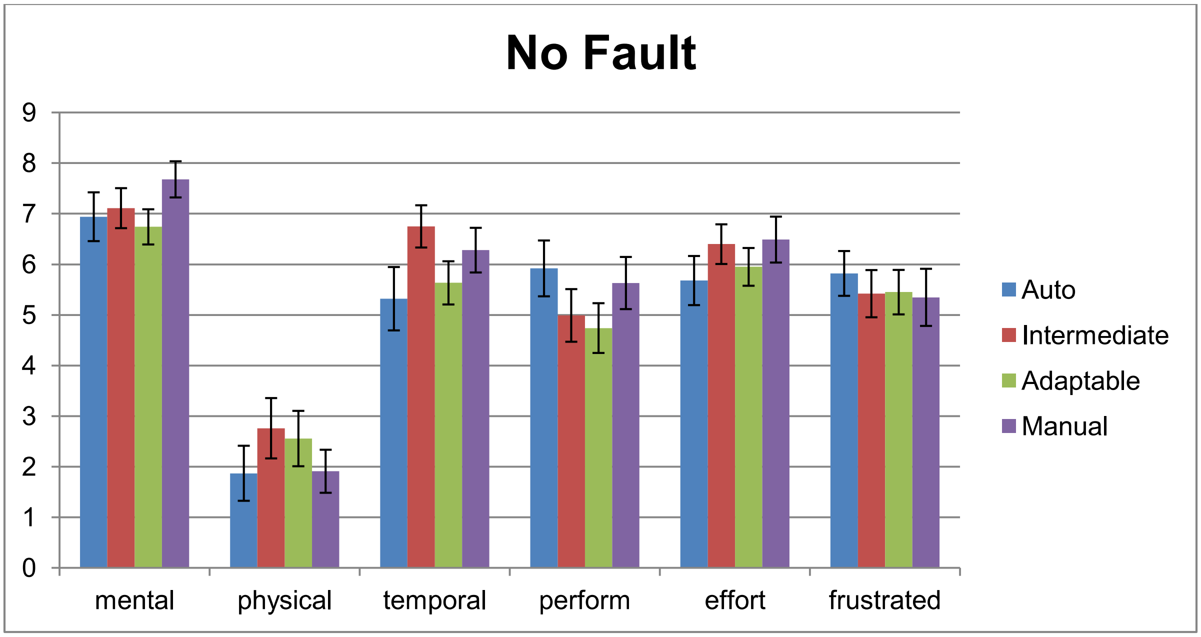

Figure 13. Mean NASA TLX Scores grouped by sub score for all four automation conditions in the No fault condition. Error bars represent standard error of the mean.

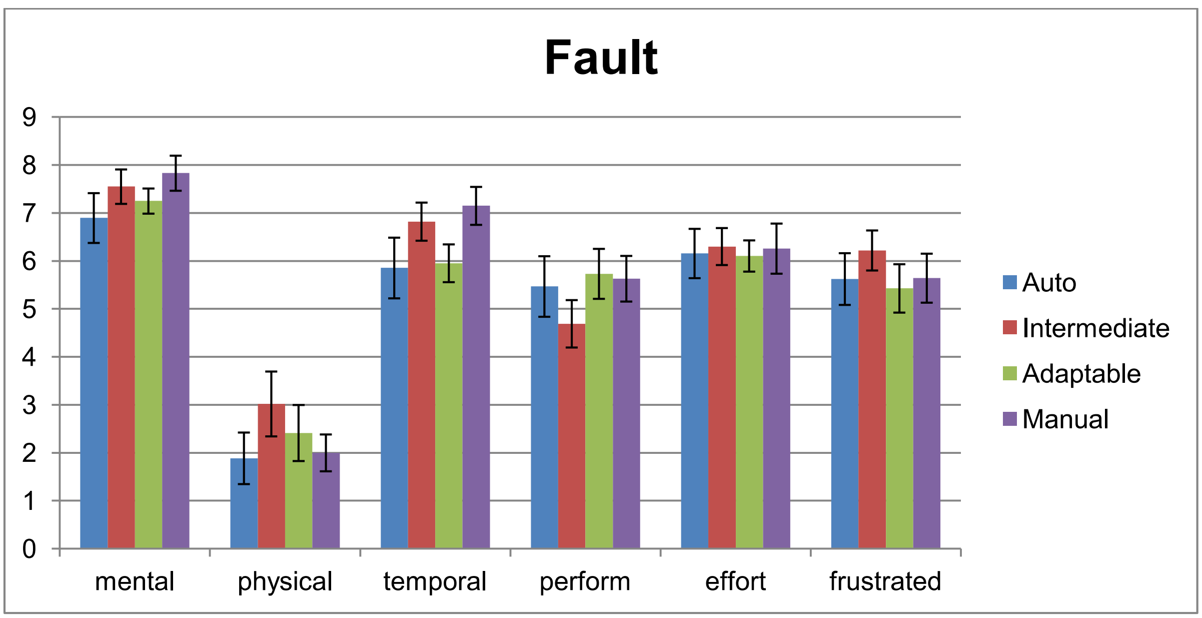

Figure 14. Mean NASA TLX Scores grouped by sub score for all four automation conditions in the Fault condition. Error bars represent standard error of the mean.

3.5 Use of Automation

The Intermediate and Adaptable automation conditions allowed participants to use automation at their discretion. It is important to understand if there was a difference in how participants used the automation in these conditions. Use of automation was characterized by calculating the fraction of the total actions carried out by automation.

- C_H = total number of control actions taken by human
- C_A = total number of control actions taken by automation
- Fraction of actions carried out by automation (Automation Fraction) = $\frac{C_A}{C_H + C_A}$

There was not a significant difference in the average fraction of actions carried out by automation between the Intermediate and Adaptable automation conditions (M = .48 and .M = .51, respectively), indicating that, on average, both groups delegated the same amount to automation.

There was not a significant difference in performance or SA between Adaptable and Intermediate automation conditions based on the automation conditions, however given that individual participants within each of the conditions chose to utilize the automation to different degrees, it is useful to investigate whether the amount of automation used in those two conditions is related to any of the measured variables.

3.5.1 System Performance

Because there was no difference in the fraction of automated actions between the Intermediate and Adaptable automation conditions, the two conditions are combined in the following analyses.

Automation Fraction is a significant predictor of performance. The line of best fit was Total Time Out = 329 + 52*Fault - 243*Automation Fraction, with all 3 terms significant. r^2 = .28 (.26 due to Automation Fraction, .02 due to fault.).

A participant in the Intermediate and Adaptable automation conditions in the No-fault situation is expected to be out of bounds 86±53 seconds if making maximum possible use of automation, and 328±51 seconds if making minimum possible use of automation. By comparison, the average Auto subject (automation fraction very near 1) was out of bounds for 75±29 seconds and the average Manual subject (automation fraction 0) was out of bounds for 400±64 seconds. [± are 95% confidence intervals, not 95% prediction intervals.]

With a fault, the regression equation predicts 138±51 seconds out of bounds with maximum possible use of automation and 381±52 seconds with minimum possible use. The average Auto subject was out of bounds 132±21 seconds and the average Manual subject was out of bounds for 458±45 seconds.

This indicates that in the conditions where participants could use automation at their discretion, using more automation resulted in better performance.

Figure 15 depicts the scatter plot of Total Time Out versus Automation Fraction for Intermediate and Adaptable Automation conditions in the no fault scenario. Figure 16 depicts the scatter plot of Total Time Out versus Automation Fraction for Intermediate and Adaptable Automation conditions in the fault scenario. The red squares plot the mean values for Fully Automatic and Fully Manual conditions for comparison.

The pattern of results was similar, but not as strong based on average discrepancy score. The line of best fit is:

Average discrepancy Score = 0.288 + .187*Fault - .246*Automation Fraction

All 3 terms are significant (t=4.73 for intercept, t=3.19 for fault, t=-2.67 Automation Fraction.). r^2 was .13 (.05 due to Automation Fraction, .08 due to fault.).

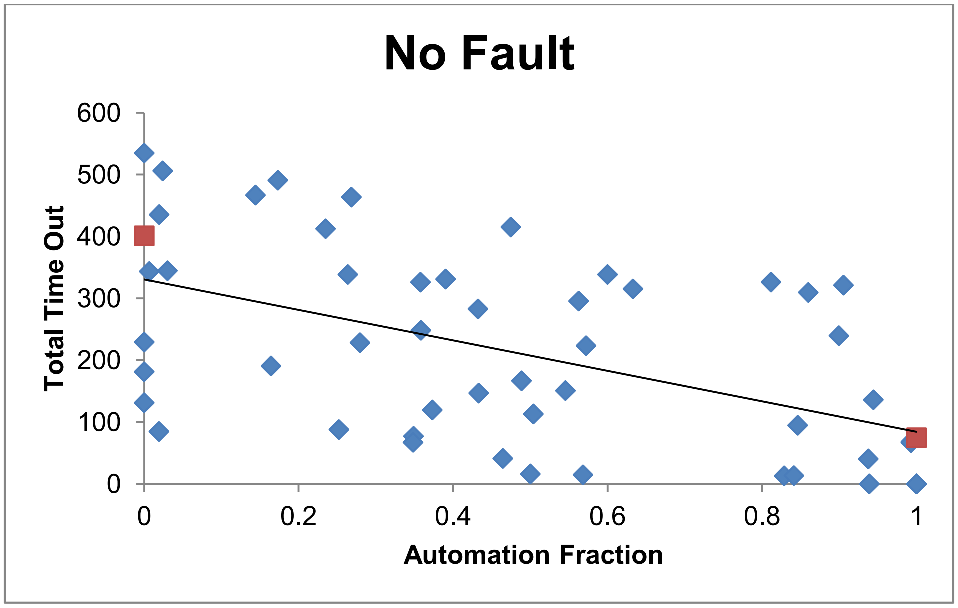

Figure 15. Scatter plot of Total Time Out versus Automation Fraction for Intermediate and Adaptable Automation Conditions in the no fault scenario. The red squares plot the mean values for Fully Automatic and Fully Manual conditions for comparison.

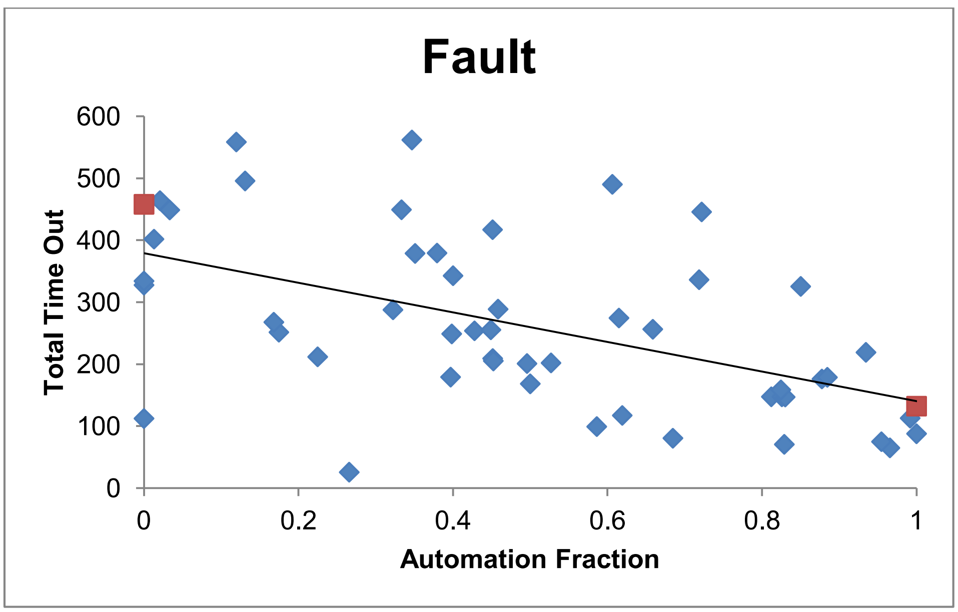

Figure 16. Scatter plot of Total Time Out versus Automation Fraction for Intermediate and Adaptable Automation Conditions in the fault scenario. The red squares plot the mean values for Fully Automatic and Fully Manual conditions for comparison.

3.5.2 Situation Awareness

Automation fraction was also a significant predictor of SA for the Adaptable and Intermediate automation groups. As automation increased, SA performance decreased (that is, SA discrepancy score increased). The line of best fit is:

Sum of SA Discrepancy Score = 8.49 + 4.62*Automation Fraction, t=4.10, r^2=.13.

Figure 17 illustrates the relationship between SA discrepancy score and automation fraction for the Intermediate and Adaptable automation conditions and Figure 18 illustrates the relationship between SA past and future trend score and automation fraction for the Intermediate condition. In the figure, a higher score indicates worse performance.

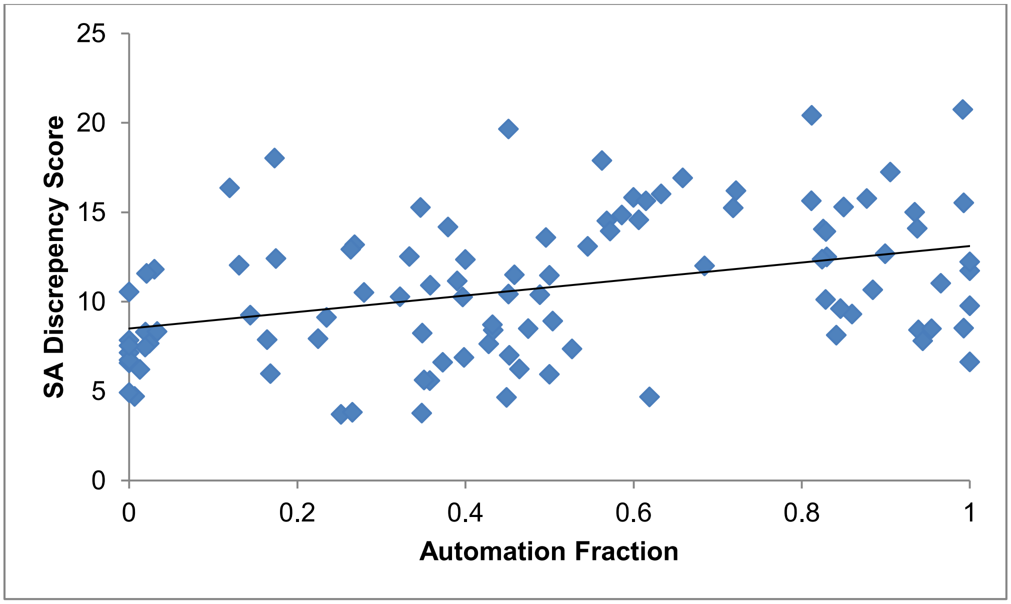

Figure 17. Scatter plot of SA Discrepancy Score versus Automation Fraction for Intermediate and Adaptable Automation Conditions.

The SA Past and Future Trend scores yielded a different pattern of results for Intermediate and Adaptable automation conditions. Automation fraction was not a significant predictor of SA Past and Future trend Scores for Adaptable automation, but it was for Intermediate automation. The line of best fit for Intermediate automation is:

SA Past and Future Trend Score = 14.56 + 7.73*Automation Fraction ($t=4.63$, $r^2=.30$)

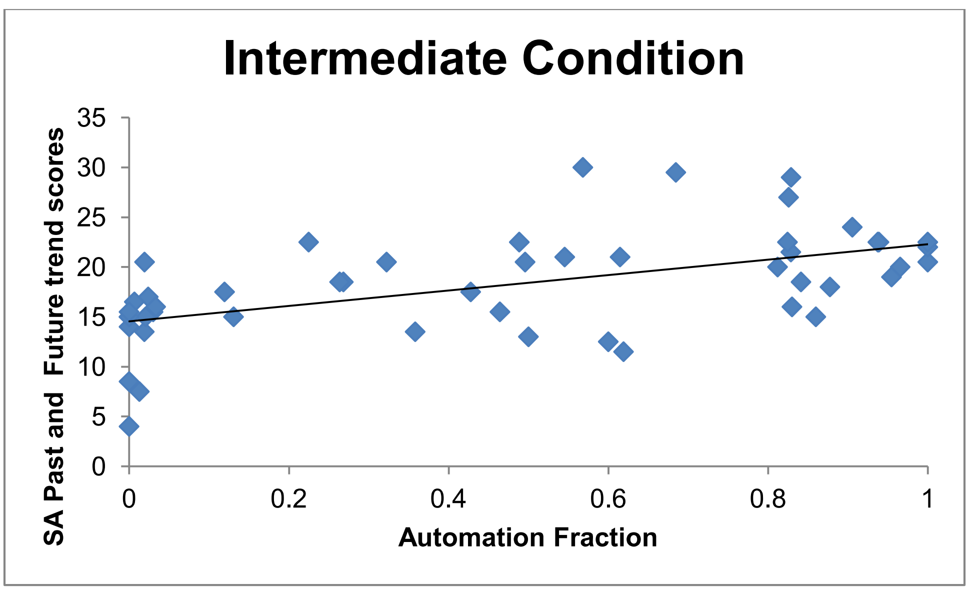

Figure 18. Scatter plot of SA Past and Future Trend Score versus Automation Fraction for Intermediate Automation Condition. Higher Scores indicate worse performance.

4. DISCUSSION AND CONCLUSIONS

Based on the human-automation collaboration literature, the researchers predicted that the adaptable automation would result in an optimum combination of human and system performance. That is, adaptable automation would yield system performance that was better than both manual and intermediate automation levels (but not necessarily as good as fully automatic), and human performance (i.e., SA, workload, and fault management) that was better than automatic and intermediate levels (but not necessarily as good as manual performance). The results of this study demonstrate that adaptable automation enhances system performance compared to manual control, but is not as good as fully automatic. Interestingly, system performance between adaptable and intermediate levels of automation did not differ significantly. This indicates that adaptable automation may be no better than intermediate levels of automation at handling the fundamental tradeoffs of increasing levels of automation from a system performance perspective.

Another way to assess the effectiveness of adaptive automation is to look at how it affects human performance compared to other levels of automation. The researchers compared fault management performance, SA, and workload in order to compare human performance across the automation conditions. More participants in the Adaptable automation condition successfully managed the fault than in any of the other automation conditions. This may indicate that fault management performance is better using adaptable automation. However, the fault management performance for the other three levels of automation was not as expected. The researchers hypothesized that participants in the Manual condition would performance the best, and automation condition would perform the worst. The results indicate that the opposite was true.

Previous research finds that SA tends to decrease with increasing automation (Endsley, 1995, 1996, 1997; Endsley & Kaber, 1999; Kaber & Endsley, 1997, 2004; Parasuraman et al., 2000; Sheridan, 2002; Wickens & Hollands, 2000; Wright & Kaber, 2005). Hence, the situation awareness findings were not as expected based on the reviewed literature. Not only was there no difference between the Adaptable and Intermediate automation conditions, there was not a significant difference across all four levels of automation. There are several possible explanations for why the SA scores did not differ based on automation condition. First, the simulation required a large number of actions in order to keep the process under control. It is possible that the participants were too busy keeping the process under control to effectively monitor the parameters. However, if this were true, then we would expect SA performance to be better in the Fully Automated condition, which it was not. Another possible explanation is that the process simply changed too quickly for participants to keep track of all of the parameters. Because the SA questionnaire was based on a technique used for nuclear power plant control room simulations, where changes typically happen much more slowly, then perhaps the method for measuring SA in this study was not appropriate.

One interesting result is that for the Adaptable and Intermediate automation conditions, the amount of automation (as measured by the fraction of total actions carried out by the automation) was related to performance on the SA questionnaire. Participants who used more automation had worse SA than participants who used less automation. These findings are in line with previous research; however it is unclear why the results for Fully Automated and Fully Manual do not follow the same pattern.

Taken together the findings from this study do not confirm the common claim that adaptable automation is an effective method to manage human performance and system performance tradeoffs associated with increasing automation, but do not necessarily refute it. Specifically, other researchers often state that adaptable automation is a better solution than intermediate levels of automation. Based on the results of this study, performance using adaptable automation is similar to that of using intermediate automation. However, as noted many of the other results were also not consistent with previous literature. Due to the limitations present in this study, some of the findings in this study are somewhat inconclusive. One limitation is in the definitions of LOA used and their specific implementations in this study.

Although the researchers adhered to previously accepted definitions of level of automation, it may not necessarily be accurate to characterize automation as strictly as it has been in the literature (or in this first experiment study). The definition of intermediate automation used in this study was taken directly from Endsley and Kaber's (1999) taxonomy of levels of automation. Their taxonomy has ten levels. Level five (called decision support) is the level we defined as our intermediate level, and is described below:

"The computer generates a list of options that the human can select from; the operator may still generate his or her own options. The computer is responsible for implementing the chosen action. This LOA is common in many expert systems or decision support systems in which the operator may use or ignore the option guidance provided by the system."

Though this definition represents an intermediate level of automation in Endsley and Kaber's (1999) taxonomy, it is not strictly static automation. That is, the operator has a choice in whether the actions suggested by the decision support system are carried out by the automation. Adaptable automation is typically defined as dynamic allocation of functions. A decision support system defined this way is inherently adaptable, because an operator can direct the automation to carry out suggested actions or ignore the suggestions. This may be the reason that the current research failed to detect a difference between Intermediate and Adaptable automation. Though the method by which tasks were delegated to automation was quite different, the number of tasks delegated to automation was roughly the same for both conditions. Based on the fraction of automated actions, in the Intermediate and Adaptive automation conditions, the *amount* of automation used predicted performance and SA better than the specific automation condition participants were assigned to. Surprisingly, the amount of automation in Fully Automatic and Fully Manual conditions did not predict SA performance. Future research should explore why this was the case.

Further, researchers designed the study to replicate other research on adaptable task delegation interfaces; however the process control simulation was not sufficiently complex to design a truly hierarchical abstraction scheme to use for the task delegation interface. Therefore, although his study compared an adaptive task delegation interface to an intermediate level of automation and found no differences, it is premature to conclude that in general adaptive task delegation interfaces do not enhance performance compared to intermediate automation. Further research needs to be conducted to replicate and extend these findings to a truly hierarchical task delegation interface.

This study was done with university students with minimal training; consequently, the participants were novice controllers of the process. As the discrepancy scores and time out of range scores illustrate, participants did not master the process before conducting the experimental scenarios. The researchers intended the task to be difficult enough to generate deviation in performance, but the task was perhaps too difficult for novice users. This limitation may explain the inconclusive findings of the SA measures. Participants may have not had an adequate understanding of the process, resulting in inadequate SA across all automation conditions. Future research efforts should ensure that participants are trained to proficiency to increase the chances of detecting potential difference in SA between automation conditions.

The same limitation may explain the surprising result that there was not a difference in subjective workload (based on NASA TLX scores). If you consider the number of manual control actions in the two extremes of LOAs, the NASA TLX results are especially surprising. Excluding the fault period, participants in the Manual condition took an average of 130 actions, while those in the Automation condition took 0 manual actions. One would expect that given such an enormous difference in the number of manual actions, the subjective workload would be quite different across the two extremes. If the participants in all of the conditions were still struggling to understand the process, it is possible that they were overwhelmed even when their only task was to monitor the process as it was controlled by automation.

Finally, the process control simulation required many actions to keep the parameters within range. Most of these tasks would be automated in a real process control situation, which may make it inappropriate to generalize the results directly to process control.

In conclusion, the results of this research indicate that although many researchers have suggested adaptive automation may enable higher levels of automation without introducing human performance costs, further research is needed to determine if it is truly superior to intermediate levels of automation. Therefore, experimental research efforts in the HAC project will investigate the effects of a modified process control simulation (with modification to the automation schemes to more accurately reflect the schemes used in other research) on expert performers. The researchers will extensively train a group of participants on operation of the simulation in order to eliminate the issues related to novice participants controlling the process.

One insight that this research provides is that the numerous methods to characterize automation may not be needed to define automation's effect on human performance. There is not a consensus in the human automation interaction literature regarding definitions for important terms like level of automation. According to these taxonomies, there is an enormous amount of nuance in the way that automation can be implemented. Careful consideration of the condition described in this research reveals that a commonly used intermediate level of automation, decision support, could actually be characterized as adaptive. The decision support prompts in this study were presented when the process parameters were in danger of going out of range. If the participants were doing a good job manually controlling the process, then he would get fewer decision support prompts. However, if the participant was having trouble controlling the process he would get more decision prompts. This situation is analogous to an operator performance-based adaptive automation scheme. In fact, the intermediate level of automation (which was designed based on a literal interpretation of Endsley and Kaber's taxonomy) was more adaptive than the adaptable automation condition, which required the operator to initiate shifts in automation. The existing taxonomies may be very useful in extensively *describing* automation, but they may not be needed to be specified in such a minute level of detail to explain differences in human performance when interacting with automation.

Another important issue related to the potential complexity of automation design impacts on human performance is that the design of the human system interface can drastically influence the degree to which a particular automation design is effective. Therefore, it is difficult to provide general guidance on how to effectively design automation, because it largely depends on the specific circumstances. Future research should aim to define the minimum set of factors that influence human performance and system performance in automated systems. This will help to provide generalizable design guidance for automated systems.

As mentioned above, empirical and analytical research will be conducted as needed to investigate the affect of automation on performance, efficiency, and safety. The design of human-system interaction to best support operation of automated systems and multiple product streams will be investigated through process simulator studies. The results from these additional studies will be used to refine the HAC model, which was initially developed by the researchers 2013. The refined model will be used to identify design requirements needed to ensure safe and effective operation. Based on the requirements a guidance document for development and evaluation of HAC design will be developed. In order to make the guidance document more practical to use, the researchers will develop a software tool for design and evaluation of Human-Automation Collaboration.

5. REFERENCES

Adams, M. J., Tenney, Y. J., & Pew, R. W. (1995). Situation awareness and the cognitive management of complex systems. *Human Factors, 37*(1), 85-104. doi:10.1518/001872095779049462.

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19* (6), 775-779.

Bedny, G. Z., & Meister, D. (1999). Theory of activity and situation awareness. *International Journal of Cognitive Ergonomics, 3*(1), 63-72.

Bedny, G. Z., Karwowski, W., & Jeng, O. J. (2004). The situational reflection of reality in activity theory and the concept of situation awareness in cognitive psychology. *Theoretical Issues in Ergonomics Science, 5*(4), 275-296. doi:10.1080/1463922031000075070.

Billings, C. (1991). Human-centered Aircraft Automation: A Concept and Guidelines (NASA Tech. Memo No. 103885). Washington, DC: NASA

Calhoun, G. L., Ruff, H. A., Spriggs, S., & Murray, C. (2012, September). Tailored Performance-based Adaptive Levels of Automation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 56, No. 1, pp. 413-417). SAGE Publications.

Dehais, F., Causse, M., & Tremblay, S. (2011). Mitigation of conflicts with automation: Use of cognitive countermeasures. *Human Factors, 53*(5), 448-460. doi:10.1177/0018720811418635

Dixon, S. & Wickens, C. (2006). Automation reliability in unmanned aerial vehicle control: A reliance-compliance model of automation dependence in high workload. *Human Factors*, *48* (3), 474-486.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*(1), 32-64. doi:10.1518/001872095779049543

Endsley, M. R. (1996). Automation and situation awareness. In R. Parasuraman & M. Mouloua (Eds.), *Automation and Human Performance: Theory and Applications*, 163-181. Mahwah, NJ: Lawrence Erlbaum.

Endsley, M. R. (1997). Level of automation: Integrating humans and automated systems. In *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting*, Santa Monica, CA: Human Factors and Ergonomics Society, 200-204.

Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M. R. Endsley & D. J. Garland (Eds.), *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Endsley, M. R., & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics, 42,* 462-492.

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors*, 37(2), 381-394.

Ha, J. S., & Seong, P. H. (2009). HUPESS: Human performance evaluation support system. *In Reliability and Risk Issues in Large Scale Safety-critical Digital Control Systems* (pp. 197-229). Springer London.

Hart, S.G., & L.E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research", Human Mental Workload, Elsevier, Amsterdam, pp.139-183 (1988).

Hogg, D. N., Follesø, K., Strand-Volden, F., & Torralba, B. (1995). Development of a situation awareness measure to evaluate advanced alarms systems in nuclear power plant control rooms. *Ergonomics, 38*(11), 2394-2413.

Jou, Y. T., Yenn, T. C., Lin, C. J., Yang, C. W., & Chiang, C. C. (2009). Evaluation of operators' mental workload of human-system interface automation in the advanced nuclear power plants. *Nuclear Engineering and Design, 239,* 2537-2542. doi:10.1016/j.nucengdes.2009.06.023.

Kaber, D. B., & Endsley, M. R. (1997). Out-of-the-loop performance problems and the use of intermediate levels of automation for improved control system functioning and safety. *Process Safety Progress, 16*(3), 126-131.

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 Ergonomic Science, 5*(2), 113-153. doi: 10.1080/1463922021000054335.

Kaber, D. B., & Riley, J. M. (1999). Adaptive automation of a dynamic control task based on secondary task workload measurement. International journal of cognitive ergonomics, 3(3), 169-187.

Kaber, D. B., Perry, C. M., Segall, N., McClernon, C. K., & Prinzel III, L. J. (2006). Situation awareness implications of adaptive automation for information processing in an air traffic control-related task. International Journal of Industrial Ergonomics, 36(5), 447-462.

Kaber, D. B., Wright, M. C., Prinzel, L. J., & Clamann, M. P. (2005). Adaptive automation of human-machine system information-processing functions. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *47*(4), 730-741.

Klein, G. A., Moon, B. M., & Hoffman, R. R. (2006). Making sense of sensemaking. Part 2: A macrocognitive model. *IEEE Intelligent Systems, 21*(5), 88-92. doi:10.1109/mis.2006.100.

Klein, G. A., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2007). A data-frame theory of sensemaking. In R. R. Hoffman (Ed.), *Expertise out of context: Proceedings of the Sixth International Conference on Naturalistic Decision Making.* (pp. 113-155). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Le Blanc, K.L, & Oxstrand, J. H. (2014). Initiators and Triggering Conditions for Adaptive Automation in Advanced Small Modular Reactors. Proceedings of the ASME 2014 Small Modular Reactors Symposium SMR2014, Washington, DC.

Lee, J., & See, K. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors,* 46 (1), 50-80.

Lin, C. J., Yenn, T. C., & Yang, C. W. (2009). Evaluation of operators' performance for automation design in the fully digital control room of nuclear power plants. *Human Factors and Ergonomics in Manufacturing & Service Industries, 20*(1), 10-23. doi:10.1002/hfm.20168.

Lin, C. J., Yenn, T. C., & Yang, C. W. (2010a). Automation design in advanced control rooms of the modernized nuclear power plants. *Safety Science, 48*, 63-71. doi:10.1016/j.ssci.2009.05.005.

Lin, C. J., Yenn, T. C., & Yang, C. W. (2010b). Optimizing human-system interface automation design based on a skill-rule-knowledge framework. *Nuclear Engineering and Design, 240*, 1897-1905. doi:10.1016/j.nucengdes.2010.03.026.

Liu, Q., Nakata K., & Furuta K. (2004). Making control systems visible. *Cognition, Technology & Work*, *6*, 87–106.

Manzey, D., Reichenbach, J., & Onnasch, L. (2008). Performance Consequences of Automated Aids in Supervisory Control: The Impact of Function Allocation. In Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting. Santa Monica, CA: Human Factors and Ergonomics Society.

Miller, C. A., Shaw, T., Emfield, A., Hamell, J., Parasuraman, R., & Musliner, D. (2011). Delegating to Automation Performance, Complacency and Bias Effects under Non-Optimal Conditions.

In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 55, No. 1, pp. 95-99). SAGE Publications.

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

O'Hara, J. & Brown, W. (2002). *The Effects of Interface Management Tasks on Crew Performance and Safety in Complex, Computer-based Systems*. (NUREG/CR-6690). Washington, D.C.: U.S. Nuclear Regulatory Commission.

O'Hara J., Gunther, W., & Martinez-Guridi, G. (2010). *The Effects of Degraded Digital Instrumentation and Control Systems on Human-system Interfaces and Operator Performance* (BNL Tech Report No. 91047-2010). Upton, NY: Brookhaven National Laboratory.

Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2013). Human Performance Consequences of Stages and Levels of Automation An Integrated Meta-Analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 0018720813501549Onnasch et al., 2013

Oxstrand, J., O'Hara, J., Le Blanc, K.L, Whaley, A.M., Joe, J.C., and Medema, H. (2013a). Development of an Initial Model of Human-Automation Collaboration – Results from a Needs Analysis. Idaho National Laboratory External Report. INL/EXT-13-28682, Rev 1.

Oxstrand, J., Le Blanc, K.L., Joe, J.C., Whaley, A.M., Medema, H., and O'Hara, J. (2013b). Framework for Human-Automation Collaboration: Conclusions from Four Studies. Idaho National Laboratory External Report. INL/EXT-13-30570, Rev. 0.

Oxstrand, J., Joe, J.C., Le Blanc, K.L., Medema, H., Whaley. A.M., and O'Hara, J. (2013c). Framework for Human-Automation Collaboration: Project Status Report. Idaho National Laboratory External Report. INL/EXT-13-29453, Rev 0.

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.

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors*, *39* (2), 230-253.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 30*(3), 286-297).

Rice, S., Hughes, J., McCarley, J., & Keller, D. (2008). Automation Dependency and Performance Gains Under Time Pressure. In *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Rook, F., & McDonnell, M. (1993). Human Cognition and the Expert System Interface: Mental Models and Inference Explanations. *IEEE Transactions on Systems, Man, and Cybernetics. 23* (6), 1649-1661.

Ross, J., Szalma, J., Hancock, P., Barnett, J., & Taylor, G. (2008). The Effect of Automation Reliability on User Automation Trust and Reliance in a Search-And- Rescue Scenario. In *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Roth, E., Hanson, M., Hopkins, C., Mancuso, V., & Zacharias, G. (2004). Human in the Loop Evaluation of a Mixed-initiative System for Planning and Control of Multiple UAV Teams. In *Proceedings of the*

Human Factors and Ergonomics Society 48th Annual Meeting. Santa Monica, CA: Human Factors and Ergonomics Society.

Sauer, J., Nickel, P., & Wastell, D. (2013). Designing automation for complex work environments under different levels of stress. Applied ergonomics, 44(1), 119-127.

Shaw, T., Emfield, A., Garcia, A., de Visser, E., Miller, C., Parasuraman, R., & Fern, L. (2010). Evaluating the Benefits and Potential Costs of Automation Delegation for Supervisory Control of Multiple UAVs. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 54, No. 19, pp. 1498-1502). SAGE Publications.

Sheridan, T. B. (2002). *Humans and Automation: System Design and Research Issues.* New York, NY: John Wiley & Sons, Inc.

Smith, K., & Hancock, P. A. (1995). Situation awareness is adaptive, externally directed consciousness. *Human Factors, 37*(1), 137-148. doi:10.1518/001872095779049444.

Smith, A. & Jamieson, G. (2012). Level of Automation Effects on Situation Awareness and Functional Specificity in Automation Reliance. *In Proceedings of Human Factors Society and Ergonomics Society, 56th Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society.

van de Merwe, K., Oprins, E., Eriksson, F., & van der Plaat, A. (2012). The influence of automation support on performance, workload, and situation awareness of air traffic controllers. *The International Journal of Aviation Psychology, 22*(2), 120-143. doi:10.1080/10508414.2012.663241

Wickens, C., Li, H., Santamaria, A., Sebok, A., & Sarter, N. (2010). Stages and Levels of Automation: An Integrated Meta-analysis. In *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting*, Santa Monica: CA: HFES.

Wickens, C. & Hollands, J. (2000). *Engineering Psychology and Human Performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Willems, B., & Heiney M. (2002). *Decision Support Automation Research in the En Route Air Traffic Control Environment* (DOT/FAA/CT-TN02/10). Washington, DC.: Federal Aviation Administration.

Wright, M. C., & Kaber, D. B. (2005). Effects of automation of information processing functions on teamwork. *Human Factors, 47*(1), 50-66.

APPENDIX A - NASA TLX

To be administered after each task

For each of the categories below, please draw a line indicating where on the scale you think the task falls. For example if you thought mental demand on this task was medium, you would draw a line as close to the center as you could. When evaluating each category, compare the task to driving a car in no traffic versus driving a car in heavy traffic.

Mental Demand -- *How much mental activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?*

Low High

Physical Demand – *How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.) Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?*

Low High

Temporal Demand – *How much time pressure did you feel due to the rate or pace at which the task or task elements occurred? Was the pace slow and leisurely or rapid and frantic?*

Low High

Performance – *How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?*

Good Poor

Effort – *How hard did you have to work (mentally and physically) to accomplish your level of performance?*

Low High

Frustration Level – *How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?*

Low High

APPENDIX B – INTRODUCTION MATERIAL

Introduction

Congratulations! You have just been hired as a process operator for SMR Chemical. At SMR Chemical, we make a variety of chemicals that are used in the pharmaceutical industry. At the plant where you will be working, we make Pantheradine™, which is used in the manufacture of a new kind of children's pain reliever. SMR chemical manufactures its own product using Pantheradine™, and sells the rest of it to Bengal Pharmaceuticals. Bengal Pharmaceuticals manufactures a full line of children's medication that uses Pantheradine™, and requires different amounts and concentrations for each product.

You will be responsible for the part of the process where we take raw Pantheradine™ that was produced in another part of the plant and prepare it for use by SMR Chemical and Bengal Pharmaceuticals. Pantheradine™ is rarely used in full concentration. In order to deliver the specific concentration of Pantheradine™ required for each product, we mix the full-strength, raw Pantheradine™ with heated water, which is necessary to ensure efficient dilution. The process you will control consists of moving the heated water into two storage tanks, and then sending the appropriate amount of heated water and concentrated Pantheradine™ to the mixing tanks. One tank is designated for SMR Chemical; the other is designated for Bengal Pharmaceuticals. The properly diluted Pantheradine™ is then sent to each manufacturer to be used in their respective processes.

Your job is to ensure that SMR chemical has a steady flow of properly mixed Pantheradine™, and to provide Bengal Pharmaceuticals with Pantheradine™ at the flow rate and concentration that they specify.

Scenario 1 – No Fault

Initial Requirements

Equipment/system	Required Parameter Value
Level of Heater Tank	20-80%
Temperature of Heater Tank	120F+-10F
Accumulator Tank A Level	15-85%
Accumulator Tank B Level	15-85%
Tank C Level	20-80%
Tank C Concentration	50%+-10%
TanK C outflow	30 lbs/hour
Tank D Level	20-80%
Tank D Concentration	50%+-10%
TanK D outflow	20lbs/hour

Time Sequence of Events

Time	Action
03:15	Freeze for SA Questionnaire
04:00	Message: Bengal Pharm is stopping production in 30 seconds. They anticipate production will restart in 2 minutes, and they will require 60% concentration at 30lbs/hour.
04:30	Message: Bengal Pharm has stopped production. Please make necessary adjustments.
06:00	Actual Message: Bengal Pharm has started production.
7:30	Freeze for SA Questionnaire
8:30	Freeze for SA Questionnaire

Scenario 2 – Automation Fault Injected

Initial Requirements

Equipment/system	Required Parameter Value
Level of Heater Tank	20-80%
Temperature of Heater Tank	120F+-10F
Accumulator Tank A Level	15-85%
Accumulator Tank B Level	15-85%
Tank C Level	20-80%
Tank C Concentration	50%+-10%
TanK C outflow	20 lbs/hour
Tank D Level	20-80%
Tank D Concentration	80%+-10%
TanK D outflow	30 lbs/hour

Time Sequence of events

Time	Action
02:45	Freeze for SA Questionnaire
04:20	Message: Bengal Pharm is stopping production in 15 seconds. They anticipate production will restart in 1 minute, and they will require 50% concentration at 20lbs/hour.
04:35	Message: Bengal Pharm has stopped production. Please make necessary adjustments.
05:20	Message: Bengal Pharm has started production.
06:45	Freeze for SA Questionnaire
07:30	Automation Fault
08:30	Pump returns to operating normally
09:45	Freeze for SA Questionnaire

APPENDIX C – SITUATION AWARENESS QUESTIONNAIRE

Instructions to the participant

- During the scenario, the simulation will occasionally freeze and prompt you to answer some questions (see right) about the parameters you are monitoring and controlling.
- Do your best to answer the questions based on your knowledge.
- If you honestly don't know the answer and can't make an educated guess, simply enter a lower case "n".
- Similarly, if you honestly don't know whether the value has increased or decreased in the last 10 seconds, then select the "?".

Situation Awareness Questionaire

Parameter	Approximate Value	Development in the past 10 secs	Development in the next 10 secs
Heater Tank Level	%	Same / Decreased / Increased / ?	Same / Decrease / Increase / ?
Heater Tank Temperature	F	Same / Decreased / Increased / ?	Same / Decrease / Increase / ?
Accumulator Tank A Level	%	Same / Decreased / Increased / ?	Same / Decrease / Increase / ?
Accumulator Tank B Level	%	Same / Decreased / Increased / ?	Same / Decrease / Increase / ?
Mixing Tank A Level	%	Same / Decreased / Increased / ?	Same / Decrease / Increase / ?
Mixing Tank A Concentration	%	Same / Decreased / Increased / ?	Same / Decrease / Increase / ?
Mixing Tank A Outflow	lbs/hr	Same / Decreased / Increased / ?	Same / Decrease / Increase / ?
Mixing Tank B Level	%	Same / Decreased / Increased / ?	Same / Decrease / Increase / ?
Mixing Tank B Concentration	%	Same / Decreased / Increased / ?	Same / Decrease / Increase / ?
Mixing Tank B Outflow	lbs/hr	Same / Decreased / Increased / ?	Same / Decrease / Increase / ?

Continue