INL/EXT-13-28682

Development of an Initial Model of Human-Automation Collaboration – Results From a Needs Analysis

Johanna Oxstrand
John O'Hara
Katya L. LeBlanc
April M. Whaley
Jeffrey C. Joe
Heather Medema

March 2013

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

SMR/ICHMI/INL-TR-2013/01
INL/EXT-13-28682

Development of an Initial Model of Human Automation Collaboration – Results from a Needs Analysis

Johanna Oxstrand[a]
John O'Hara[b]
Katya L. LeBlanc[a]
April M. Whaley[a]
Jeffrey C. Joe[a]
Heather Medema[a]

March 2013

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

[a] Idaho National Laboratory
[b] Brookhaven National Laboratory

EXECUTIVE SUMMARY

The U.S. Department of Energy (DOE) and the commercial nuclear power industry are exploring alternatives to meet energy demands in the United States. As part of this mission, they are looking at small modular reactors (SMRs) and advanced small modular reactors (aSMRs). The DOE, and in particular, the Office of Nuclear Energy (NE) is sponsoring research and development (R&D) on small reactors, as evidenced by NE's 2010 Report to Congress, Nuclear Energy Research and Development Roadmap. In short, DOE-NE's mission is to assist in revitalization of the U.S. nuclear industry, including development of advanced designs, through R&D. By doing so, NE can help accelerate deployment of new plants in the short term, support development of advanced concepts for the medium term, and promote design of revolutionary systems for the long term.

All aSMR designs will employ advanced digital instrumentation, controls, and human-machine interfaces (ICHMI), technology that is significantly more advanced that existing analog systems in the light water reactor fleet. The U.S. DOE recognizes that ICHMI research, development, and demonstration is needed to address the specific technical challenges and technological gaps of ICHMI for aSMR designs. The new aSMRs will be designed to utilize new automation and instrumentation and control technologies, and there are a number of concerns about how those technologies will affect human performance and the overall safety of the plant. It is expected that aSMRs will rely on automation to a greater extent than the current nuclear power plant fleet. However, there are many issues and concerns that still need to be addressed related to how automation should be designed and implemented. For example, further researcher is needed to address how humans and automation will collaborate under various operational conditions.

The Human-Automation Collaboration (HAC) research project is one of three research efforts related to investigating how the advanced technologies planned for aSMR designs will affect human factors and human performance. Given the increased use of automation in aSMR designs, the HAC research project is investigating the consequences of allocating functions between the operators and automated systems. The research effort addresses the questions of what the collaboration level should be and how it should be implemented to have the greatest positive impact on overall plant performance and safety. The research project is also developing a model of HAC, which will support aSMR designers when evaluating their proposed approach for conduct of operations in terms of how humans and automation collaborate. The research results will inform the integration and cooperation between plant staff and automation, with the purpose of maximizing productivity and safe operations of aSMRs. One key research goal is providing a technical basis to support the reduction in aSMR operations and maintenance costs through reduced staffing per unit, which is made possible by greater integration and cooperation between plant staff and automation. Additionally, aSMR vendors will be able to use the results of this research effort to inform development of the technical basis for their licensing case in submittals to the Nuclear Regulatory Commission.

This report documents the work conducted to date in this research effort. The research team conducted a review of the human factors, psychological, and automation literature to identify and characterize the current state-of-practice in human-automation interaction, and to identify factors that influence HAC. Additionally, current standards that are applicable to HAC were reviewed.

While conducting the review of human factors literature regarding automation and human performance, the research team identified and analyzed key contributing factors, such as levels of automation, reliability, the cognitive functions that the automation is responsible for, and how those aspects of automation affect operator performance (e.g., the out-of-the-loop phenomenon) and system performance. These activities combined provided critical information for better understanding of the current HAC state-of-practice and the means to construct the initial framework for HAC.

While human factors and psychological research has gone a long way in identifying the factors that influence HAC, there are clear gaps in the current state of knowledge for addressing the needs of aSMRs. First, the majority of the human factors literature (with a few exceptions) defines performance problems associated with certain HAC configurations, but the literature does not necessarily illuminate the circumstances that lead to successful HAC.

Second, taken together, findings from the existing literature would recommend using intermediate levels of automation for most functions in order to keep the operator in the loop. However, aSMRs will employ much higher levels of automation to meet the need of reducing operations and maintenance costs to a per kilowatt cost that is comparable to the existing fleet of reactors. Therefore, extensive research needs to be conducted to investigate how to enable higher levels of automation, while still keeping the operator actively engaged in operation of the plant.

Based on the literature review, the research team developed an initial HAC model, with a focus on the conceptual interaction between humans and automation. The model is a means of identifying the gaps that need to be bridged by new research in order to establish a technically sound future state of practice for aSMR plants. The HAC model defines (1) the important design dimensions of automation that impact automation's use by personnel and integrated human-automation performance, and (2) what aspects of human cognition, behavior, and performance mediate automation's use by personnel (i.e., the model identifies how human cognition and behavior interact with the design dimensions of automation to affect overall human-system performance).

The initial HAC model (presented in this report) will be updated as additional research is conducted, and will ultimately inform the development of procedures and guidance that will support aSMR designers when evaluating their proposed design of the human-system interaction in terms of HAC.

The research team concluded, based on the activities described in this report, that aSMR designs will benefit from more empirical research on how to maximize the use of automation to achieve cost savings, but at the same time avoid an adverse effect on safety or performance. The team proposes three main topics for further research; (1) Impact of Highly Automated Advanced Small Modular Reactors on Operator Awareness, (2) Regaining/Reacquisition of Operator Awareness, and (3) Effect of Human-Automation Collaboration Characteristics on Operator's Use of Automation.

CONTENTS

FIGURES

TABLES

ACRONYMS

aSMR	advanced small modular reactor
DOE	Department of Energy
HAC	human-automation collaboration
HSI	human-system interface
ICHMI	instrument and control, human-machine interaction
I&C	instrument and control
LOA	level of automation
MWe	megawatt of electricity
NE	Office of Nuclear Energy
NPP	nuclear power plant
NRC	U.S. Nuclear Regulatory Commission
O&M	operations and maintenance
R&D	research and development
SA	situational awareness
SG	steam generator
SMR	small modular reactors

Development of an Initial Model of Human Automation Collaboration – Results from a Needs Analysis

1. INTRODUCTION

1.1 Background

Small nuclear reactors were first employed in the late 1940s and 1950s by the U.S. military to power military operations and naval vessels. The U.S. Navy launched the first nuclear-powered submarine, the U.S.S. Nautilus, in 1955 and today operates 82 vessels powered by small reactors. During a period of experimentation, the U.S. Army used small reactors to power military bases and provide electricity to remote areas before ceasing exploratory use in the mid 1970s. New proposals for small modular reactors (SMRs), an emerging category of new nuclear power plant (NPP) design, which contribute to future energy needs, are very similar to these early reactors and the naval reactors still in use today (U.S. Department of Commerce, 2011).

Unlike typical, current U.S. NPPs that may generate over 1,000 megawatts electricity (MWe), SMRs generate far fewer MWe per unit, many producing less than 100 MWe. According to the classification adopted by the International Atomic Energy Agency, a "small reactor" is one with a total possible electrical power of 300 MWe or less (IAEA, 2005, 2006). The U.S. Department of Energy (DOE) also has used the "300-MWe or less" threshold to define SMRs. In addition to their ability to serve as a source of energy for electricity generation, the new reactor types also can carry out other critical functions, including hydrogen production and industrial process heat applications such as desalination, water purification, and production of both liquid transportation fuels and petrochemicals. According to recent estimates from the International Atomic Energy Agency, more than 45 SMR designs currently are being developed (IAEA, 2009).

DOE divides SMR designs into two major technology classes: designs based on existing light water reactors, such as integral pressurized water reactors, and designs that are not based on existing light water reactor technologies. This second class is known as advanced SMR (aSMR) designs and includes designs that rely on a coolant other than water (such as helium, sodium, lead-bismuth, or molten salt) or have fuels other than uranium oxide (such as metallic or triso fuels).

DOE and the commercial nuclear power industry are exploring alternatives to meet the energy demands in the United States. As part of this mission, they are looking at SMRs and aSMRs. DOE and, in particular, the Office of Nuclear Energy (NE) is sponsoring research and development (R&D) on small reactors, as evidenced by NE's 2010 Report to Congress, *Nuclear Energy Research and Development Roadmap*. The report states, "NE's objective is to assist in the revitalization of the U.S. industry through R&D. By advancing technologies through R&D, NE can help accelerate deployment of new plants in the short term, support development of advanced concepts for the medium term, and promote design of revolutionary systems for the long term."

Regardless of the specific reactor design, all aSMR designs will employ advanced digital instrumentation, controls, and human-machine interfaces (ICHMI), which consists of technology significantly more advanced than existing analog systems in the light water reactor fleet. DOE recognizes that ICHMI research, development, and demonstration is needed to address the specific technical challenges and technological gaps of ICHMI for aSMR designs. Consequently, the DOE aSMR Program has established a critical ICHMI research pathway, consisting of research projects in the following five technical areas (Wood, 2012):

- SMR assessment methods
- SMR ICHMI

- SMR materials, fuels, and fabrication
- SMR licensing support
- SMR advanced concepts evaluation.

In the ICHMI research pathway, a particular focus of the R&D effort for aSMRs is in the area of human factors. Given the plans for using new technologies in aSMRs, there are a number of concerns about how these technologies will affect human performance. As will be explained in Section 1.2 of this report, it is expected that in order to be cost-competitive, aSMRs will utilize automation to a greater extent than is presently employed in the current fleet of U.S. NPPs. However, aSMR plants will not be fully automated, which raises the unanswered question of how to best design and implement automation to ensure optimal interaction of automation and human operators. Collaboration between humans and automation will be necessary to operate the advanced controls of aSMR plants, but the specific manner in which the collaboration occurs and the way to optimize both human and system performance are unresolved issues that need further research before any aSMR control technologies are implemented.

To address these issues, the DOE ICHMI research pathway includes three related research efforts that investigate separate but related building blocks needed for the design and operation of a highly automated plant. These three research projects are as follows:

- **aSMR Concepts of Operations:** A plant's concept of operations generally is understood to be a high-level description of the plant, its systems and their functions, and how operating personnel will work and interact with the system to achieve their responsibilities. SMR and aSMR plants will require defining non-traditional concepts of operations to address the unique operating scenarios that aSMRs will involve, all of which are expected to have an effect on human performance, staffing, training, and reliability. The aSMR Concept of Operations Project is investigating the impact of new operational concepts on human performance and responsibilities and the implications for effective aSMR plant operations and safety. Specifically, this project will involve research activities encompassing analysis of operational scenarios based on prospective plant configurations, functional analysis of operational tasks, evaluation of function allocation options, simulator-based testing for analysis and verification of operational concepts, assessment of minimal staffing requirements for a multi-module design, and assessment of the impact on human performance and reliability.
- **aSMR Human-Automation Collaboration (HAC):** Given the increased use of automation in aSMR designs, the HAC Project is developing a framework for integrating humans with automation to maximize the productive and safe operations of aSMRs. Automation that is all-or-none, or that is implemented without consideration to the impact on human operators, often produces problems of overall system performance. This project is investigating how to best employ a modern approach that uses collaboration of personnel and automation, thereby capitalizing on the strengths of each. The HAC Project will produce guidance on how to define levels and implementation of automation in a way that support successful aSMR operations.
- **aSMR Supervisory Control:** This project addresses supervisory control capabilities that enable high levels of automation to help ensure the economic viability of SMRs through optimal staffing. The research involves development of fundamental capabilities (such as control, decision, diagnostics and prognostics) coupled with demonstration within an architectural framework suitable for integrating these capabilities. The first-phase development involves definition of a supervisory control strategy and establishment of an architecture that allows implementation of hierarchical hybrid control of multi-modular plant systems. In addition, foundational modules and architectural structures will be generated to enable initial demonstration of supervisory control capabilities in the next phase of research. The demonstration focus will involve plant configurations based on near-term SMR designs, in which multiple units are coupled through common balance of plant systems (e.g., multiple reactors feeding a single turbine generator). In subsequent phases of the research, the supervisory control

architectural framework will be adapted to multi-unit plants with more complex architectures (e.g., reconfigurable product streams), based on other SMR designs.

Together, these three efforts set the context, determine the impact and consequences, and define the capabilities related to collaboration between human operators and automated systems. See Figure 1 for an illustration of the relationships between these three research efforts. Specifically, the Concept of Operations Project sets the context by identifying plant functions and defining operating scenarios, strategies, and requirements that establish operational functions and staffing requirements to inform function allocation and human-machine collaboration. The Supervisory Control Project identifies and develops the technological capabilities to enable necessary automation given the input from HAC.

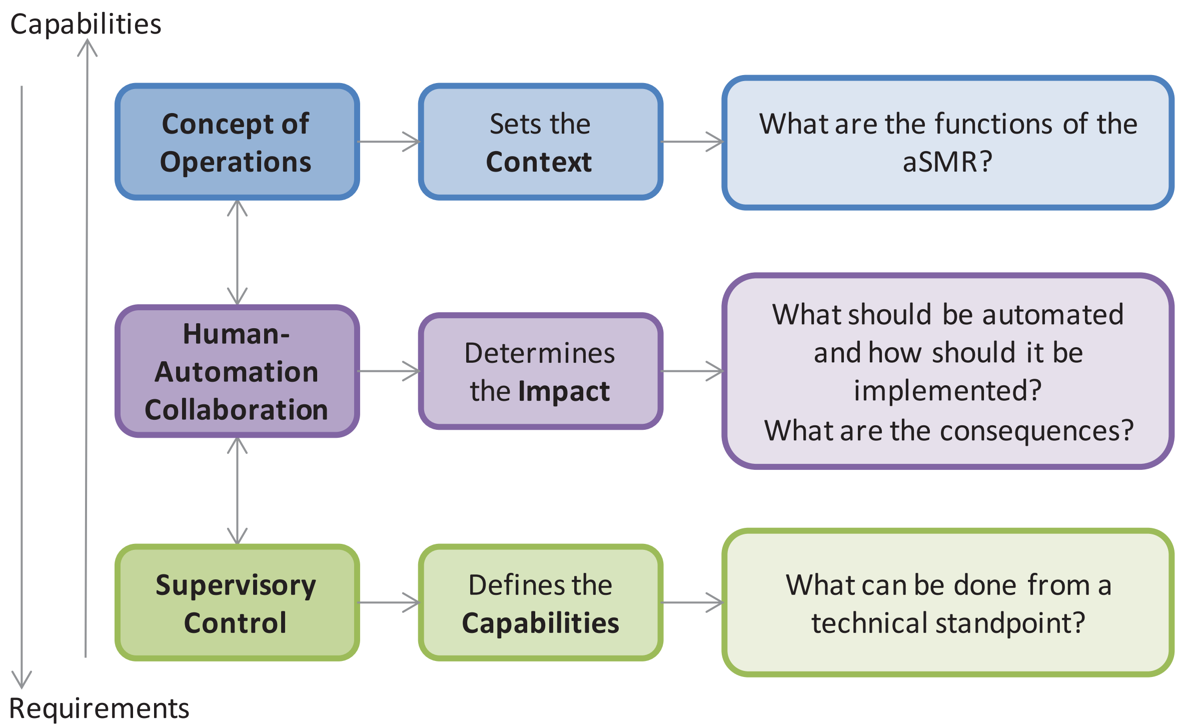


Figure 1. Interrelationships between advanced small modular reactor research efforts.

Between the context of collaboration and the capabilities available to support collaboration is an integration layer (i.e., the HAC Project). The HAC research effort determines the impact of using automation (given the aSMR context and the automation's capabilities) on human and system performance. In making this determination, questions (such as what should be automated and how and what are the consequences) will be answered.

The above discussion provided a high-level overview of the HAC Project. Before discussing the details of the project, it is essential to document the reasons why automation is necessary for aSMR designs and what the consequences of such high levels of automaton are for human performance. The remainder of Section 1 will detail the reasons why high levels of automation are required for aSMR designs (Section 1.2), provide an overview of the impact of automation on human performance in complex systems (Section 1.3), and detail the specific scope and goals of the HAC Project (Section 1.4).

1.2 Automation as an Economic Requirement for Advanced Small Modular Reactor Plants

As mentioned above, it is expected that aSMR plants will be highly automated. For example, NUREG-1368 states that the Power Reactor Innovative Small Module (PRISM) liquid-metal (sodium) cooled reactor will be highly automated. The primary reason this automation is necessary to reduce staffing costs. One of the goals of the new aSMR designs is more economical electrical power generation. One approach to making aSMRs more economical in comparison to current plants is to reduce operations and maintenance (O&M) costs. In existing light water reactors, O&M costs are significantly higher than fuel costs and labor is over 50% of that cost (Thomas, 2012). Thomas noted:

> Nuclear power could be at a considerable disadvantage if it continues to rely on an operating model that requires a large plant staff. The largest component [of] a typical nuclear plant's operating and maintenance (O&M) cost is labor, representing well over 50% of the cost structure. Labor will continue to be a rising cost over time while technology will generally be a falling cost. Thus, generation sources that are more technology-based could significantly erode the cost advantage that the nuclear power industry has enjoyed. Digital technology provides the opportunity to transform the operating model of the nuclear power plants (NPPs) from one based on a large staff performing mostly manual activities to an operating model based on highly integrated technology operated by a smaller staff (p. 883).

One approach to achieving an "operating model based on highly integrated technology operated by a smaller staff" is by increasing automation. O&M costs can be reduced significantly if fewer personnel are needed to achieve production and safety goals.

The aSMR designs feature a number of characteristics that are vastly different from traditional power plant designs. These features have substantial consequences for the economic viability of the plant and for the design of HAC. aSMR plants are modular and scalable (i.e., the plant can be constructed module-by-module and is scalable to different sizes). Economically, the advantage to doing this is that completed modules may operate while other modules are still in construction and capital investment costs can be recovered more quickly. However, having multiple modules could impose an economic consequence that may be unsustainable if staffing and instrumentation and controls do not scale with size. In other words, for example, if each module or unit must be operated by a four-person crew, the O&M costs for multiple modules will be far greater than the O&M costs of traditional light water plant designs. If aSMR plants are required to maintain the same level of staffing per reactor as the traditional light water plants (i.e., are not granted a waiver from 10 CFR 50.54(m)(2)(iii)), then the aSMR design becomes economically unviable. For this reason, it is economically necessary to increase automation and reduce staffing.

However, it should be noted that reduction in aSMR plant staffing, whether through automation or other means, has been identified as a potential safety issue by the U.S. Nuclear Regulatory Commission (NRC), following an Issue Identification and Ranking Program the NRC used to independently assess and identify potential technical and regulatory issues (Smith & Moore, 2009), and in NRC's report to Congress on advanced reactor licensing (NRC, 2012). For NRC to consider waivers to staffing requirements in 10 CFR 50.54(m)(2)(iii), SMR vendors must be able to present technical data demonstrating how reactors could be safely and securely operated with fewer control room operators and security personnel. The aSMR Concepts of Operations Project is addressing the appropriate level of staffing for aSMR designs and the regulatory guidance for reduced staffing; the HAC research effort is focused on developing a framework for HAC within the context of increased automation and reduced staffing.

The next section will provide a brief overview of some of the issues that are raised by high levels of automation[1].

1.3 Automation and Human Performance in Complex Systems

Given the necessity for high levels of automation in aSMR designs, it is imperative that research be conducted to develop a comprehensive understanding of how humans interact with automation and what the consequences of automation are for human performance. This section provides examples of the challenges that automation presents to human operators, which explains why additional research on the best way to design and implement automation to avoid performance problems is necessary.

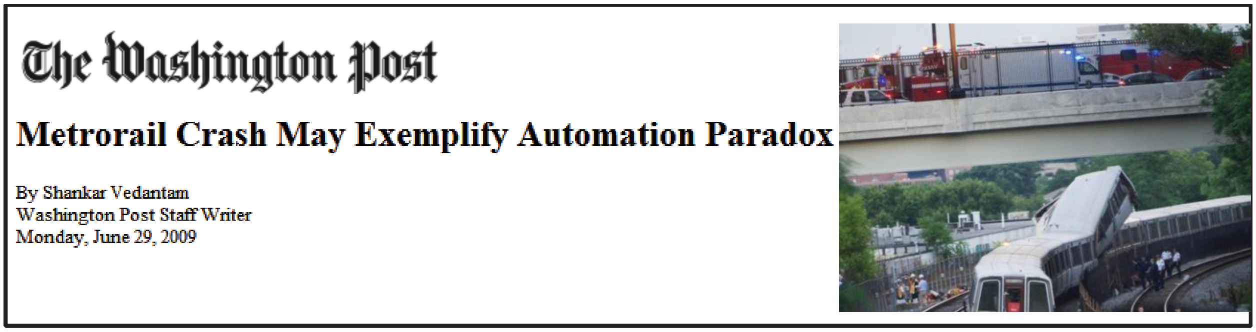
As the design of human-machine systems became more complex, designers often viewed personnel as the "weak link," or as the most unreliable and unpredictable aspect of the system. From an engineering perspective, the solution to making systems more operationally reliable was automation. A prevailing philosophy emerged to automate all functions that could be automated, leaving personnel to manage what could not be automated. In essence, automation was viewed as a means to make system performance safer and more reliable and as a means to reduce operator workload.

However, research and operating experience soon revealed that simply considering whether humans or machines were more capable agents for performing a specific function was not sufficient. Significant human performance problems were observed in highly automated systems such as the operator's loss of awareness of the system's state, the high workload associated with recovering from automation failure, and the loss of skills for manually performing tasks usually performed by automation (see Section 2 of this report for a detailed discussion of these issues). In a seminal article discussing such issues, Bainbridge (1983) referred to this as the "ironies of automation."

Further studies have shown that a significant contributing factor to the difficulties operators encounter in highly automated systems is that the interfaces between operators and automation are poorly designed (Billings, 1997a, 1997b; Endsley, 1996; Funk & Lyall, 2000; Hollnagel, 1999; Lyall & Funk, 1998; Parasuraman, Sheridan, & Wickens, 2000; Parasuraman & Riley, 1997; Thurman, Brann, & Mitchell, 1977; Wiener & Curry, 1980). It is not just a simple high-degree of automation problem, but involves human-automation interaction through the human-system interface (HSI).

The "ironies of automation" are still with us in modern systems, and automation still challenges operators, sometime resulting in serious consequences for safety. The 2009 Metrorail accident in Washington, D.C. (see Figure 2) occurred when the automatic impact avoidance system failed to detect an idle train on the track ahead due to a faulty track circuit; the operator was unable to stop the moving train before it collided with the idle train, despite applying the emergency brakes. Following this accident, as several accident investigation experts noted, investigators separately focused on either the malfunction of the computer system or whether the driver applied the brake on the speeding train, "…the discrete aspects of machine or human error; whereas the real problem often lies in the relationship between humans and their automation systems" (Vedantam, 2009). Additionally, this accident was subject to intense public scrutiny, bringing a high level of public attention to the challenges presented by humans interacting with automation.

[1] The reader should note that this report includes text that has been previously published in other U.S. Government reports written by at least one of the co-authors. To simplify the formatting of this report, text excerpted from other U.S. Government reports (i.e., O'Hara and Higgins, 2010; NUREG/CR-7126) is not quoted or otherwise delineated from other text. This is within the copyright granted to the U.S. Government, which states it has a nonexclusive, royalty-free license to publish, republish, or reproduce the work or to allow others to reproduce this work for U.S. Government purposes.



The Washington Post

Metrorail Crash May Exemplify Automation Paradox

By Shankar Vedantam
Washington Post Staff Writer
Monday, June 29, 2009

Figure 2. Washington Post article following 2009 Metrorail Accident, Washington, D.C.

Another recent example, also in 2009, that has received considerable discussion in the public media is the crash of Air France 447in the Atlantic Ocean. This accident typifies many of the problems crews face with automation. Icing of the airspeed sensors caused a loss of airspeed information and led to failure of the autopilot. Therefore, the pilots had to take over manual control of the aircraft; however, they did not have adequate training on manual control at high altitude during turbulence or training on how to use procedures for handling unreliable airspeed indications. Additionally, the stall alarm behaved in a manner that contradicted actual flight circumstances and confused the pilots. The official accident report stated:

> The occurrence of the failure in the context of flight in cruise completely surprised the pilots of flight AF 447. The apparent difficulties with aeroplane handling at high altitude in turbulence led to excessive handling inputs in roll and a sharp nose-up input by the PF. The destabilisation that resulted from the climbing flight path and the evolution in the pitch attitude and vertical speed was added to the erroneous airspeed indications and ECAM messages, which did not help with the diagnosis. The crew, progressively becoming de-structured, likely never understood that it was faced with a "simple" loss of three sources of airspeed information. In the minute that followed the autopilot disconnection, the failure of the attempts to understand the situation and the de-structuring of crew cooperation fed on each other until the total loss of cognitive control of the situation (BEA, 2012, p. 199).

In light of these persistent issues, designers and researchers continue to work to improve the means by which human and automation interact to accomplish plant functions and tasks. In some cases, functions and tasks are accomplished using varying levels of automation. In others cases, a function is performed primarily by personnel with automation assisting some aspects of the tasks to be performed. In other cases still, a function may be performed primarily by automation, with personnel performing some aspects of the task, such as to provide authorization to perform one subtask when a prior subtask is completed. However, little guidance is available to designers to implement such approaches to HAC. Thus, additional research is needed to identify successful approaches and to develop design guidance supporting their implementation. Currently, little such guidance is available to designers.

Research on HAC has been identified as a significant need in the nuclear industry and, in particular, as it pertains to aSMRs. In 2007, DOE published a study providing a technology roadmap on ICHMI to support DOE advanced NPP programs (Dudenhoeffer et al., 2007). Seven areas of research were identified as essential elements for advancing ICHMI technologies in NPPs to resolve the challenges and needs. One area was Human-Automation Interaction Models and Analysis Tools. It was defined as follows:

> Human factors must be a key consideration in any upgrade or paradigm of operation shift. This includes integration of plant automation, new information systems, new procedures, and any other aspect that changes the human machine

> interaction expectation. Current human activity within nuclear power plant operation and maintenance is studied by the U.S. NRC and industry itself. This topic addresses the development of new models of human-automation interaction based on emerging control technologies, such as automation that adapts to operator workload. Models should be defined and methods of analysis for allocation of functions, including dynamic allocation, should be formalized. The user interface requirements for each model should be specified. A test program should be included to evaluate concepts (Dudenhoeffer et al., 2007, p. 28).

In 2012, NRC published a study outlining human-performance issues related to the design and operation of SMRs (O'Hara, Higgins, & Pena, 2012). The study identified several issues related to HAC, one of which was entitled "High Levels of Automation for All Operations and its Implementation." Their findings emphasized automation as the key enabling technology for multi-unit operations. As crews manage increasing numbers of units, automation must take responsibility for tasks traditionally performed by operators. aSMRs are no exception and their degree of automation will be high as both normal and safety operations will be automated. The "automate everything that can be automated" philosophy often dominates programs for developing advanced reactors to improve their performance and decrease operational costs. However, as discussed previously, there is a complex relationship between automation and human performance, which often fails to confirm common-sense expectations. For example, one might expect that high levels of automation will lower workload; instead, it shifts workload and creates other human-performance difficulties (O'Hara, Gunther, & Martinez-Guridi, 2010). The authors suggested that flexible approaches to using different levels of automation in a single system should be explored. This is discussed in greater detail in Section 2.

Human-automation interaction also was identified in NRC's report to Congress on advanced reactor licensing and, in addition, research needed to support licensing (italics added for emphasis) was identified:

> The future designs will generally rely on passive rather than active safety features and may involve concurrent control of multiple modules from a common control room. In general, these designs will employ digital information and controls technology as opposed to the predominantly analog information and controls C technology used in the current fleet of operating nuclear plants. *These systems will provide the capability for increased automation that makes greater use of interactions between personnel and automatic functions. Automation can change the operators' role in monitoring, detection, and analysis of off-normal conditions, situation assessment, and response planning. Research is needed to determine the effect of these changes on operator safety performance and on plant safety* (DOE, 2010, p. 28).

Thus, the nuclear industry, from both design and regulatory perspectives, has identified significant issues related to HAC in development of future commercial NPPs in general and aSMRs in particular.

Given the public visibility of new plant construction and awareness of the "ironies of automation," it is imperative that research is performed to ensure that design of aSMR automation is based on sound scientific and engineering principles that support HAC, efficient performance, and safety.

1.4 Human-Automation Collaboration Research Effort

As shown in Sections 1.2 and 1.3, high levels of automation are a requirement for aSMR designs, and automation presents challenges to human operators. In order to ensure safe and productive aSMR plant operation, research is needed to determine how to best automate the plant functions without causing human performance detriments. Given the design specifications aSMR vendors have provided, it is anticipated that the means by which humans and automation interact will be significantly different from typical reactor control room operations. No blueprint is in place for how this kind of HAC will best be carried out. Therefore, it is imperative that empirical research be conducted to determine how to maximize the use of automation to the greatest degree, without adversely affecting safety, efficiency, and performance of the human operator. This is a delicate balance and only investigation by various experimental designs will produce the key to best design implementation.

The HAC research project is investigating the consequences of the allocation of functions between the operators and automated systems. The research effort addresses the question of how to best design the collaboration of personnel and automation to capitalize on the strength of each. Hence, focus is on what the collaboration level should be and how it should be implemented to have the greatest positive impact on overall plant performance and safety. It also is developing a model of HAC, which will support aSMR designers when evaluating their proposed approach for the conduct of operations in terms of how humans and automation are collaborating. The research results will inform integration and cooperation between plant staff and automation, with the purpose of maximizing productivity and safe operations of aSMRs. One key research goal is providing a technical basis to support reduction in aSMR operations and maintenance costs through reduced staffing per unit, which is made possible by greater integration and cooperation between plant staff and automation. Additionally, aSMR vendors will be able to use the results of this research effort to inform development of the technical basis they will develop for their licensing case in submittals to NRC.

An initial step to establish a technical basis for HAC in aSMR designs is creation of a HAC model, with a focus on interaction between humans and automation. This HAC model is one means to identify the gaps that need to be bridged in order to establish a technically sound future state of practice for aSMR plants. As an initial step, the research team conducted a review of research literature related to human factors and automation. The path of moving the state-of-practice forward consists of a selection process part and an investigation part. For each identified knowledge gap, the researchers will ask a series of questions to determine if the particular gap is relevant to the research effort and the aSMR industry (the the selection process). If the answer to the five questions is "yes," the gap should be investigated further via empirical studies (the investigation part). The research team will reiterate the process of asking these questions thoughout the span of the project. The questions are as follows:

1. Is there a knowledge gap?
2. Can it be addressed with experiment(s) (i.e., can it be demonstrated empirically)?
3. Is it relevant to the aSMR field?
4. Can the result be generalizable to most aSMR designs?
5. Can the results also be the technical basis for development of HAC-related engineering procedures and aSMR guidance?

The overarching question that current research efforts aim to answer is how multi-modular aSMRs can be safely and economically operated with a reduced operating crew that controls the aSMR through a greater reliance on digital automation technologies. In order to address this question and the identified gaps, the research team will conduct research to develop potential solutions.

The principal focus of this report is to summarize the current state-of-practice in HAC, identify a preliminary HAC model (which will be updated as this research continues), and identify research that needs to be addressed to accomplish the research project's overall objective. The HAC model will, among other things, serve as the basis for further model development and development of engineering procedures and guidance for use by designers in developing human automation collaborations.

1.5 Organization of this Report

The remainder of this report details the literature review and to-date development of a HAC model. The project team analyzed research findings pertaining to the effects of automation on human performance to identify lessons learned, best practices for supporting performance, and unresolved issues. Information from a variety of sources was used, including the basic literature on automation and human performance, consisting of papers from research journals, technical conferences, and operational experience with automation in the nuclear and other industrial domains. The literature review and conclusions are described in Section 2.

An important step in creating aHAC model is to develop a characterization of automation and performance. Characterization provides an important structure needed to develop and organize the design procedures and guidance later in the research effort. The initial model of HAC, including characterization of automation design and HAC design characteristics, is described in Section 3.

Section 4 discusses the path forward based on development of the initial HAC model. Preliminary research issues and proposed research topics are described.

2. STATE-OF-PRACTICE IN HUMAN-AUTOMATION-COLLABORATION

This research effort aims to identify the factors that designers must consider in order to facilitate effective HAC. Human factors researchers have explored a wide variety of factors that influence human collaboration with automation. This section contains a review of the existing literature on HAC and serves as a starting point for the framework for HAC and an initial description of the current state-of-practice in HAC. The literature is divided into two subsections: Subsection 2.1 describes how researchers have characterized what automation is and what it does. Subsection 2.2 describes empirical research that investigates how the factors identified in Subsection 2.1 affect performance in HAC.

2.1 Characterization of Automation

In order to frame the discussion of HAC, it is helpful to first understand what automation is and how researchers typically characterize some of the important factors related to automation design. Automation is part of the plant's instrumentation and control (I&C) system. Modern complex digital I&C systems afford a great deal of functionality that is vital to the plant's performance and safety. New reactor designs use a diversity of digital I&C architectures. Figure 3 is a generic block-diagram of a representative I&C system identifying the generic components that it encompasses. The blocks represent the types of digital instrumentation component types required to process a signal from the system to its end use. The arrows represent the communication links between each component. The signals are processed through a "data processing and error checking unit" (sometimes termed an I/O unit), and transformed to an appropriate format that might include some high- level calculated parameters. These input parameters then pass through a communication link to a computerized logic unit, often containing internal software that processes the parameters, and compares them to a set of criteria to decide if a series of systems and components must be actuated. Thereafter, the actuation signal is transmitted to actuator devices that complete the desired operation. The communication links can be provided from both the data- processing and actuation-signal logic units through a separate communication bus to generate information displays in the control room. Displays for the operators can be fed from any of these components, although the logic units and signal processors are the commonest links to the control room displays.

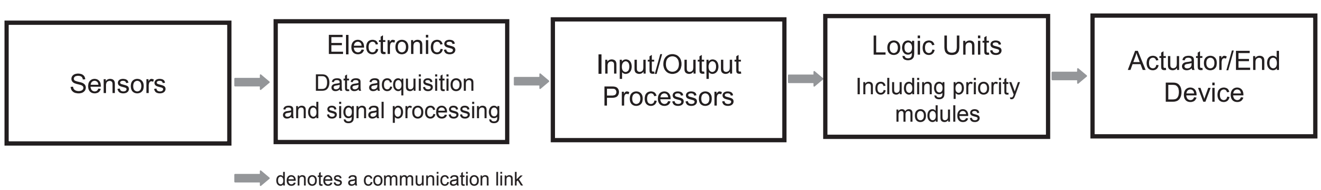


Figure 3. Digital I&C system components.

Figure 4 illustrates the I&C system from a functional point of view. This way of describing the I&C system, and its subsystems, was developed for use in the I&C roadmap for the DOE's advanced NPP programs (Dudenhoeffer et al., 2007).

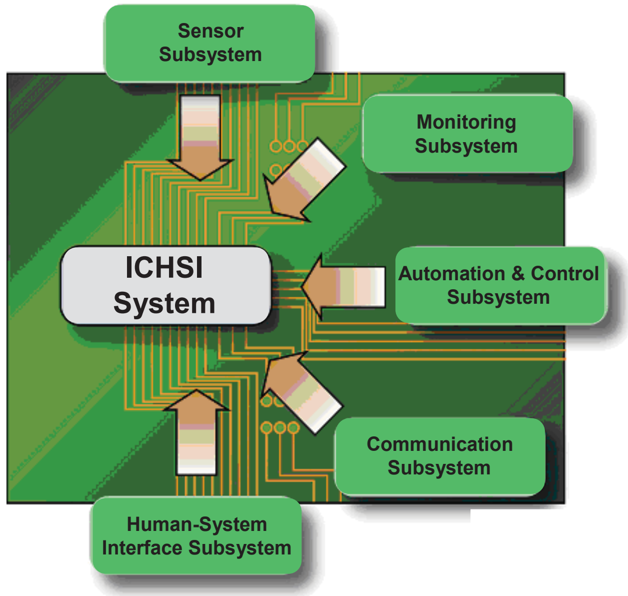


Figure 4. I&C subsystem representation employed by the DOE for advanced NPPs (Dudenhoeffer et al., 2007).

The subsystems are described as follows:

- *Sensor subsystem* – Nearly every plant process uses some form of physical measurement taken by sensors that detect parameters in the plant, such as neutron flux, temperatures, pressures, flow, valve positions, electrical current levels, and radiation levels. Some new nuclear-energy production technologies employ new, different types of sensors and instruments to measure physical processes. In some reactor designs, they include electronic sensors with imbedded software that are required to work in high-temperature environments and measure and analyze process parameters quite different from those in today's operating light water reactors.
- *Monitoring subsystem* – These subsystems monitor the signals and other information produced by sensors and evaluate that information to determine whether and what type of response is needed. They can contain sophisticated diagnostic and prognostic functions. Diagnostics refers to techniques for identifying and determining the causes of deviations or faults in the plant's systems or processes. Prognostics refers to methods for using sensor data to estimate the rate of physical degradation and the remaining useful life of systems, predicting time to failure, and applying this information to more effectively control processes.
- *Automation and control subsystem* – Digital control systems offer the ability to implement more advanced control-algorithms than those presently used in U.S. NPPs that rely primarily on single-input, single-output, classical control schemes to automate individual control loops. Advanced control schemes include matrix techniques for optimal control, nonlinear control methods, fuzzy logic, neural networks, adaptive control (a control that modifies its behavior based on plant dynamics), expert systems, state-based control schemes, and other schemes combining multiple control methods. Applying these advanced techniques will assure a more integrated control of plant systems and processes (versus separate, non-interacting control loops) and greater complexity.

- *Communications subsystem* – A variety of communication systems assure information flow throughout the I&C system and to devices being monitored and controlled; they may include wireless technology. A classical I&C architecture provides point-to-point wiring of measured variables to the monitoring and control systems. The communications subsystems for a modern I&C system are configured into a flexible network architecture' their greatly expanded functionality enables "smart" transducers to signal their service condition to the engineering staff.

In this view, the role of automation in the I&C system and plant operations is more readily apparent. For the purposes of our project, a more detailed characterization of automation is needed. This project takes the view that humans and automation

Sheridan (2002) defined automation as (a) the mechanization and integration of the sensing of environmental variables (by artificial sensors); (b) data processing and decision making (by computers); and (c) mechanical action (by motors or devices that apply forces in the environment) or information action by communication of processed information to people. At the simplest level, an automation system is designed to accomplish a goal that can be predetermined by designers or set by operators, based on their current needs. The automatic system processes inputs from the plant and operators to meet the goal (Figure 5).

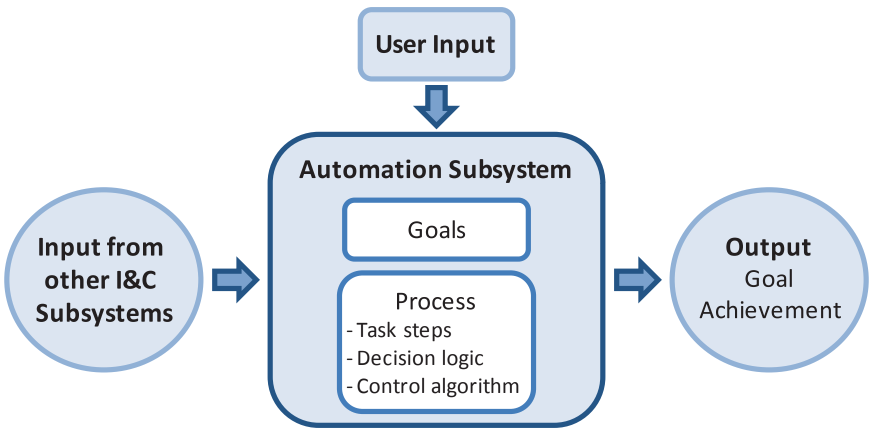


Figure 5. Overview of an automation system.

Because automation is applicable to many aspects of the plant's operations, from analyzing procedural steps to controlling the plant's systems, the specific processes used to accomplish automation's goal vary depending on its particular usage. Thus, modern approaches to automation emphasize the value of multi-agent teams monitoring and controlling complex systems (Christoffersen & Woods, 2002; Hollnagel & Woods, 2005; Woods & Hollnagel, 2006). The teams consist of human, software, and hardware elements working together, sharing responsibilities, and shifting responsibilities to support the plant's overall production and safety missions (see Figures 5 and 6). In this context, the term "agents" refers to who/what is performing an activity (i.e., agents are entities that do things). An agent will monitor the plant to detect conditions, indicating that a function must be performed. An agent will assess the situation and plan a response. Having established the response plan, it must be implemented by sending control signals to actuators. The agent will continue monitoring the activity to determine that the function is being accomplished and to plan again if it is not. Finally, the agent must decide when the function is completed satisfactorily. Human or machine agents can perform any one or all of these activities. Uhrig, Gao, and Tsoukalas (2004) suggest that for advanced NPP designs, multi-agent systems will be the first line of defense against degraded conditions, assuring continuous surveillance and predictive diagnosis.

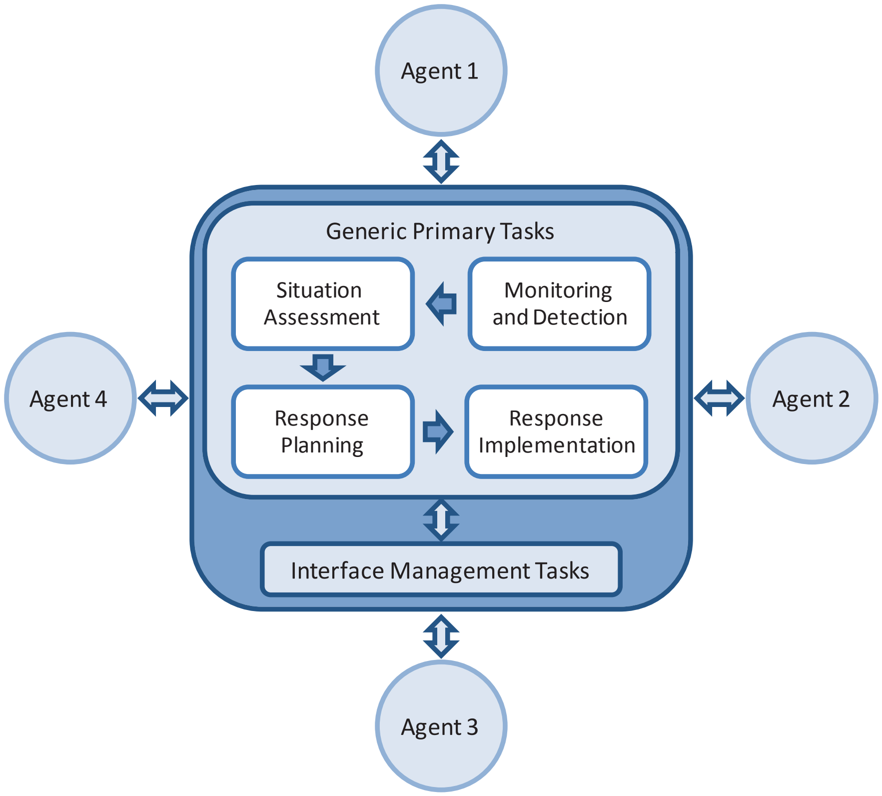


Figure 6. Multi-agent teams to accomplish plant functions and operator tasks.

In recent research, O'Hara and Higgins (2010) have identified a six-dimensional characterization of automation: levels, cognitive function, processes, modes, adaptability, and reliability. Each automation dimension is described in the following subsections. Note that a "type" of HAC is defined by its characterization along these dimensions. An adapted version of this six-dimensional model serves as a preliminary automation characterization for the HAC research effort. The following subsections describe the dimensions of cognitive function, level of automation (LOA), reliability, degradation, process mode, and adaptability.

2.2 Automation Design Dimensions

This section defines the factors that researchers use to characterize automation and HAC. The factors presented in this section were initially identified by O'Hara and Higgens (2010). The contribution of these factors to HAC is discussed in the context of the empirical literature in Section 2.3.

2.2.1 Function of Automation (Cognitive Function)

It is important to note that the word 'function' is term that is used in psychology, engineering, and other fields, and is defined and used in a variety of ways. Psychological and human factors literature uses the word function to refer to information processing functions that operators perform as part of executing their tasks in the control room such as monitoring, detecting problems, gathering information, making a diagnosis, generating plans or alternative strategies, making decisions, and implementing the chosen action. Automation also serves these same information processing functions. These are the low-level functions that must be accomplished to successfully execute higher-level plant functions, regardless of whether they are performed by the operators or automation. However, the word function also has specific meaning to designers related to the *plant* functions (e.g., produce electricity, steam, and process heat) and *system* functions (e.g., provide core cooling). Human factors engineering also uses the word function when describing a standard human factors engineering process called function allocation (i.e., how decisions are made regarding whether a plant function is allocated to the human operator, the automated systems, or both). To avoid confusion regarding which type of function is being discussed, this report will use plant function or system function when referring to the plant or system functions and cognitive function when referring to the lower-level information processing functions, whether they are performed by the operators or automation. It may seem counterintuitive to call these low-level functions cognitive functions when they are performed by automation (because automation does not "think" in a strict sense); however, it can be argued that, in most cases, automation executes these cognitive functions in place of the operator, thus serving as the cognitive function that the operator would normally fulfill.

The list of cognitive functions varies slightly depending on the researchers and model of cognition. Parasuraman et al. (2000) characterized the following four main cognitive functions:

6. Information acquisition (e.g., the gathering of process information)
7. Information analysis (e.g., calculations)
8. Decision and action selection (e.g., evaluating step logic, conditions, or providing recommendations)
9. Action implementation (taking a control action such as opening a valve).

Endsley and Kaber (1999) identified four slightly different generic cognitive functions that are applicable across domains:

1. Monitoring: scanning displays and indications to perceive system or process status
2. Generating: formulating options or strategies to achieve operational goals
3. Selecting: making a decision on a particular option or strategy
4. Implementing: carrying out the selected option.

The two lists of cognitive functions are very similar, yet differ slightly. Parasuraman et al. (2000) combine generating response options with decision making and have a separate function of information analysis, whereas Endsley and Kaber (1999) separate generation of options from decision making and incorporate information analysis with monitoring. Both of these models have received wide use, in part, because both models are generic enough to apply across industries.

However, in developing an HAC model for the aSMR domain, it is important to use a model that is more tailored to the specific domain due to the unique design characteristics and the anticipated advanced automation of aSMRs. O'Hara et al. (2010) have proposed a new taxonomy that is more relevant to the aSMR domain than either of the Endsley and Kaber (1999) or Parasuraman et al. (2000) models. Specifically, the O'Hara et al. (2010) taxonomy of cognitive functions:

- Is widely used in the nuclear industry
- Is representative of the types of process control task activities that NPP operators commonly engage in
- Includes interface management tasks, which are not typically considered in the other taxonomies
- Is more directly associated with issues related to automation, such as failure to monitor automation and loss of situational awareness (SA).

For these reasons, the research team will utilize the O'Hara et al. (2010) taxonomy of cognitive functions in developing a model of HAC for aSMR designs. This taxonomy is briefly described in the following paragraphs.

In fulfilling their responsibilities, agents perform primary tasks (i.e., cognitive functions), which 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). O'Hara et al. (2010) proposed a set of cognitive functions that consider all of the factors. These cognitive functions 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 an NPP, 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, 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.

A potentially important facet of cognitive function that is not explicitly represented above is the fact that some processes may be discrete and others may be continuous. For example, a single control action (or a series of control actions) could be defined as discrete processes, while monitoring parameters are considered continuous processes. There may be important implications for allocating function to either humans or automation, because they have different capabilities when it comes to continuous versus discrete processes.

2.2.2 Level of Automation

LOA describes the amount of automation used in a given situation. There are nearly as many taxonomies of LOA as there are researchers who investigate it; however, each taxonomy typically varies from fully manual (the human operator does everything) to fully automatic (the automatic system does everything), with intermediate levels typically including some collaboration between automation and human. In 1992, Sheridan defined three global levels of automation (O'Hara et al., 2010):

1. Manual control (all control is accomplished by humans);
2. Supervisory control (some or all of the control is performed by the computer, but the human supervisor can assert control); and

3. Fully automatic control (all control is automatic and the human cannot interfere in the process except, perhaps, to terminate it).

As technology has evolved, more fine-grained distinctions between these levels have evolved. This section describes and discusses the extant taxonomies of LOA.

Billings (1991, 1997a) proposed one of the best-known frameworks for LOAs based on his work in the aviation industry. Table 1 illustrates how Billings (1997a) characterized the division of responsibilities for functions and tasks across humans and automation.

Table 1. Billings' levels of automation.

Level	Role of Automation	Role of Human
Autonomous Operations	Fully autonomous operation. Human not usually informed. System may or may not be capable of being disabled.	Human generally has no role in operation and monitoring is limited.
Operation by Exception	Essentially autonomous operation unless specific situation or circumstances are encountered.	Human must approve of critical decisions and may intervene.
Operation by Consent	Full automatic control under close monitoring and supervision.	Human monitors closely, approves actions, and may intervene.
Operation by Delegation	Automatic control when directed by human to do so.	Human provides supervisory commands that automation follows.
Shared Control	Automatic control of some functions/tasks.	Human controls some functions/tasks.
Assisted Manual Control	Primarily manual control with some automation support.	Human manually controls with assistance from partial automation.
Direct Manual Control	No automation is used.	Human manually controls all functions and tasks.

While early taxonomies of LOA treated it as independent from cognitive function, several researchers have acknowledged that the consequences of LOA for HAC are largely dependent on which cognitive functions are automated. As a result, these researchers have combined level and cognitive functions into a single taxonomy. 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 four cognitive functions described in Section 2.1.1 (i.e., monitoring, generating, selecting, and implementing) to the human operator, automation, or a combination of the two to develop the 10 levels of automation shown in Table 2 (Endsley & Kaber, 1999). Note that as the LOA increases, automation takes over progressively more of each cognitive function.

Table 2. 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

Levels of Automation	Roles: Monitoring	Generating	Selecting	Implementing
(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

Each of these levels are described in more detail as follows (Endsley & Kaber, 1999):

1. Manual control: The human performs all cognitive functions.
2. Action support: The system aids the operator in performing the selected action; however, some operator control actions are required.
3. Batch processing: The human generates and selects the options to be performed and then turns the actions over to the computer to carry out. The automation at this level is primarily in terms of implementing actions.
4. Shared control: Both the human and computer generate possible options. The operator makes the decision on which option to implement, and then shares responsibility with the system for carrying out the action.
5. Decision support: 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.
6. Blended decision making: The computer generates a list of decision options, selects one, and implements the action if the operator consents. The operator may approve or disapprove the computer's choice and may provide her or his own option.
7. Rigid system: This system presents a limited set of actions to the operator, who must choose an option from the set and cannot generate any other options. The computer then fully implements the chosen action.
8. Automated decision making: The computer generates a list of options to which the operator may add suggestions. The computer then makes a decision and carries out the chosen action.
9. Supervisory control: The computer generates options, makes decisions, and carries out the chosen actions. The operator's role is primarily to monitor the system and intervene only when necessary.

10. Full automation: The computer carries out all cognitive functions. The human operator is out of the loop and cannot intervene. This is a fully automated system, where human intervention is not considered to be necessary.

Parasuraman et al. (2000) also proposed an LOA taxonomy that incorporated cognitive functions. Their taxonomy includes the four cognitive functions described in Section 2.1.1 (i.e., information acquisition, information analysis, decision selection, and action implementation). Based on these four functions, Parasuraman et al. developed their own 10 LOAs (Table 3).

Table 3. Levels of automation (Parasuraman et al., 2000).

Levels of Automation of Decision and Action Selection		
High	10	The computer decides everything, acts autonomously, and ignores the human.
	9	The computer executes automatically and informs the human only if the computer decides to.
	8	The computer executes automatically and informs the human only if asked.
	7	The computer executes automatically, then necessarily informs the human.
	6	The computer allows the human a restricted time to veto before automatic execution.
	5	The computer executes the suggestion if the human approves.
	4	The computer suggests one alternative.
	3	The computer narrows the selection of decision/action alternatives to a few.
	2	The computer offers a complete set of decision/action alternatives.
Low	1	The computer offers no assistance, the human must take all decisions and actions.

Their levels differ somewhat from Endsley and Kaber (1999); however, there are some overlaps (e.g., Parasuraman et. al Level 5 corresponds to Endsley and Kaber LOA 6 [blended decision making]). Endsley and Kaber's model does not include levels that specify when and how automation notifies the operator of actions taken, as is the case with Parasuraman et al.'s Levels 7, 8, and 9. Endsley and Kaber's model also includes levels that detail more of the operator's supervision of the system (e.g., LOA7, LOA8, and LOA9) than Parasuraman et al.'s model.

Parasuraman et al. (2000) provide a model in which they advocate different LOAs, depending on the cognitive function and other automation or task characteristics (Figure 7).

Parasuraman et al. (2000), Endsley and Kaber (1999), and Billings' (1997a) LOAs have all been used widely in empirical research. However, it remains to be seen which taxonomy best fits automation that will be employed in aSMR NPPs.

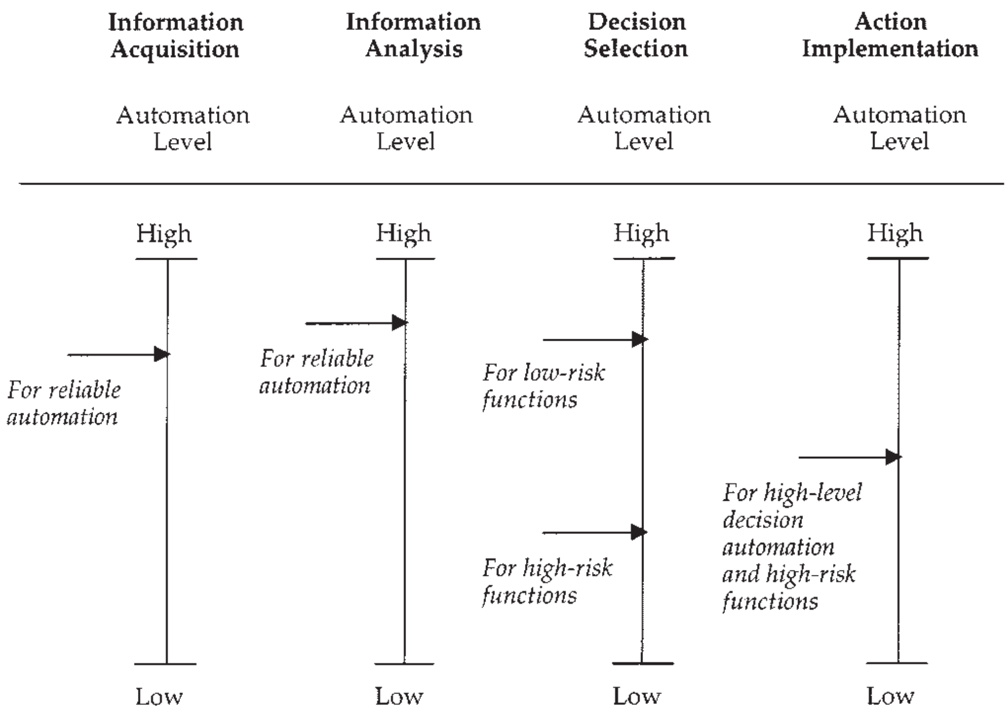


Figure 7. Recommended levels of automation for different cognitive functions (Parasuraman et al., 2000).

The majority of LOA taxonomies have been developed in the context of industries other than nuclear power (e.g., aviation). To ensure relevance to the nuclear industry, O'Hara et al. (2010) adapted existing taxonomies to account for the types of automation used in the nuclear industry (see Table 4).

O'Hara et al.'s Level 3 corresponds to Endsley and Kaber's LOA6 and Parasuraman et al.'s Level 5, and O'Hara et al.'s Level 4 corresponds to Parasuraman et al.'s Level 6 and Endsley &and Kaber's LOA8 and LOA9 (Endsley and Kaber's model breaks down into finer detail than O'Hara et al.'s levels).

Though this taxonomy was tailored to the type of automation that would likely be employed in 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 aSMRs; therefore, this taxonomy may need to be updated to accurately reflect the aSMR context.

Table 4. Preliminary levels of automation for nuclear power plant applications.

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, under close monitoring and supervision	Operators monitor closely, approve actions, and may intervene to provide 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

Research on the impacts of LOA on operator performance (see discussion in Section 2.2.3) indicates that it is not possible to determine the impact of LOA on operator performance without knowing which tasks the automation is controlling. In this manner, level and cognitive function are not independent of each other, which is reflected in research where level and cognitive function are confounded: higher levels of automation have the automation conducting more of the cognitive functions that operators typically are responsible for in non-automated settings. Both Endsley and Kaber's (1999) and Parasuraman et al.'s (2000) LOA taxonomies illustrate this point. As LOA increases, not only does the amount of automation increase, but automation takes over increasingly more of the operator's role and activities.

2.2.3 Reliability

Automation is imperfect. In fact, all engineered systems have less than perfect reliability. Automation's reliability is defined as how well automation accomplishes its task. Reliability is a characteristic of properly functioning automation. For example, an alarm system's reliability is a function of the alarm conditions correctly detected, false alarms (signaling an alarm when no alarm condition exits), and missed alarms (failing to signal an alarm when an alarm condition does exists). This relationship is illustrated in Table 5.

Table 5. Automation reliability performance matrix.

Situation	Event Occurs	Event Does Not Occur
Alarm Triggers	Correct Performance	False Alarm
Alarm Does Not Trigger	Miss	Correct Performance

This type of relationship characterizes many types of automation. The reliability of such a system can be expressed as (Cullen, Roders, & Fisk, 2012):

$$\text{Reliability} = \frac{\text{Hits}}{\text{Hits + Misses + False Alarms}}$$

However, when automation's tasks are complex (as is the case for many decision support systems), defining the measures of reliability is more involved. Further, automation's reliability may differ across different contexts of use or modes of operation.

An important feature of automation's reliability is degradation. Automation is a part of the overall instrumentation and control system, which is composed of four subsystems: sensor, monitoring, automation and control, and communications subsystems. Automation depends on the other subsystems to function properly. Problems arising in the instrumentation and control infrastructure can lead to degradation or failure of any aspect of automation (O'Hara et al., 2010). Thus, for example, loss of sensors or significant delays in the transmission of information along the data highways can cause degradations of any aspect of automation.

For purposes of the HAC research effort, degradation and failure fall on a continuum from degradations resulting in minor loss of functionality to complete loss of functionality (or an automation failure). 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. For the purposes of this section, we are not considering automation degradation and failure s that are due to incorrect operator inputs, configuration, or usage of automation.

Degradation and failure can lead to two types of problems for operators:

- Automation does not do what it is supposed to do when it should do it
- Automation does something that it is not supposed to do such as causing abnormal operating conditions due to erroneous automatic action or providing erroneous information.

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 are potentially subject to human performance issues.

2.2.4 Process

Automation uses input from the plant (and perhaps the operator) and processes the information to accomplish a goal. These processes are an important aspect of automation in that they are the means by which automation performs its tasks. Automation processes can include control algorithms, decision logic (such as the use of Boolean logic), and virtually any other type of information processing routine suited to its tasks (O'Hara et al., 2010).

2.2.5 Mode

Automated systems may have different modes of operation. Modes define sets of mutually exclusive behaviors that describe the relationship between input to the automation and the response to it (Jamieson & Vicente, 2005). A system can have multiple modes, but only one is active at a time. Modes do not imply differing levels of automation; rather, they involve performing the same function in different ways. Modes are beneficial in providing the capacity for a system to do different tasks, or to accomplish the same task using different strategies under changing conditions

A global positioning system device is a simple example of modes. After the user specifies a destination, the global positioning system device automatically plans the best route. Users can select driving mode or pedestrian mode. In a city environment, where there are many one-way streets, the route suggested by each mode may be completely different. In driving mode, the one-way streets constrain the route selected; in pedestrian mode, one-way streets have no impact on the route selected. Therefore, the task is the same, but the solution depends on the mode selected (O'Hara et al., 2010).

2.2.6 Adaptability

A system can be designed such that the human or machine agent responsible for performing a task always is the same (i.e., the so-called static allocation). Alternatively, a task can be performed either by automatic systems or by personnel based on situational considerations such as the operator's overall workload. For example, automation may assume control over lower priority tasks when the operator's workload increases to a level where all current work becomes difficult to complete. This approach ensures operators can focus their attention on high-priority tasks because their workload levels remain within acceptable limits. A simple example is alarm reset, when during a major plant disturbance workload is very high and many alarms are coming in, operators can reallocate the alarm-reset task from manual to automatic. When human or machine agents can flexibly perform tasks, automation is said to be adaptive (O'Hara et al., 2010).

A key consideration for adaptive automatic systems is the "triggering" condition (i.e., the condition that causes the adaptive automation shift), and which agent makes the change: the human operator, the automation, or whether the system is flexible enough that the operator or the automation can initiate a change to the automation.

Adaptability of automation is an issue that is not entirely separate from the LOA or the cognitive functions for which the automation is responsible. The concept of adaptive automation involves changing the LOA (and/or the cognitive functions allocated to the automation) depending on the specific circumstances of the situation. The triggers that initiate a change to LOA or cognitive function can vary. Parasuraman et al. (1996) and Yoo (2012) have identified five main categories of techniques for classifiying triggering conditions:

- critical events (that will change demands on ops – like an emergency operating procedure initiator)
- operator performance measurement
- operator physiological assessment
- modeling
- hybrid methods combining one or more of the above

It should be noted that if triggers are to be made on operators' states, such as poor SA or high workload, a means to measure those states in real-time systems will be needed (Salmon et. al. 2008).

2.3 Review of Empirical Literature on Human-Automation Collaboration

As the first step in developing a model of HAC, researchers reviewed empirical literature on human automation interaction. The researchers identified the prominent human performance characteristics that are investigated in the human factors literature on HAC. The researchers then investigated what conclusions can be drawn from the existing literature regarding how the automation design dimensions affect those human performance characteristics.

Before going into a detailed review of the human factors literature, it is worth pointing out the distinction this research is making between 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 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.

Human-system interaction involves analysis and design of the interaction between people (i.e., users) and technological systems. The goal is to design the human-system interaction in such a way that the user's performance and, as a result, the overall human-system's performance is optimized. This can be measured with a variety of metrics, including time to learn, speed of performance, rate of errors by the users, retention over time, and subjective satisfaction (Schneiderman, 1998).

With respect to human-system interface methods and models to evaluating the design of user interfaces or HSIs, there are a number of widely accepted approaches and methods, including the goals, operators, methods, and selection rules model (Card, Moran, & Newell, 1983) and user-centered design (Norman & Draper, 1986). In general, these methods model (a) the human in terms of their cognitive, physical, and social attributes, (b) the technological system in terms of its capabilities and limitations, and (c) their dynamic interaction as they work together to perform various tasks to achieve a defined goal. More importantly, because these methods often focus on the design of user interfaces that allow the user to effectively use the technology, human-system interaction is considered a subset of HAC.

2.3.1 Performance Characteristics in Human-Automation Collaboration

Though a large portion of the human factors' literature addressing HAC focuses on the human performance consequences of automation (Endsley, 1996, 1997; Endsley & Kaber, 1999; Endsley & Kiris, 1995; Jou, Yenn, Lin, Yang, & Chiang, 2009; Kaber & Endsley, 2004; Lin, Yenn, & Yang, 2009, 2010a, 2010b; van de Merwe, Oprins, Eriksson, & van der Plaat, 2012), the overall characterization of

HAC needs to consider the consequences to overall system performance. Overall system performance can be characterized as a combination of each individual agent's (human or automatic) individual performance in accomplishing the overall plant function in addition to each agent's ability to compensate or recover from another agent's failure. Thus, the characterization of overall system performance can be broken down into two levels: integrated system performance and the individual performance of automation and the human (which also encompasses the quality of HAC). Even though overall performance of automation is an important factor in system performance (and is likely to be of great importance to designers), we will only consider it here as it applies to human performance and human interaction with automation.

Automation and personnel work together to accomplish a mission, function, or purpose. The overall accomplishment of that plant function reflects the success of HAC. Therefore, measures of plant function accomplishment ultimately are the bottom line from an operations standpoint. Measures of plant function performance are scenario specific. While they are an important criterion for success, these measures typically are not diagnostic. They do not provide an indication as to whether successful plant function accomplishment was achieved in an undesirable way (such as with poor SA or high workload) that may bring into question the reliability of performance or why plant function accomplishment failed.

Even though overall system performance ultimately is the measure by which HAC will be deemed successful, it also is essential to consider the human performance consequences of automation. As stated above, a plant function that has been performed satisfactorily may not reflect satisfactory performance of the individual agents involved in executing the plant function (i.e., the human and/or automation may have failed to perform individual tasks). The majority of empirical studies investigating human-automation interaction characterize human performance in terms of the following dimensions: objective task performance (which is situation dependent), SA, and use of automation. These dimensions characterize the human performance aspects of HAC, and are often affected by the automation design dimensions reviewed in section 2.2. Thus the review of the HAC literature is organized by human performance dimension and discussed in the context of the specific automation design dimensions that are investigated.

2.3.2 Objective Task Performance

Many researchers have investigated the effect of automation's reliability on performance (both human performance and system performance). Wickens and Dixon (2007) reviewed the findings of 20 studies that used automation's reliability as an independent variable to quantify its effects on task performance. The automation in these studies supported the generic tasks of monitoring/detection and SA (not response planning or implementation). A regression analysis of the data from these studies indicated the following:

- There was a "strong linear function," relating reliability and performance.
- Below a reliability of 0.70, providing automation led to poorer performance than no automation.
- The effect of reliability on performance was stronger in high-workload conditions.

To provide a richer understanding of these general findings, the researchers examined some specific studies addressing how reliability affects performance. While these studies are from a variety of industrial domains, most of the tasks are target detection–target recognition tasks.

Metzger and Parasuraman (2005) examined decision aid reliability and its relationship to using automation, with two reliability levels: high and low. Reliability levels were not quantified. The participants were professional air traffic controllers. The aid supported detection of air traffic conflicts. The researchers assessed task performance and workload. They found that when the aid was very reliable, task performance improved, and workload was lowered. However, when automation was less reliable, conflicts were missed and manual detection was better. The authors suggested applying automation of lower reliability to support less important tasks, thereby keeping operators available to address the more important tasks.

Skitka, Mosierer and Burdick (1999) compared simulated flight performance with and without an automated decision aid that monitored specific gauges and recommended actions to be taken when the gauge's readings entered the red zone. The participants, all students, were told the gauges were always 100% accurate, but that automation did make mistakes. When working properly, automation increased task accuracy and reduced errors. When automation failed to monitor properly, operator errors rose relative to the no automation condition (i.e., errors of omission [missing an event that was not detected by the aid] and of commission [doing what the aid said even when it contradicted their training and the information in the gauges] were higher). The authors suggested that in trying to reduce cognitive effort, operators tend to accept what the decision aids tell them.

Ruff et al. (2004) explored the impact on task performance, workload, and trust in automation of unmanned aerial vehicle team size with two different LOAs and two levels of automation reliability (low and high). Sixteen participants controlled unmanned aerial vehicle teams of two or four in a planning and targeting task. The automation planned new routes and identified targets, then either waited for user input to proceed (operation by consent) or waited for a time and then continued unless the participant gave a stop command (operation by exception). Performance measures included targeting task performance, workload, and trust. The low reliability of automation decreased task performance and trust, but had no effect on workload.

Goh, Wiergmann, and Madhavan (2005) evaluated the effects of reliability as an aid (70 vs. 90%) and the type of cue (direct vs. indirect) on the competence of students in identifying targets in a security luggage screening task. The direct cue was a green circle around a suspect target in the security display; the indirect one was a green border around the display suspected to contain a target. Their study also examined the participants' performance after an automation failure. Target identification was better with the direct cue than the indirect cue. Performance was better with the aid that was 90% reliable compared with the one that was 70% reliable. However, performance with the 70% reliable aid was not significantly better than that of a control group lacking an aid. The authors concluded that the 70% reliable aid did not sufficiently support performance and participants did not rely on it.

Dixon and Wickens (2006) examine the effect of the different reliability levels of an automatic alerting system. Pilots flew unmanned aerial vehicles and the system provided auditory alerts for system failures, route changes, and other mission updates. The reliability of the auto-alert system was either 100% detected (15 correct alarms); 67% detected (10 correct alarms, 5 false ones); or 67% detected (10 correct alarms and 5 that the alert system missed). A manual condition (no auto-alert system) was included. They found that detection accuracy and response time worsened with the automation's declining reliability.

de Visser & Parasuraman (2007) conducted a study in which 12 student participants performed a simulated target recognition task with unmanned ground vehicle teams of three or six. An automated target recognition system with three levels of reliability was used: low, medium, and high. They employed a variety of performance measures, including target detection, SA, workload, and trust in automation. For comparison, they assessed the target recognition performance of the user alone and the automated target recognition alone. Performance of the joint human- automated target recognition system was better than performance of either agent by themselves. Thus, even under the condition of low reliability, the automated target recognition system supported overall task performance. Reliability of automation affected trust and the task performance. Both decreased with lower reliability. SA and workload were unaffected by reliability.

In de Visser and Parasuraman's (2011) Experiment 1, student participants performed a target detection task under two task loads in a high-fidelity, multi-unmanned vehicle simulation. A target detection aid was provided in one of three levels of reliability. The study showed that even extreme levels of imperfect automation can still be beneficial to the overall system performance. In their study, human-automation performance at a target detection task was better than either automation or operators alone. Each agent is imperfect and their interaction led to the highest performance. These studies generally support the findings of Wickens and Dixon (2007) presented at the beginning of this section.

While higher reliability improves task performance, it does not improve an operator's ability to detect automation failures. In fact, the higher the reliability of automation, the less likely it is that operators will recognize when it fails (Dixon & Wickens, 2006; Wickens et al., 2010). In fact, this is one of the ironies of automation identified by Bainbridge (1983) many years ago.

In summary, highly reliable automation improves task performance, but not the detection of automation failures. When operators know the actual reliability of the system, they can make use of that knowledge to more properly adjust their automation usage. Based on the research above, we hypothesized the relationship as follows. As automation becomes less reliable, its support for task performance becomes less and performance declines. At some reliability threshold (perhaps.70 for monitoring, detection, and SA task automation), 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.

Objective task performance of a HAC also depends largely on the LOA and the function that is automated. Though the dimensions of level and function are discussed separately in section 2.2, they are discussed together in this review because very few empirical studies have invested them separately.

Prevot et al. (2012) tested the effect of a newly designed, highly automated system for air traffic control on overall human system performance. They varied the LOA (i.e., fully automated, interactive, or manual) and the level of air traffic. They found that the best performance (as measured by operator workload, time to resolve conflicts, and the number of loss of separations) was obtained when the tasks were executed in an interactive manner, rather than a fully manual or fully automatic manner. Specifically, the best performance was obtained when the human operator was still involved in the high-level decision making. Additionally, they found that as the level of traffic increased, the impact of interactive automation increased. These results indicate that it may be preferable to include operators in the loop of the high-level decision making, even in situations that require high levels of automation for most tasks.

One study found an interaction between the cognitive function that is automated and the reliability of automation. According to Rovira, McGarry, and Parasuraman (2007), the reliability of automation differentially affected acquisition of information and decision-making cognitive functions. Student participants identified targets as part of a simulated command-and-control operation. They had to identify the most dangerous target, deciding on an engagement strategy (i.e., identify which "friendly" resources should be used to attack the enemy). The task was performed manually or with an automated aid that supported information gathering or three levels of decision functions. The aid was either unreliable (60% correct) or reliable (80% correct). The dependent measures were accuracy and speed of identifying the most dangerous target and the correct engagement strategy. Their results revealed that reliable automation improved performance compared with manual performance. If the 80% reliable automation aid failed, the participant's performance was worse with decision support than with information-acquisition support. Performance was poor when the aid was only 60% reliable, regardless of the type. The authors concluded that should decision-support automation not be highly reliable, designers should provide users with information automation only because it is easier to compensate for loss of that function in comparison with decision functions.

Level of automation also plays a role in performance under conditions of automation degradation or failure. Wickens et al. (2010) conducted a meta-analysis of 17 studies that examined the effect of level of automation on failure detection. They identified a "routine-failure tradeoff." Simply stated, "more automation yields better human-system performance when all is well, but induces increased dependence so that it will produce more problematic performance when things fail" (p. 389). This relationship is illustrated in Figure 8. The tradeoff is acceptable until point A in the figure. As the automation increases beyond Point A, the negative effects of failure performance become significant. The legitimacy of the routine-failure tradeoff has been supported in more recent research (Smith & Jamieson, 2012).

Manzay, Reichenbach and Onnasch (2008, evaluated the effects of degree of automation (analogous to LOA) and automated decision aids on human performance. Specifically, they measured task performance when using the automated aids and return-to manual performance upon failure of the automated aids. They found that task performance was better with a higher degree of automation. However, return-to-manual performance was poorest under the highest degree of automation. This suggests that performance may improve with higher levels of automation, but perhaps at the cost of a decrement to manual recovery of the automation fails.

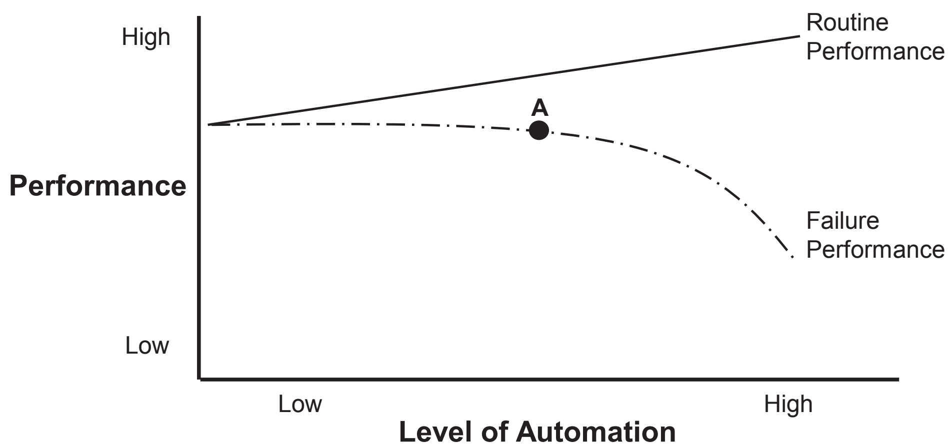


Figure 8. Routine-failure tradeoff (adapted from Wickens et al., 2010).

Wickens suggested that as designers make decision about automating tasks, they should assess this performance tradeoff in the light of automation failure probabilities. That is, there may be a point where the costs of increasing the level of automation are too great if the automation fails. Thus, one approach to manage this tradeoff may be to refrain from higher levels of automation (Kaber & Endsley, 2004). Higher automation reliability exaggerates this effect (probably due to complacency), impairing the operator's ability to detect automation degradation or failures (Dixon & Wickens, 2006; Wickens et al., 2010). Operators are less likely to monitor automation they consider reliable. This situation is considered "overtrust" and is likely to be the root cause for this finding.

Jou et al. (2009) evaluated operator performance and workload on a simulated reactor control task (using two separate tasks: alarm reset and reactor shutdown), with a secondary target detection task with an advanced digital HSI. Operators worked under either low automation or high automation (LOA not specified). Low automation produced significantly longer primary and secondary task performance and significantly higher workload than high automation. Additionally, the authors found performance differences between the types of operational tasks and concluded that the fact the mental workload was significantly higher in the alarm reset task than the reactor shutdown task shows that the task type plays a critical role in mental workload; LOA does not necessarily influence workload, but depends on the task type.

The research on LOA and HAC performance indicates that in most cases, intermediate levels are automation produce the performance. In many of the studies LOA is confounded with cognitive function, and thus it is difficult to make generalizable conclusions regarding the optimal level of automation without considering function. For some cognitive functions, higher levels of automation enhance performance without producing any negative consequences. However for decision making functions, higher levels of automation may enhance performance, but at the cost of reduced ability for the operator to return to manual performance.

Some studies have investigated how adaptive automation affects performance compared to a static LOA. Kaber and Endsley (1997) studied how LOA and adaptive automation affects human performance, namely workload and vigilance demands, as participants were presented multiple tasks and goals to be completed simultaneously in a complex and dynamic simulation and a secondary monitoring task. The results showed that LOAs had an effect on the operator's performance of the primary task, but no effect on the secondary monitoring task. The adaptive automation had more of an effect on the secondary monitoring task and less of an effect on performance of the primary task.

Clamann, Wright, and Kaber (2002) found a relationship between adaptability and cognitive function. The authors studied how adaptive automation affects performance when applied to different stages of human information processing. Their results showed that the participant's performance with adaptive automation was better compared to fully manual operation, but that adaptive automation also was more effective when supporting less cognitively demanding stages of information processing (e.g., information acquisition and action implementation) than more cognitively demanding stages (e.g., information analysis and decision making).

Parasuraman et al. (2009) investigated the effect of an adaptive automation system for a human operator supervising multiple unmanned vehicles. The adaptive automation was adjusted based on the human operator's performance in a change detection task. They found that performance and SA were higher under conditions of adaptive automation than in manual or static automation. These results indicate that operator performance may be a useful method for activating adaptive automation, provided that the measure of performance is validated first.

Though some studies have successfully employed adaptive automation using a specific triggering condition (e.g., Parasuraman et al., 2009), there has not been a great deal of research on the relative merits of the various triggering conditions for adaptive automation (Kaber, 2012; O'Hara et al., 2010), yet their importance to human-automation collaboration is widely acknowledged (Kaber, 2012; Sheridan & Parasuraman, 2005). Since there are relative benefits and disadvantages to each type of techniques mentioned in Section 2.2.7, several authors have recommended the use of hybrid methods to ensure that automation is initiated (or changed) when it should be (Parasuraman et al., 1996; Sheridan & Parasuraman, 2005). Hybrid may lead to a more robust, resilient system that is less subject to potential problems or errors of individual triggers.

That the initiation of automation can be based on machine authority may contradict the often stated principle in the automation literature that human should always be in charge (e.g., Billings, 1991, 1997a). In actual complex systems, this is often not the case and should not be the case (Inagaki & Sheridan, 2012). For example, in the recently published Revision G to MIL-STD-1472 (U.S. Department of Defense, 2012), the following Requirement of automaton is stipulated:

> Requirement 4.12.2, Human involvement. Irrespective of the level of automation, system and task design shall ensure that the human user is in command, involved in ongoing operations, and appropriately informed to maintain awareness of the situation and other status of automated functions.

However, while this is generally true, some system functions are allocated to automation because they cannot be performed by personnel within time requirements (O'Hara, 2012). For example, if a safety function has to be performed within a second or two, the user is not really in command of it. As another example, there may be situations where automation initiates a critical action because users have failed to do so. What does it mean for the user to be in command of these types of automation applications? They will be performed with no input or opportunity for the user to intervene.

When automation can be initiated under either human or machine authority, the situation is called "co-agency" (Inagaki & Sheridan, 2012). When humans are in control, we refer to the situation as "supervisory control" (Sheridan, 2011). Note that this usage of "supervisory control" is different than the LOA of Supervisory Control. Inagaki and Sheridan (2012) explain that the advantage of co-agency is that automation can be authorized to respond if it can detect the operator's failure to respond appropriately.

Another consideration related to adaptive automation is that there is a tradeoff between operator-initiated and automation-initiated triggers (Kaber, 2012). Requiring operators to initiate automation or changes to automation levels increase workload because they have to take action. This may come at a time when they want to initiate automation because their workload is already high. Automation-initiated triggers do not require additional operator workload. However, operators may become disoriented or distracted by the initiation of automation or a change in the degree of automaton. This disorientation can cause transient performance decrements. The cognitive cost of initiating automation cannot outweigh its benefits or operator will not use it (Parasuraman et al., 2009). Thus when operator initiated triggers are used, designer should seek to minimize the workload associated with it. When automation-initiated triggers are used, designers should seek to design etiquette strategies to alert operators to the change in a manner that minimizes distractions and interruptions.

In general, adaptive automation tends to enhance automation compared with static automation; however additional work is needed to define the optimal strategy for adaptive automation for the aSMR context.

Automation design dimensions (such as reliability, LOA and adaptability) clearly have an effect on human performance. However, for any given set of the automation design dimensions, the HSI may have a substantial impact on the HAC performance. Skjerve & Skraaning (2004) evaluated a human-automation interaction display in two experimental studies. The key features of the experimental interference were that it provided the following:

- Representation of the key automatic devices on the overview display
- Verbal feedback associated with the activity of the key automatic devices; dedicated displays for the automatic controllers available on the operators' workstations
- Computer-based logic diagrams available on the operators' workstations.

Both studies found that the quality of the collaboration between the human and automation was greater in the experimental display compared with the conventional display.

2.3.3 Research Related to Effects of Automation on Operator Awareness

Situation assessment is the evaluation of current conditions to determine that they are acceptable or to determine the underlying causes of abnormalities when they occur (e.g., diagnosis). Operators actively try to construct a coherent, logical explanation to account for their observations. This cognitive activity involves two related concepts: the mental model and the situation model. The mental model refers to general knowledge governing the performance of highly experienced individuals, and it consists of the operator's internal representation of the physical and functional characteristics of the plant and its operation. The mental model is built up through formal education, training, and operational experience. Operators develop and update a mental representation of the situation based on the factors known, or hypothesized, to be affecting the plant's state at a given point in time. This situation assessment process produces a mental representation that is referred to as a situation model, the person's understanding of the specific current situation; alternately this may be described as "situation awareness," the understanding that personnel have of the plant's current situation. SA also is used more generally to refer to an operator's awareness and understanding of what is going on (Endsley, 1995).

To develop SA, operators use their general knowledge and understanding about the plant and how it operates to interpret the information they observe and understand its implications. Operators constantly update their SA as they receive new information. The HSI provides alarms and displays that are used to obtain information in support of situation assessment. The HSI may provide additional support to SA in the form of operator support systems. Limitations in knowledge or in current information may result in incomplete or inaccurate SA.

Situation assessment and accurate SA is critical to taking proper human action. This is noted in an IAEA report (1988) with respect to events involving incorrect human actions: "Frequently such events have occurred when plant personnel did not recognize the safety significance of their actions, when they violated procedures, when they were unaware of conditions of the plant, were misled by incomplete data or incorrect mindset, or did not fully understand the plant in their charge" (p. 19). If operators have an accurate situation model, but mistakenly take a wrong action, they have a good chance of detecting it when the plant does not respond as expected. However, when an operator has a poor situation model, they may take many "wrong" actions because, while the actions are wrong for the plant state, they are correct for their current understanding of it.

SA is a term that generally refers to an operator's awareness of what is going on (Endsley, 1995). As a research construct, SA has received extensive attention over the last 20 years. 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, 2004; Bedny, Karwowski, & Jeng, 2004), but the model that has received the most empirical investigation and support by far is Endsley's (1995) model, which was developed through work in the aviation industry and has been applied in numerous additional industries, including, but not limited to, air traffic control, military command and control, power plant operations, National Aeronautics and Space Administration missions, rail system operations, equipment maintenance, and medicine. Endsley's SA model allows operator SA to be measured and quantified, which makes it a supremely useful model for empirical research.

Endsley's SA model (shown in Figure 9) 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 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.

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).

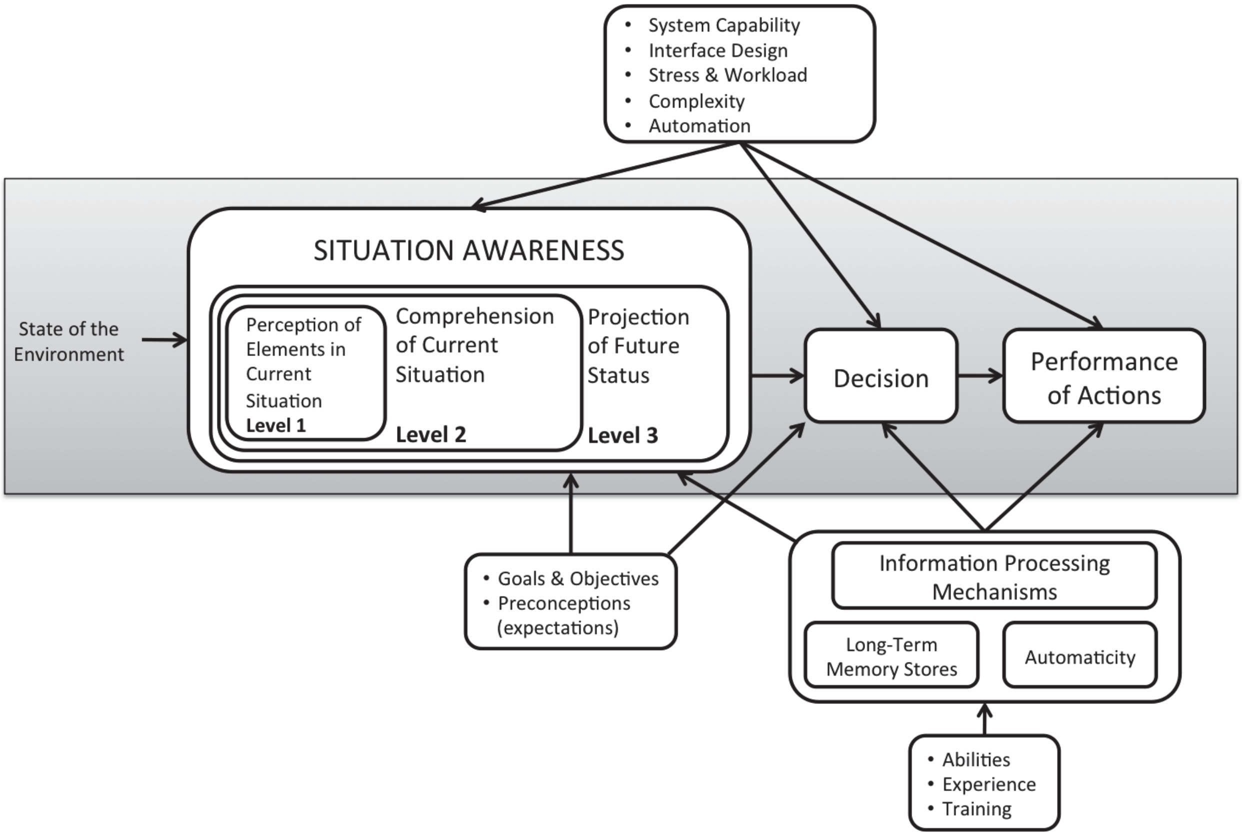


Figure 9. Endsley's model of situational awareness.

When working with a system that is highly automated, the operator's role changes. Instead of interacting frequently with the system and process, the operator must engage in extensive amounts of monitoring of the automated process instead and remain vigilant to identify any system changes or indications of system upset. However, research has shown that vigilance, which is an operator's ability to maintain focused attention, awareness, and alertness over a prolonged amount time, is actually a very difficult task (Warm et al., 2008). Performance on a vigilance task declines over time, sometimes as early as 5 minutes into the vigilance task. Vigilance tasks tend to tax limited cognitive resources and are associated with high levels of perceived workload and stress (Warm et al., 2008).

Therefore, forcing the operator into the role of monitoring has significant consequences for the operator's ability to maintain awareness of the state of the plant and automation.

Furthermore, Parasuraman & Manzey (2010) have shown that under conditions of high task load, operators preferentially allocate their attention to manual tasks at the expense of monitoring the automated system. This leads to reduced awareness of what the system is doing and increases complacency. Moray & Inagaki (2000) showed that in highly reliable automated systems, it actually is a sensible and suitable strategy for operators to not maintain constant SA, meaning that the operator's behavior is well calibrated to the reliability of the system.

One of the most significant and persistent 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; Lee, 2006; 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.

Endsley (1996) asserts that the out-of-the-loop phenomenon occurs through the following three primary mechanisms:

- Vigilance failures and complacency related to monitoring, including either over reliance on or distrust of automation
- Assumption of a passive monitoring role rather than an active controlling role, which makes the operator less actively engaged in the decision-making process
- Changes in feedback provided by the system to the operator, such as providing information in a different form or manner than what the operator is used to, changing the salience of critical information, providing too much information overall, information of lower quality than what the operator needs, or providing integrated displays without the underlying data on which the display is based.

Additionally, Endsley notes that one of the major obstacles to successful automation is the difficulty that operators have in understanding automated systems, which are typically more complex than manual systems; this makes achieving an accurate mental model of the system very difficult and often is due to poorly designed interfaces or inadequate training.

It has been suggested that automation can improve operator SA by reducing operator workload. Studies have shown a detriment to SA when workload is either too high or too low. As such, operator workload seems to be a major consideration for designing automation. However, some studies have shown that automation does not always decrease operator workload; in fact, it can increase it, depending on how it is designed and the cognitive functions automation is responsible for (Lin et al., 2010a). Furthermore, automation may reduce operator workload at the expense of SA (van de Merwe et al., 2012, Miller & Parasuraman, 2007).

Despite all of this, well-designed automation can provide benefits to SA by reducing visual clutter and providing integrated displays and, for the most part, automation has worked well and accompanied a reduction in many types of human errors. 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).

As discussed in Section 2.2.1 and 2.2.2, the impact of automation on operator SA and workload depends on LOA and the cognitive function (i.e., monitoring, planning or generating options or strategies, making decisions, and implementing actions) automation is providing. Increasing amounts of automation produces different effects, depending on which cognitive function is automated and in what manner. A number of studies have been conducted to investigate this relationship between level and cognitive function.

Lin et al. (2009) conducted a simulated reactor shutdown with a secondary detection task at two different levels of automation (Endsley and Kaber's LOA2 and LOA9; see Table 2). They found that LOA9 produced lower operator workload than LOA2. While they did not find any significant differences in operator SA, depending on LOA, some operators in the high automation mode reported feeling out-of–the-loop. The authors concluded that existing automation design still is not sufficient to eliminate out-of-the-loop problems. Subsequent research by the same group of researchers did find an effect of LOA on both workload and SA (Lin et al., 2010a). They compared workload and SA for four levels of automation (Endsley and Kaber's LOA2, LOA5, LOA6, and LOA9; see Table 2). LOA6 (an intermediate LOA) produced the best SA and lowest workload of all four LOAs. LOA2 produced the highest workload, and

LOA9 produced the out-of-the-loop performance problem. Participants in this experiment preferred lower LOAs for generating options and decision making and higher LOAs for monitoring and implementing (Lin et al., 2010a).

Lin et al. (2010b) expanded on this research in a survey of participants familiar with their control room simulator task. The authors used Rasmussen's (1983) skill-rule-knowledge taxonomy to identify the types of human errors that can occur in various levels of automation. Depending on LOA (Endsley & Kaber, 1999; see Table 2) and the cognitive function provided by automation, different types of human errors are more likely. Participants reported that skill-based slips and lapses are the most likely human errors in LOA2. However, in LOA9, participants indicated that the most likely human errors are knowledge-based mistakes. Participants also reported feeling out of the loop in LOA9 because so many tasks are automated and the operators complete so few tasks. As a result of being out-of-the-loop, knowledge-based mistakes may occur when dealing with unfamiliar alarms or unexpected conditions (Lin et al., 2010b).

van de Merwe et al. (2012) conducted research on an air traffic control task with an automated decision support tool called SARA that provides controllers a speed and route combination for inbound flights. They did not use a particular LOA taxonomy; what van de Merwe et al. varied was the type of information that the decision support tool provided (nothing beyond a standard air traffic control system, delta time (change in estimated arrival), speed advisory only, and speed and route advisories). As a decision support tool, SARA corresponds to Endsley & Kaber's (1999) LOA5; this study found that differences within a single LOA between types of information presented may have an impact on operator performance. Specifically, the researchers found that speed and route information produced the best overall performance, negligible impacts on workload, and a decrease in operator SA due to the out-of-the-loop phenomenon. Two studies were performed. SARA did assist controllers in delivering traffic more accurately, but interestingly produced lower SA. The participants mentioned that with SARA they felt less engaged in the traffic situation, or in other words, they felt out-of-the-loop. The authors conclude that particular care needs to be taken when designing automated decision support tools to avoid out-of-the-loop phenomenon and support operator SA.

In an in-depth study of LOA effects on performance, Kaber & Endsley (2004) investigated whether interspersing manual with automatic control (at five different LOAs: LOA3, LOA4, LOA6, LOA9, and LOA10; see Table 2) of a target-processing task similar to an air traffic control task (with a secondary gauge monitoring task) would impact operator performance, SA, and workload. They found that LOA3 and LOA9 produced the worst operator SA across conditions. LOA3 had better target-processing performance than LOA4 and LOA6 and, when combined with longer times on automation, LOA3 produced better target processing performance than LOA6. LOA3 also yielded the greatest number of target collisions, indicating operators were out of the loop in terms of avoiding target conflicts. LOA4 was the worst in terms of target processing performance across all conditions. When interspersed with periods of manual control, LOA10 was associated with improved SA, but when LOA10 was used without periods of manual control, operator SA was significantly poorer. LOA9 also produced low SA that can be attributed to the out-of-the-loop performance problem. In terms of workload, the cycle time between periods of manual control and automation had the greatest impact: higher cycle times (longer periods on automation) produced the lowest operator workload. LOA3, in particular, showed marked detriments to workload when there were no periods of manual control; in other words, full-time monitoring of the automation involved more workload than interspersing LOA3 with periods of manual control. An important conclusion from this study is that interspersing periods of manual control with fully automatic operation can enable operators to maintain adequate SA. Depending on LOA, different cycle times between periods of manual control and automated control can produce markedly different overall performance, operator workload, and operator SA automation in the primary task; however, LOA was not a driving factor in secondary task performance.

Another factor contributing to the failure to properly monitor automation is that automation often performs tasks independently from plant personnel (O'Hara & Higgins, 2010). Personnel often have other tasks for which they are responsible. While personnel do play a role in monitoring the performance of the automation, that responsibility often becomes compromised in the face of workload pressures. This problem is exacerbated when automation is reliable, and personnel trust and depend on it to function properly (Parasuraman & Riley, 1997). Thus because of workload management strategies, personnel may continue to use automation, even when it does not correctly fulfill its functions.

Across these multiple studies, there appears to be a general consensus that intermediate levels of automation (i.e., Endsley & Kaber's LOA6, Parasuraman et al.'s LOA5, and O'Hara's automation by consent) are best to ensure proper operator SA. There also is agreement across many studies that high levels of automation are not recommended due to the out-of-the-loop performance problem that high levels of automation continue to produce. This presents a particular problem for aSMR designs in that aSMR plants will be highly automated as a requirement for financial viability. This fact highlights an area in which additional research is needed: given high levels of automation, how can system designers ensure adequate operator SA (or design in sufficient time and means for operators to recover SA)?

The HSI have can have an important impact on operator SA in the same manner that it can affect objective performance. Dehais, Causse, and Tremblay (2011) studied how "cognitive countermeasures" presented through the automation HSI can help mitigate cognitive errors (such as attentional tunneling [Wickens & Alexander, 2009]), the operator can commit when automation behaves in an expected 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.

Jou et al. (2011) applied content category analysis and performance evaluation matrix methods to explore the potential operator errors that can be caused by advanced digital HSI in light water reactor NPP control rooms. They identified that multiple accidents, pressure level, number of available operators, and other environmental factors are key issues that impact the likelihood of operator errors. They recommend providing guidance on prioritizing tasks in multiple accident scenarios, increasing staffing to reduce the individual operator workload, and taking care with implementing automation to ensure operator workload is not increased.

HSI is critical in cases of automation degradation or failure. HSI is key to whether operators detect the automation degradation or failure. Operator SA can minimize the "routine-failure tradeoff" discussed in Section 2.3.2.2 above; 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).

As an example of this issue, O'Hara, Gunther, and Martinez-Guridi (2010) analyzed the possible effect of failures of an NPP's digital feedwater system on HSIs and operator performance. A previous study determined the risk significance of digital I&C failures on this system (Chu et al., 2008). They developed a detailed failure modes and effects analysis (FMEA) for the system. Using that analysis, they extended the effects to how they would impact the HSIs used by operators to monitor and control the system and how those impacts might affect operator performance.

Seventeen of the degraded conditions are latent failures because they do not cause loss of automatic control of the system, but lower its functionality to some extent. If other degraded conditions occur and/or the operators make a mistake(s) after a latent failure, the outcome can range from negligible to severe. In eight out of these seventeen degraded conditions, the HSI gives *no* indication that the degraded condition exists.

In fourteen of the degraded conditions, one or more of the HSIs give some indication that a failure occurred. Sometimes, the HSI only informs the operators that there was a failure, but does not specify the condition. Operators generally would need technical support from maintenance personnel to troubleshoot the specific cause of the failure. One interesting case, the failure, the system provides an analog input signal for the steam generator (SG) level to a valve controller. The information is displayed to the operators, but the controller does not use it for any calculations or decisions. Accordingly, this failure mode does not directly affect the system's operation. However, the displayed SG level will be (incorrectly) low, and may mislead operators to take erroneous actions to increase the SG level, e.g., increasing the flow of feedwater to the SG. This can lead to a high SG level, and should the high-level set point be reached, the reactor will be tripped. This example illustrates how a system failure can be designed to lead operators to conclude a control failure has occurred when it has not.

Five of the degraded conditions cause a loss of automatic control of a valve that requires operators to take manual control of the system. The failure to do so may result in a reactor trip due to an incorrect SG level. In these five cases, the operators have available information about the degraded condition, but it is not alarmed. Hence, some time may elapse before they become aware that a degraded condition exists, potentially allowing the problem to worsen. A reactor trip is a transient that challenges the operators, and potentially, the safety systems. Should some components or trains be unavailable at the time of the trip, the transient may evolve into a serious safety challenge, e.g., the accident at Three Mile Island Unit 2 in1979 started with a reactor trip with a loss of feedwater.

Thus the analysis of selected degradation and failure failure modes in a digital feedwater system revealed the following:

- Operators can be mislead about the plant's state when automation uses different information than is made available to the operators, and, while responding appropriately to the situation, may appear to be malfunctioning to operators in view of their information and understanding of the situation. Further, operators may take inappropriate actions based in the erroneous information.
- Important degradation of the digital system may not be alarmed nor communicated to operators in a timely way. This can cause a delayed response, and possibly none at all.
- Degraded conditions may not affect the system's functionality and may not be communicated to the operators. This might create latent failures and subsequently more serious events should there be new failures or certain changes in conditions.
- Loss of automatic control places demands on operators, and can lead to significant transients, such as a reactor trip.

The authors identified strategies that might be adopted to minimize the potential impact of degraded automation in this system:

- Displays are needed to support operator awareness of degraded components within complex systems, such as the digital feedwater control system (8 of 17 degraded conditions in it are not communicated to the control room).
- The HSIs should indicate to the operators the same information as is used by the automaton, otherwise their SA may be compromised

- Five of the degraded digital I&C conditions cause the loss of automatic control and the need for manual action. Therefore, an alarm should alert the operator of the automatic-manual status of the system.

The authors further suggested that potential I&C failures should be analyzed to determine their effects on HSIs and the operator performance that may be needed in response. Extending the designer's failure modes and effects to include how failure modes are processed through the HSI might identify potential impacts on human performance that could be addressed in system design. Using such a method, the authors uncovered weaknesses in the HSIs and offered opportunities for improvements.

Consistent with this finding, Parasuraman and Riley (1997) noted that there is evidence to indicate that automation failures were better detected when the behavior of automation can be easily determined in the HSIs, especially those that minimize attentional demands (such as integrated displays and emergent features).

2.3.4 Research Related to Use of Automation

When HAC is designed such that use of an automated aid is discretionary rather than mandatory (i.e., the operator has the choice to use the automation or not), there a number of factors that influence whether a human operator will choose to use the automated aid.

One of the most important factors that influences whether an operator chooses to use automation is the perceived reliability of the automation. Related to perceived reliability is the concept of trust in automation, which has been a concept that has been researched extensively in the human factors literature. For a review of research that has looked more directly at trust, see (Beck, Dzindolet, & Pierce, 2007; Cook, Lacson, & Manes, 2012; Endsley & Strauch, 1997; Ezer, Fisk, & Rogers, 2005; Ghazizadeh, Lee, & Ng Boyle, 2012; Itoh, 2012; Jian, Bisantz, & Drury, 2000; Madhaven, Wiegmann, & Lacson, 2006; Meyer, Feinshreiber, & Parmet, 2003; O'Hara, Brown, Higgins, & Stubler, 1994; Riley, 1996; Sarter, Mumaw, & Wickens, 2007; Spain, Bustamante, & Bliss, 2008; and Strauch, 1997). However, the research team does not view trust as its own independent factor in the sense that it is a very similar construct to perceived reliability, which itself is an intermediating variable between reliability and use of automation.

Ross et al. (2008) examined the relationship between automation reliability, operator trust, reliance (use of automation in place of manual performance), and performance. Student participants performed a simulated unmanned ground vehicle task to identify the locations of terrorists, civilians, and improvised explosive devices. They were told that using the aid was optional, but were not told about its reliability (the levels of which ranged from 75 to 99% calculated as a function of misses and false alarms). Some participants performed the task with no aid. Even though the participants did not know the system's reliability, the results showed that participants' reliance on automation was a function of its reliability. As reliability increased, perceived trust was greater and task performance better.

In a study by Rice et al. (2008), student participants performed a simulated security screening task with the support of a decision aid for detecting weapons that was 95%, 80%, or 65% reliable. The aid never missed a target, but could produce false alarms. The participants made target-detection decisions in both time-pressured and non-time-pressured scenarios. The former were intended to increase reliance on automation. Overall detection performance was best when the reliability was 95% and worst when it was 65%. A higher LOA dependence (agreement with the recommendations of the decision aid) was registered in the time-pressure scenarios despite the occurrence of false alarms. The authors concluded that using time-pressure scenarios encouraged automation use (i.e. operators will use less reliable automation if their workload is sufficiently high).

Riley (1996) demonstrated that perceived risk has a significant influence on operator use of automation. Risk in this context refers to the consequences of failing to achieve the mission or accomplish the task. Riley points out that this is a significant difference between studies conducted in laboratories/simulators when compared with real-world systems. In laboratory/simulators studies, there are no real consequences for failing, while in real-world systems; the real consequences of failure have a different impact on an operator's use of automation.

Carsten et al. (in press) found that participants in a driving study had different levels of trust and acceptance of automaton depending on the tasks to which it was used, even when the reliability was the same. In this study, driver response was not the same for automation supporting longitudinal and lateral driving tasks under identical reliability conditions. This finding emphasizes the need to understand general research findings within the context of each specific application.

Parasuraman and Riley (1997) offered a number of recommendations for how to encourage proper use of automation and discourage improper use. They suggest proper use can be facilitated by making sure the automation's processes and mode are transparent to the operator and to make it easy for the operator to be able to engage and disengage automation as the operator sees fit. To avoid misuse, they recommend making sure the operator's workload is not too high, that operators are given ample training on how to use automation so they have confidence in their ability to use it or to perform the task manually, to use automation cues as heuristics for the operator's decision making, and to have automation provide clear feedback on its processes, mode, and intentions. To avoid disuse, they recommend that designers remain vigilant about aspects of automation that lead the operator to mistrust it. These aspects are generally related to when automation fails to perform as the operator expects and include specific issues such as automation having a high false alarm rate and low reliability.

In general, operators use automation more if they perceive it to be reliable. However, highly reliable automation may contribute to issues of complacency and inappropriate use. Though different researchers refer to many different types of inappropriate use (automation bias, complacency, misuse, disuse overreliance, etc.) these different constructs can be distilled into two categories of inappropriate use: first operators can fail to use functioning automation (thus wasting their own resources on doing a task manually) or they can use malfunctioning automation (thus failing in the task execution).

The cognitive function that is executed by automation, along with LOA for that function, may affect whether an operator will use automation. Bekier, Molesworth & Williamson (2012) conducted a survey of air traffic controllers that inquired about their hypothetical acceptance of automated air traffic management tools. The survey revealed that air traffic controllers were willing to accept automated aids as long as the ultimate decision making was still delegated to the human operator. However, if decision making was delegated to the automated tools, then air traffic controllers reported that they would likely reject automation.

Sheridan and Parasuraman (2005) identified "perceived understandability" of automation as a factor impacting the operator's perceived reliability of automation and its acceptance and use. Understandability refers to the ability of the operator to "form a mental model and predict future system behavior." This is analogous to the automation design dimension of "process."

One of the important factors that contributes to successful HAC is whether the HSI informs the operators about what the automation is doing, what its reliability is, and so forth.

The content and format of the HSI design impacts operator trust. (Lee & See, 2004; Parasuraman & Riley, 1997). 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.

Operator trust in automation can increase if they are provided with information about the processes automation uses to accomplish its functions such as control algorithms and decision logic. For example, Oduor and Wiebe (2008) gave information about the way a decision aid arrives at its decisions and recommendations to students performing a simulated task involving adjustment of a city's resources and tax rates. Information about its algorithms was given as a graphic display, a textual display, or, at times, none was given. Also, the reliability of the decision aid was varied from low to high. A between-subject design was used. The perceived reliability of the decision aid was greatest for textual information, followed by graphic information, and the lowest in the no algorithm condition. The results for measures of "understandability" followed a similar pattern. Overall, perceived reliability was low in the low-reliability condition when compared with the high-reliability condition. The authors concluded that presenting an automated aid algorithm supported the appropriate calibration of trust and a better understanding of automation.

Wang, Jamieson, and Hollands (2008b) studied students undertaking a target identification task; targets identified as hostile were to be shot and friendly targets were not shot. This task was performed in one of three conditions: no aid, 67% reliable aid, or 80% reliable aid. The aid classified targets as friends or unknown. Its classification of friendly targets was always correct, but it was fallible when classifying targets as unknown. Reliability was quantified as the percent of hostile targets classified as unknown. One group received information about the aid's reliability (the informed group), while the other did not (uninformed group). The informed group was given specific information as to the aid's reliability in its classification of targets as unknown, while the uninformed group only knew the reliability was not 100%. Measures of performance were obtained (false alarms – shooting targets that are friendly; misses – failing to shoot hostiles), trust, reliance, and belief in the aid's reliability (for the uninformed group). It seemed that informing participants of the aid's reliability helped in establishing proper reliance.

McGuirl and Sarter (2006) obtained similar results in examining the effect of providing confidence information about an aid's task performance on the operator's use of information. Instructor pilots undertook simulated flights during which icing was encountered. The decision aid supported the detection and management of the icing conditions. One group of pilots received only overall confidence information (system is 70% accurate), while a second group saw a confidence trend display of continuously updated confidence information over time. The latter experienced fewer icing-induced stalls and were more likely to modify their approach to the icing conditions when it was not effective. The authors concluded that providing more precise information about the decision aid's confidence improved the pilots trust and, consequently, their use of the automated system. The authors were concerned that the continuously updated confidence information they received might constitute an information overload. After examining its impact on other flying tasks, they uncovered no negative effects.

The way the information is presented is important; for example, Lacson, Wiegmann, and Madhavan (2005) demonstrated the effect of the presentation mode of reliability information to operators on their automation-utilization strategies. Students performed a signal detection task with the support of one decision aid. Three approaches were used to communicate the aid's reliability: positive framing (the aid is 80% reliable), negative framing (the aid is 20% inaccurate), and neutral (the aid is 80% reliable and 20% inaccurate). Performance was better in the group receiving neutral information than in the two other groups, suggesting that more complete information may be the best approach for improving the use of automation utilization.

Another important aspect of the way information is presented is the automation's perceived etiquette. The design of automation's etiquette has been found to impact the operator's trust in automaton (Atkinson et al., 2012). Automation that is intrusive is trusted less than automation that interacts with the operators in a more "civilized" manner.

2.4 Conclusions

In order to fully comprehend all facets of HAC, it is imperative that the research team identifies and analyzes key contributing factors, including LOA, reliability, and out-of-the-loop. Carrying out this analysis, along with a corresponding review of the extensive collection of human factors' literature geared toward automation, provided critical information for better understanding of the current HAC state-of-practice and the means to construct the initial framework.

Reliability of automation is an important feature of HAC performance due to its direct effect on system performance and its mediating effect on human performance in HAC. Generally, as reliability increases, an operator's trust and use increase, resulting in increased system performance. Unfortunately, higher reliability also tends to decrease the operator's monitoring of automation, leading to inappropriate use when automation fails (this is referred to as complacency, misuse, and automation bias in the literature).

Level and cognitive function also are important features of HAC that affect human performance. For many cognitive functions, high levels of automation produce better system performance than lower levels. However, for some cognitive functions, such as decision making, higher levels of automation produce poorer performance. Additionally, higher levels of automation for any function tend to reduce operator monitoring and contribute to out-of-the loop issues, which reduce an operator's ability to regain manual control after an automation failure.

Many studies indicate that adaptive automation may be a promising way to increase the overall level of automation while mitigating some of the negative consequences of high LOAs. However, it is unclear what the best strategies for initiating adaptive automation are, and further research is needed to establish them.

In addition to the automation design dimensions, there are several contextual factors that interact with the design dimensions to affect HAC performance. One of those factors is task load. As task load increases, many of the effects of automation design dimensions increase. For example, operators are less likely to monitor highly reliable automation under conditions of high task load.

Similarly, the design of the HSI can interact with the automation design dimension to influence the success of HAC. Poorly designed HSIs are implicated in many automation failures. However, in many cases, the HSI can provide a means to mitigate some of the negative effects of using higher reliability and higher levels of automation on human performance. For example, if HSI provides a simple means to monitor automation, the operator is more likely to do so, thus eliminating monitoring problems like complacency and automation bias. Additionally, providing information about automation (such as reliability, the process, and current mode) through the user interface can improve HAC and reduce the feeling of the operator being out-of-the-loop.

While human factors and psychological research has gone a long way in identifying the factors that influence HAC, there are clear gaps in the current state of knowledge for addressing the needs of aSMRs. First, the majority of the human factors' literature (with a few exceptions) defines performance problems associated with certain HAC configurations, but the literature does not necessarily illuminate the circumstances that lead to successful HAC.

Second, taken together, findings from the existing literature would indicate that using intermediate levels of automation for most functions is ideal for keeping operators in the loop. Additionally, many studies indicate that intermediate LOAs are also ideal for performance. However, as stated in Section 1, aSMRs are likely to employ much higher levels of automation to meet the need of reducing O&M costs to a per kilowatt cost that is comparable to the existing fleet of reactors. Therefore, extensive research needs to be conducted to investigate how to enable higher levels of automation, while still keeping the operator actively engaged in operation of the plant.

Finally, a potentially important limitation in the existing research is the fact that none of it was conducted in the aSMR context and very little was conducted in the nuclear domain. Most of the studies were from a military or aviation context. Some of the conclusions presented in the literature simply are not applicable in the nuclear context. For example, the recommendation to use lower reliability automation to enhance human performance is inappropriate given the potential consequences of low reliability in the nuclear domain.

3. MODEL OF HUMAN-AUTOMATION COLLABORATION

As stated previously, the overall objective of this project is to research the issues that contribute to the optimization of HAC in order to maximize the productive and safe operations of aSMRs. To do this, a model of HAC for aSMRs was developed. This preliminary HAC model serves as the basis for future empirical research that will inform further refinement of the HAC model and the development of engineering procedures and guidance for use by designers in developing HACs. These products of the HAC model (i.e., planned empirical studies, engineering procedures, and guidance) are discussed in more detail in Section 4. The HAC model defines (1) the important design dimensions of automation that impact automation's use by personnel and integrated human-automation performance, and (2) what aspects of human cognition, behavior, and performance mediate automation's use by personnel (i.e., the model identifies how human cognition and behavior interact with the design dimensions of automation to affect overall human-system performance). In doing this, the model identifies what aspects of the human-automation interaction are important for the designer to consider in developing automation for SMR systems. This model is based on existing research and operational experience from aviation, military, and existing NPPs. The importance of a HAC model has been expressed by a number of researchers, including Parasaruman and Wickens (2008), Bruni et al. (2007), Kaber et al. (2009), Sanchez (2009), and Linegang et al. (2006).

3.1 Description of the Human-Automation Collaboration Model for Advanced Small Modular Reactors

The nature of HAC is dependent on many factors related to the characteristics of HAC, and a number of contextual factors that mediate the interaction. All of these factors interact to produce either a successful HAC outcome or one of a number of different kinds of unsuccessful HAC outcomes (i.e., HAC failure modes). Furthermore, those HAC performance outcomes may or may not have direct effects on overall system performance.

The research team has developed a model of HAC that shows the relationship between these factors and outcomes (see Figure 10). This model describes the factors that affect HAC; the potential outcomes of HAC, which depend on the ways in which the factors interact; and the ultimate consequences on system performance. In experimental psychological terminology, the characteristics of HAC [dark blue box] are independent variables, and the contextual factors [light blue box] are mediator variables. The research team hypothesizes that the ways these variables interact when they are applied to and used in the aSMR design will affect a number of dependent variables [big blue diamonds], including the operator's SA, use of automation, workload, skills, and abilities (e.g., to take manual action). Furthermore, this model hypothesizes that, depending on the state of the operator's SA, use of automation, workload, skills, and abilities, the resulting HAC and system performance (which are also dependent variables) will either be successful (i.e., functional or satisfactory) [green boxes] or unsuccessful [black box], leading to a loss of money and/or failure of the mission goal. The model also specifies that there are general HAC failure modes [red boxes] that will describe the specific human errors that the operators are expected to commit when their SA, use of automation, workload, skills, and abilities are adversely affected by the independent and mediating variables. Finally, the model represents the defense in depth and resilience of the engineered safety systems present in all aSMR designs as balance of system variables [small blue box], which is included to convey the idea that even in the event of a human error and dysfunctional HAC, the overall system is designed to operate safely.

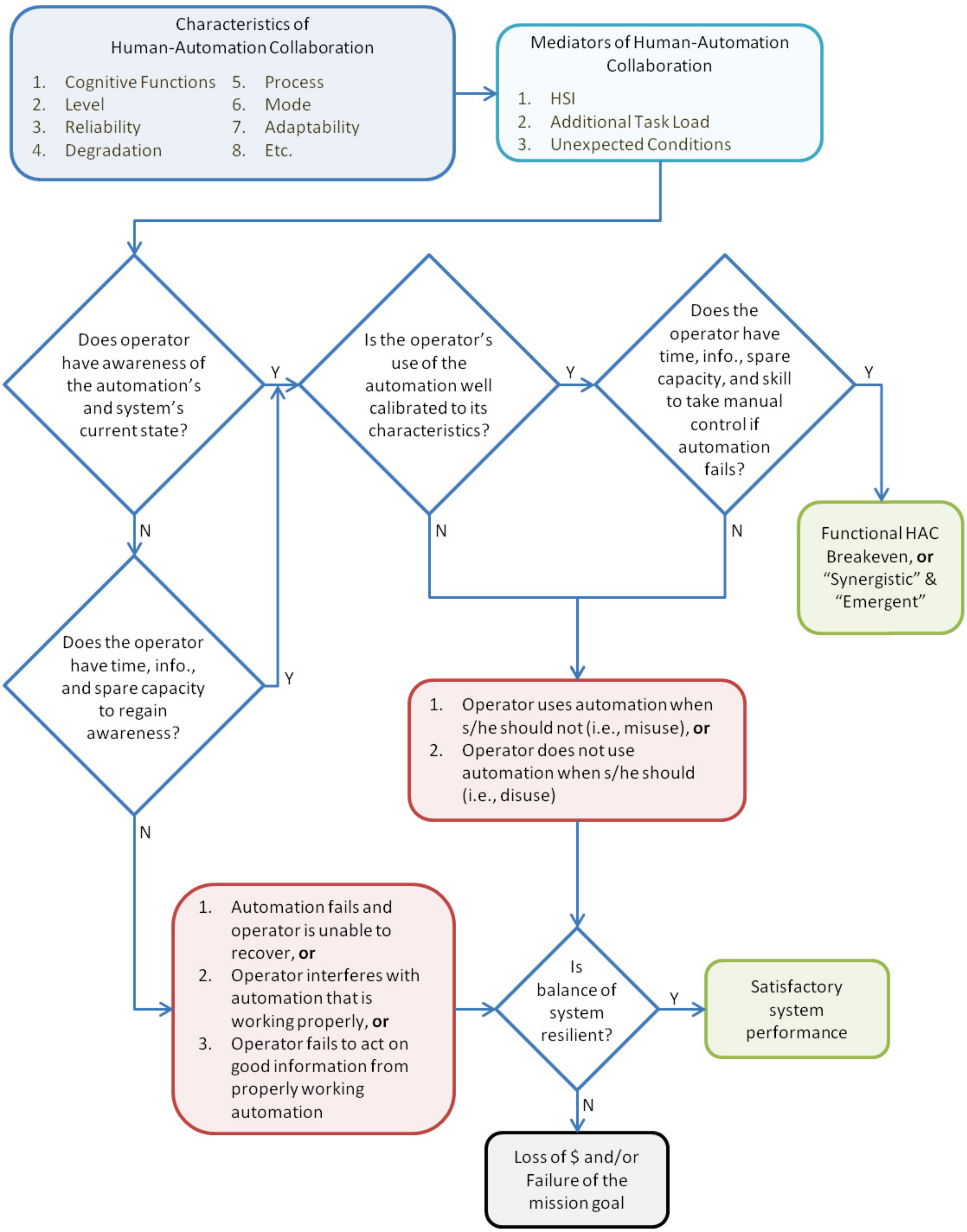


Figure 10. High-level model of human-automation collaboration for advanced small modular reactor designs.

Note that the model in Figure 10 is a preliminary model of HAC for aSMRs. In particular, the research team recognizes that this is a high level model and that lower levels of detail that identify more specific relationships between factors and outcomes is needed. As more knowledge is gained as a result of empirical studies conducted, information acquired on aSMR designs and other surrogate facilities (e.g., planned concepts of operations), and new research findings in the literature, the model will be updated (i.e., refined and more detail added), and documented in future reports and research papers.

3.2 Characteristics of Human-Automation Collaboration

There are several characteristics of the automation and how the HAC is designed. These factors need be represented in the model and need to be considered by aSMR designers in order to gain a complete understanding of the important variables that contribute to successful HAC.

3.2.1 Cognitive Functions

Cognitive functions typically are described in terms of human information processing analogies. This HAC model uses the breakdown described in O'Hara and Higgins (2010), which is monitoring and detection, situation assessment, response planning, and response implementation. As described in Section 2 showed, there is relationship between these stages of information processing that are automated and human performance. Thus, in order to provide guidance on HAC design, this characteristic is included in the HAC model in order to understand how it interacts with other factors and how it affects system performance.

3.2.2 Level

Multiple taxonomies describe levels of automation (for a review, see Section 2.2.2; or Sheridan, 1992). O'Hara and Higgins (2010) proposed a taxonomy that reflected NPP operations rather than aviation (Billings, 1997a). All taxonomies generally vary from fully manual (i.e., the human does everything) to fully automatic (i.e., the automatic system does everything) to intermediate levels, typically including some collaboration between automation and the human. LOA is included in this model because as numerous studies described in Section 2 show (e.g., Prevot et al., 2012; Lin et al., 2010b; Bekier et al., 2012), it has an impact on human performance and ultimately system performance.

3.2.3 Reliability

The current working definition of reliability used in the HAC research effort is the same as what was stated in Section 2.2.3. Reliability is the recognition that automatic systems might fail entirely or in part, compromising their ability to achieve their intended function. Overall reliability may be expressed in terms of probability that the system will correctly perform its function. The effect of automation's reliability on the human operator, HAC, and system performance is well documented in the literature review provided in Section 2 (e.g., Wicken & Dixson, 2007; Ross et al., 2008); therefore, it is clearly an important causal factor to include in this HAC model for aSMRs.

3.2.4 Degradation

Automation depends on the other instrumentation and control subsystems (i.e., sensors, monitoring, and communications) to function properly. Malfunctions in these other subsytems (e.g., loss of sensors or significant delays in the transmission of information along the data highways) can degradation or can automation to fail (O'Hara et al., 2010). Degradation and failure can lead to two types of problems for operators (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 such as causing abnormal operating conditions due to erroneous automatic action or providing erroneous information. It is important for aSMR designers to understand the nature of these dependencies within the instrumentation and controls subsystems, because their degrading effect on automation can change how automation behaves and performs, which can subsequently affect HAC and system performance.

3.2.5 Process

Process refers to the way an automatic system uses input from sensor feeds in the plant and from the human operator and assesses it relative to its programmed information processing routine (e.g., control algorithms and decision logic) to initiate its preprogrammed response. When well designed, the automation's preprogrammed responses are appropriate, or correct, for the input it receives and assesses. This is how automation performs correctly and how it is able to accomplish the tasks assigned to it.

However, not all automatic processes are going to be readily comprehensible by humans, and research summarized in Section 2 document the consequences of opaqueness (e.g., Odour & Wiebe, 2008). Sometimes, the process is very complex or occurs too quickly for the operator to detect. Other times, the process would be comprehensible, but it is not communicated to the operator through the HSI. Successful HAC performance frequently depends on the operator knowing what the automation processes are. The operator not knowing what the automation processes are contributes to the operator losing SA. Thus, in order to facilitate HAC performance, the system design should accommodate operator comprehension through any of a number of design modifications such as making the automation process transparent to the operator and simplifying complex processes for human operators.

3.2.6 Mode

Automation can be designed to have different modes and the behavior of and processes automation executes can be substantially different depending on which mode it is in. One example of a relatively simple mode error is when a human is using graphics editing software and the user expects the mouse pointer to behave like it does when controlling the computer's operating system, when, in fact, the mouse pointer has been set to behave like a paintbrush or an eraser. A more nuclear power relevant example of the potential for mode error is the automatic power control system. The automatic power control system can be set in (1) turbine following mode for load following or (2) reactor following mode for full power operations. If an operator takes a local controller out of automatic into manual, she/he will need to know in what mode the automatic power control system is functioning. If this is not communicated to the operator, their manual actions could result in degraded system performance.

Sarter and Wood (1995) provides a review of how the operator losing SA on what mode automation is in can lead to human errors and dimished system performance. HAC performance can depend greatly on whether the operator is aware of which mode the automation is in and how well the operator understands what behaviors and processes automation will exhibit in that mode, which is why it is an important characteristic in this model of HAC for aSMRs.

3.2.7 Adaptability

The level and cognitive functions that are allocated to automation, the human, or both can be done statically or dynamically. The decision criteria for when tasks or functions should be reallocated can be based on factors such as the operator's workload level (i.e., triggering conditions). In addition, aSMR designers need to consider the iniators for changing the allocation of functions. Three examples of initiators include operator initiated, automation initiated, and hybrid (i.e., either the operator or automation can initiate).

However, a number of studies have shown that the dynamic allocation of functions and tasks between automated systems and operators can positively and negatively affect HAC and system performance (e.g, Parasuraman et al., 2009; Clamann et al., 2002). Another example is if the adaptation mechanism fails and functions can no longer be dynamically allocated or the operator is unable to regain control of key safety functions, this will not only adversely affect the operator's trust and future use of automation, but also overall system performance. Given that the goal for this project is to provide aSMR designers the information needed to optimize HAC and that adaptability will be widely used in aSMR designs, it is essential that this model of HAC include this characteristic in the model, because the model is the basis that informs this research and the subsequent guidance that this project produces.

3.3 Mediating Factors

A number of factors may enhance or degrade the effect of the characteristics of automation design and the HAC design characteristics (i.e., factors identified in Section 3.2) on HAC and system performance. Considering only the effects of the characteristics of automation design and HAC design characteristics on HAC and system performance without recognizing important contextual factors does not provide a complete picture of how and under what conditions these factors will affect HAC performance and whether their effect can be amplified or attenuated.

3.3.1 Human System Interface Design

HAC is executed through the HSI or graphical user interface. Therefore, design of the HSI and the effectiveness of the human-system interaction (also called the human-computer interaction) via the HSI will have important effects on the nature of the interaction. An HSI that has been modeled and designed through a well-established, human-system interaction methodology, such as the goals, operators, methods, and selection rules model (Card, Moran, & Newell, 1983), or the user-centered design approach (Norman & Draper, 1986) will make up for deficiencies created in the earlier stages of the HAC design that would have otherwise produced unacceptable HAC performance (e.g., an automated process previously obscured to the operator and thereby creating a human out-of-the-loop issue could be corrected with the proper indications presented in the HSI). Similarly, functions may be allocated to best take advantage of each agent's capabilities, but if the HSI design is poor, it will attenuate the positive effect the sound function allocation decision-making affect has on system performance. These two examples illustrate why HSI is treated as a mediating factor, because the extent to which the characteristics of automation design and HAC design characteristics affect HAC and system performance depends to a large degree on the design of the HSI.

3.3.2 Additional Task Load Unanticipated by Function Allocation

A good function allocation method will allocate functions to agents (i.e., automation, human, or both) in a way that optimizes performance. In many cases, the function allocation also will balance workloads and achieve cost efficiencies. However, unanticipated events may produce additional workload that was not considered in the original function allocation. This additional workload may change the way the human interacts with automation (e.g., it may reduce monitoring performance). This is an example of how context or changing circumstances in the operating environment may enhance or degrade the effects of the factors identified in Sections 3.2.

3.3.3 Unanticipated Operational Conditions

It is impossible to predict all possible operating conditions a system will encounter; therefore, it is safe to assume that, at some point during the life of a plant, it will be operating under conditions that were not considered in the original function allocation and HAC design. These changes may positively or negatively affect the ability of both the human and the automation to perform in the manner expected.

3.4 Human-Automation Collaboration Performance Consequences

Whether interaction of the factors described in Sections 3.2 and 3.3 results in successful or unsuccessful HAC performance depends on a number of key intermediate outcomes, including (1) the operator's awareness of the automation's and plant's current state, (2) whether the operator's use of automation is well calibrated to its capabilities and intended use, (3) the operator's current workload level, and (4) the operator's proficiency in performing tasks manually if the automation fails, or lack thereof (e.g., skill loss).

3.4.1 Functional Human-Automation Collaboration

When the operator is (1) aware of the automation's and plant's current state, (2) using automation appropriately, and (3) can successfully intervene if automation fails (i.e., has sufficient time, a manageable workload level, the necessary information, and necessary skills to take over manual control), then the consequence, or outcome, is called Functional HAC.

Furthermore, when these prerequisites are met, the model specifies that there are at least two different kinds of acceptable or successful HAC. One of the early findings cited in the human factors literature was that there was the expectation that automation was going to improve system performance or at least achieve system performance equivalent to less-automated systems at a reduced cost. However, multiple studies have found that this is not always the case; for example, automation intended to reduce operator workload can actually produce increased workload (Bainbridge, 1983).

The first kind of successful HAC is where automation and operator perform at the level expected. The "expected level" of performance can be defined in different ways. For example, it can be defined as what the aSMR designers would predict given how they addressed the factors described in Sections 3.2 and 3.3 or relative to the optimal performance of a similar system that relies on more manual operator control.

The second kind of successful HAC is where HAC performance exceeds the expected level of performance. This outcome has been seen in some studies on automation (de Visser & Parasuraman, 2011) and, generally speaking, is an extension of the literature on group performance, where multiple studies (for a review see Surowiecki, 2004) have shown how groups of people outperform or produce more than the same number of individuals working independently (i.e., the group's performance is synergistic and the ephemeral properties of "teamwork" emerge or become emergent). However, it is recognized that a number of studies show how groups perform worse than individuals working independently, but the key preconditions to synergistic or emergent team performance are well understood. To the extent that automation and humans operating as a team have the same team dynamics issues and challenges that all human teams have, these preconditions for emergent team performance should apply to when humans and automation collaborate to achieve a common goal.

3.4.2 Three Failure Modes Related to the Operator's Lack of Awareness of the Automation's and Plant's Current State

As stated previously, whether interaction of the factors described in Sections 3.2 and 3.3 results in acceptable or unacceptable HAC performance depends on a number of key intermediate outcomes. When the operator is unable (or unmotivated) to perceive, comprehend, and project what the automation's processes and activities are or will be, it is unlikely the operator will be able to recover when automation fails. It is only under conditions where the operator has the prerequisites of adequate time to recover, the necessary information, and the spare capacity to take on additional workload that this failure mode can be mitigated. If the operator does not have all three of these prerequisites, then one of the following failure modes can occur:

1. Automation fails and the operator is unable to recover.
2. Operator interferes with automation that is working properly, thereby potentially introducing a fault into the system that could propagate into a system failure.
3. Operator fails to act on good information from automation that is working properly.

3.4.3 Two Failure Modes Related to the Miscalibration of the Operator's Use to Automation's Capabilities

The operator's use of automation is a well-researched topic in the human factors literature. One of the most often used lexicons (e.g., use, misuse, disuse, and abuse), developed by Parasuraman and Riley (1997), served as a starting point for identifying the remaining two failure modes. In this model of HAC, the operator's use of the automation is recharacterized in terms of whether it is well calibrated to the automation's capabilities and the intended use of automation. The extent to which the operator's use is not well calibrated will result in one of the following two use-related failure modes:

1. The operator uses automation when she/he should not.
2. The operator does not use or rely on automation when she/he should.

3.4.4 Moderating Effect of the Resilience of the Balance of the System

This model of HAC recognizes that HAC performance outcomes can be and often are different from overall system performance. Sometimes poor HAC performance leads to unacceptable overall system performance; other times poor HAC performance is mitigated by other features of the overall system. This difference is due to the fact that aSMRs are complex systems that are designed with defense-in-depth, which often is achieved through diversification of instrumentation and control system design and engineered passive safety features. Most of the time, the defense-in-depth is well coordinated and robust (i.e., resilient) and at other times uncoordinated and fragile (i.e., brittle).

When the balance of the remaining system is resilient and can mitigate the effects of poor HAC, it is still possible to achieve overall satisfactory system performance.

When the balance of the remaining system is brittle, it is unable to mitigate the effects of poor HAC and under these circumstances will lead to a loss of money and/or failure of the mission goal.

4. PATH FORWARD

This section discusses the research studies that this project is planning to conduct to fill in the knowledge gaps about HAC, and the engineering procedures and guidance that will be produced with respect to helping aSMR designers with their planned use of automation.

4.1 Framing the Advanced Small Modular Reactor Context for Human-Automation Collaboration Research – Assumptions

As stated in the introduction of this report, SMRs and aSMRs are expected to have design features that are vastly different from existing light water reactor NPPs, which means that the way operators and automation interact (i.e., collaborate) will be significantly different than current conduct of operations in U.S. NPPs. While NPPs currently in operation have become increasingly more automated, the fundamental concepts of operations have remained the same. The approach to concepts of operations in existing NPPs is very human labor intensive and requires multiple operators per reactor unit/core. However, SMR (Petrovic et al., 2012) and aSMR vendors (Tsuboi et al., 2012) alike have indicated that in order to be cost competitive[2], aSMRs will need to operate with fewer operators and use more automation under both normal and emergency operating conditions[3]. Yet, questions such as what should be automated, how should the characteristics of automation be combined, what should be manually controlled, what should be jointly controlled by automation and human operators, and why (i.e., what technical basis was used to answer these questions) have not yet been answered. While some research has been conducted in other domains to answer these questions, including unmanned aerial vehicles (NUREG/CR-7126) and air traffic controllers (van de Merwe et al., 2012), considerably less R&D has been performed for nuclear power. Clearly, aSMR designs will benefit from more empirical research on how to maximize the use of automation to achieve cost savings, but at the same time not adversely affect safety or performance.

4.2 Preliminary Research Issues

Given framing of the aSMR context in Section 4.1, the relationship of this research project to other DOE ICHMI research pathway aSMR projects (as described in Section 1.1), and the preliminary model of HAC that has been developed for this research project (see Figure 10), the HAC research issues discussed in the following subsections were identified as the highest priority research needs. The research issues were identified through a selection process where the research team asked the following questions (also mentioned in Section 1.4):

1. Is there a knowledge gap?
2. Can it be addressed with experiment(s) (i.e., can it be demonstrated empirically)?
3. Is it relevant to the aSMR field?
4. Can the result be generalizable to most aSMR designs?
5. Can the results also be the technical basis for development of HAC-related engineering procedures and aSMR guidance?

Three knowledge gaps identified made it through the selection process. These gaps are presented as research topics in Section 4.2.1 to 4.2.3.

It is important to note, however, that these research issues are only a subset of known issues with HAC that have not been studied thoroughly in the research literature. These other knowledge gaps are described in more detail in the Appendix.

---

[2] For a general overview of the economic issues for SMRs and aSMRs, see Boarin et al. (2012).

[3] For a discussion of the regulatory perspective on staffing for SMRs and aSMRs, see Trikouros (2012).

4.2.1 Research Topic 1: Impact of Highly Automated Advanced Small Modular Reactors on Operator Awareness

Based on the fact that aSMR designers have said they plan to use automation extensively in the conduct of operations (Tsuboi et al., 2012) and some, including NuScale (Reyes, 2012) and the Pebble Bed Modular Reactor (Nicholls, 2001) also have proposed multi-unit operation, how can operator and team awareness of the automation's and plant's state (i.e., operating condition) be maintained? This research proposes to investigate whether it is possible to implement high levels of automation in the aSMR design, require operators to monitor multiple independent units simultaneously, and still keep the operator sufficiently in the loop so he/she can perform his/her functions or responsibilities when required.

The goal of this study is to identify, in the context of high levels of automation and multi-unit aSMR operations, whether an aSMR with this concept of operations can still allow an operator to perform the functions or responsibilities allocated to them (e.g., initiate a safety system or correct a malfunctioning safety system that was automatically actuated). Additionally, this study proposes to evaluate the effect of high levels of automation for different stages/modes of plant operation (e.g., start-up, normal operations, off-normal operations, emergency operations, and shutdown operations) on the operator's SA of both the automation's and plant's state.

In the context of the HAC model for aSMRs, this question relates to the three failure modes related to the operator's lack of awareness of the automation's and plant's current state, as described in more detail in Section 3.5.1. This study will elucidate the hypothesized causal relationship between (1) HAC characteristics, including, but not limited to, level, cognitive function, mode, and HSI, and (2) failures of awareness (i.e., automation fails and the operator is unable to recover, operator interferes with automation that is working properly, or operator fails to act on good information from automation that is working properly).

4.2.2 Research Topic 2 – Regaining/Reacquisition of Awareness

This study proposes to identify how quickly operators can regain or reacquire SA after having lost SA and under what circumstances the process of reacquisition of SA can be expedited. This study assumes that it is not reasonable to expect operators to maintain perfect SA of multiple units over the duration of their shift. Given that it is inevitable that operators will have reduced SA at times during their workday, the following research questions need to be investigated:

1. What do HSI and automation need to do to facilitate a successful reacquisition of SA?
2. What is the necessary information that the HSI must provide the operator?
3. Can information that predicts the future state of the plant be generated and presented on the HSI so that operators have some advanced warning?

In the context of the HAC model for aSMR, this question relates to how the operator can recover SA and what preconditions are needed in order for the operator to recover. This study will provide insights into what aSMR designers will have to do in their designs to ensure the operator has sufficient time, the necessary information, and a workload level that is not too high so the operator can take on additional tasks to perform the recovery action.

4.2.3 Research Topic 3 – Effect of Human-Automation Collaboration Characteristics on Operator's Use

Section 2 of this report presented a number of studies that showed how HAC characteristics (e.g., reliability and level) can affect use. However, not all causal relationships are known, particularly when considering the unique aspects of the aSMR context. This study proposes to identify, from the perspective of aSMR design needs, the most critical gaps in what is known about how HAC characteristics affect use and conduct studies to fill those gaps.

For example, implicit in the increased use of automation in aSMRs is the fact that this automation will be based on newer and more complex technology than the automation technology used in existing light water reactor NPPs. As the complexity of automation increases (e.g., automation that can dynamically vary level, mode, and process versus an automated system that has only one level, mode, and process), it can affect operator use in new and more complex ways. Understanding how more sophisticated automation technology affects operator use and, subsequently HAC and system performance, is important for an aSMR designer to know.

Furthermore, as Vicente et al. (1996) and others have documented, existing NPP displays typically use a design philosophy where a one-to-one representation of the parameters on the displays is used, which requires the operator to mentally synthesize the data to give it meaning. Advanced automation displays used in aSMRs will likely adopt an ecological interface approach (Vicente & Rasmussen, 1992) or use information-rich displays that present aggregated or grouped information, reducing the inherent information complexity, and rely on representational aiding (Woods, 1991) principles to effectively map lower-level data to higher-order functional information that gives the lower-level data meaning. How this change in HSI philosophy affects operator use, HAC, and system performance also is important to understand.

4.3 Review of Experimental Approaches and Apparati to Support Human-Automation Collaboration Empirical Studies

As Section 2 of this report has shown, the human factors literature on automation is vast. The sheer amount of research conducted on this topic is a good indicator that HAC is a significant issue. However, it is also clear that the existing literature does not provide all of the guidance that aSMRs designers need to know in order to implement automation in their designs without adversely affecting optimal system performance. Some additional empirical research is needed, not only because there are some gaps in what is known about HAC, but also because the context in which a significant portion of the literature has been conducted (e.g., military and aviation) is markedly different from the aSMR context. Furthermore, a number of reports describing (Tsuboi et al., 2012) or reviewing (NUREG/CR-7126) aSMR designs have stated that there will reduced staffing levels in aSMRs for economic viability reasons, implying the extensive use of automation in the daily operation and control of aSMRs. Extensive use of automation is a significant departure from the conduct of operations at currently operating large light water reactor NPPs in the United States. As such, using existing NPP operating experience also will likely prove to be an insufficient source of information for designers of aSMRs.

Given the anticipated need to conduct cost-effective and well-designed research that is targeted at effectively filling in the knowledge gaps for HAC in an aSMR context, the research team investigated what empirical approaches might be suitable for these research requirements. The research team investigated full-scope simulation, microworlds (a kind of part task simulation), and other approaches that have been developed and used as an experimental platform/medium for other studies that could be adapted for this project. Full-scope simulator studies are not being considered at this time due to their high costs and for the practical reason that it is unlikely that any aSMR vendor yet has had a full-scope simulation of their design programmed.

A microworld is a computer simulation that is used as an experimental tool that can be used to increase the fidelity of controlled laboratory studies. Microworlds have been characterized by Brehmer and Doerner (1993), as being (1) complex in that they require an operator to try to achieve multiple goals simultaneously, which often leads to the operator having to make trade-off decisions; (2) dynamic in that the state of critical variables the operator must monitor and control are changing over time; and (3) opaqueness in that some variables and processes are not directly observable and require the operator to make inferences from other available information. Thus, while not as complex and realistic as a field study, a microworld provides some additional realism to the study and allows the experimenter to still maintain control over key variables that may confound the experimental design.

While numerous studies that have used microworlds can be found in the human factors literature, the software used in these studies is very application or research domain specific. For example, a number of microworlds have been developed for aviation. A few have been developed for simulating wild fires or forest fires, and many generic process control industry microworlds have been developed. In general, microworlds or part-task simulation have the advantages of being targeted and cost effective. Clearly, the relevant parts of the anticipated uses of automation in aSMRs need to be simulated, but the research team will need to decide whether an existing microworld, or one that is developed, is the more ideal experimental apparatus. Besides the publicly available microworlds, there are a variety of part-task simulator development software tools that allow researchers to design software-based simulations for designing, prototyping, and deploying control systems and other embedded monitoring applications. These software development programs allow researchers and engineers to (1) design and then operate a dynamic simulation of various plant processes on a computer, (2) develop the behind-the-boards instrumentation and controls logic to monitor and control the simulated plant process, and (3) design the in-front-of-the-boards graphical user interface that the human operator will use to interact and control the system. In short, they allow the researcher to develop a customized complex simulation that includes the critical human-system interactions that are likely to be present in real aSMRs. Additional features of the simulation, such as HSI (i.e., graphical user interface), complexity of the process, complexity of automation used, and how functions are allocated to either the automation or human operators, also can be manipulated to determine their effects on human and overall system performance.

Computation-task simulation techniques (e.g., discrete event simulation) also are an option. This allows Monte Carlo methods of parametric simulation of the interaction between entities to be used to identify the probabilistic triggering of alternatives. Meta-analyses, engineering analyses, and other subject matter approaches are also possible options. As the research team proceeds with this phase of the project, all experimental approaches and apparti will be evaluated for how well they help facilitate the research project’s overall goals.

4.4 Advanced Small Modular Reactor – Specific Engineering Procedures and Design Guidance for Human-Automation Collaboration

The HAC model developed for this project will be used to develop engineering procedures and design guidance (i.e., to support designers in making decision as to where, how, and when to use automation) and to design the human-computer interaction and the HSI/procedures/training to support that interaction. Examples of these engineering procedures include the following:

- Procedures for identifying the type of HAC for a specific application (e.g., deciding on the levels of automation to be implemented for a given function and whether the automation should be adaptive)
- Procedure for evaluating, verifying, and validating HAC to ensure the plant’s production and safety goals are achieved.

Guidelines provide the principles for implementing a selected type of HAC in the design (e.g., to build the user interfaces with which personnel monitor and interact with automation based on the level of HAC selected).

Thus far, the focus of the research team has been concentrated on developing the technical basis upon which the HAC model can be built.

The developed HAC model characterizes the nature of human and automation interaction and its affect on performance. Characterization is important because it provides a structure for developing and organizing the design procedures and guidance that will be developed later in this project.

4.5 Summary

The needs analysis and development of the initial model of HAC described in this report are important activities toward developing a technical basis. As discussed in Section 4.2, the research team will continue to explore and investigate concerns related to HAC by conducting studies that are designed based on insights gained so far.

Additionally, researchers currently are in contact with aSMR vendors to gain a better understanding of their specific design and HAC needs. However, because most aSMR vendors are in their plant design and licensing processes, it is a challenge to find detailed information regarding their plans or need for HAC. In fact, research efforts (such as the HAC effort) will be great support in moving the aSMR field forward. Vendors currently are focusing on technical concerns (e.g., how to design the core and the appropriate material for a small reactor); therefore, they focus very little on how to actually operate the plant in an optimal and economic manner. By this research addressing this concern in advance, it allows the vendors to focus on designing their plants. When they are ready to consider operation of the plant (i.e. design collaboration between the operator and automation), the vendor can use the research results as a spring board; therefore, reducing resources spent to get up to speed on how to best design the collaboration.

Because of the current state of the aSMR field, the researchers will look to other industries to find operational experience that can be generalized to HAC for aSMRs. For example, researchers aim to further investigate and leverage experience and lessons learned from the coal power industry. Coal power plants are highly automated; therefore, their control rooms share characteristics that are thought to be applicable for aSMRs.

One of the objectives within the framework for HAC is to specify how HAC should be evaluated to ensure integrated human-automation system performance acceptably meets design performance requirements for both production and safety. A methodology for this objective will be formalized in subsequent research. The studies to be conducted to support development of the methodology will provide methods, data, and lessons learned that will help address this objective.

The framework for HAC will be evaluated and verified through its application to an aSMR design or to a mock-up design, which is as close to reality as the research team can predict it to be based on lessons learned throughout the research effort. The framework and methodology will be used to develop engineering procedures and guidance for implementation of the framework.

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.

Atkinson, D., Hancock, P., Hoffman, R., Lee, J., Rovira, E., & Stokes, C., (2012). Trust In Computers And Robots: The Uses And Boundaries Of The Analogy To Interpersonal Trust. *In Proceedings of Human Factors Society and Ergonomics Society, 56th Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society.

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

BEA (2012). Final Report on the accident on 1st June 2009 to the Airbus A330-203 operated by Air France (flight AF 447 Rio de Janeiro – Paris). Paris, France: Bureau d'Enquêtes et d'Analyses.

Beck, H., Dzindolet, M. & Pierce, L. (2007). Automation Usage Decisions: Controlling Intent and Appraisal Errors in a Target Detection Task. *Human Factors,* Vol. 49 (3), 429-437.Bedny, G. Z., & Karwowski, W. (2004). A functional model of the human orienting activity. *Theoretical Issues in Ergonomics Science, 5*(4), 255-274. doi:10.1080/1463922031000086681.

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.

Bekier, M., Molesworth, B. R. C., & Williamson, A. (2012). Tipping point: the narrow path between automation acceptance and rejection in air traffic management. *Safety Science*. *50*, 259-265.

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

Billings, C. (1997a). *Aviation Automation: The Search for a Human-centered Approach*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Billings, C. (1997b). Some Questions about Advanced Automation. In M. Mouloua & J. Koonce (Eds.), *Human-automation Interaction: Research and Practice* (pp. 314-320). Mahwah, NJ: Lawrence Erlbaum Associates.

Boarin, S., Locatelli, G., Mancini, M., & Ricotti, M. E. (2012). Financial case studies on small and medium-size modular reactors. *Nuclear Technology*, *178*(2), 218-232.

Brehmer, B., & Doerner, D. (1993). Experiments with computer-simulated micro-worlds: escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior*, 9(2-3), 171-184.

Bruni, S., Marquez, J., Brzezinski, A., Nehme, C. & Boussemart, Y. (2007) Introducing a Human-Automation Collaboration Taxonomy (HACT) in Command and Control Decision-Support Systems. In*12TH ICCRTS - Adapting C2 to the 21st Century*.

Carsten, O., Lai, F., Barnard, Y., Jamson A., & Merat, N. (in press). Control Task Substitution in Semiautomated Driving: Does It Matter What Aspects Are Automated? Human Factors.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Christoffersen, K., & Woods, D. (2002). How to Make Automated Systems Team Players. *Advances in Human Performance and Cognitive Engineering Research* (Vol. 2, pp. 1-12): Elsevier.

Chu, T., Martinez-Guridi, G., Yue, M., Lehner, J. & Samanta, P. (2008). *Traditional Probabilistic Risk Assessment Methods for Digital Systems* (NUREG/CR-6962). Washington, D.C.: U.S. Nuclear Regulatory Commission.

Clamann, M. P., Wright, M. C., & Kaber, D. B. (2002). Comparison of performance effects of adaptive automation applied to various stages of human-machine system information processing. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 46(3), 342-346.

Cook, M., Lacson, F. & Manes, D., (2012). Guided Attention For Autonomous System Supervision. *In Proceedings of Human Factors Society and Ergonomics Society, 56th Annual Meeting.* Human Factors Society.

Cullen, R., Rogers, W., Fisk, A., & Grado. (2012). The Effects of Automation Reliability and Experience on Attention in a Computer Environment. *In Proceedings of Human Factors Society and Ergonomics Society, 56th Annual Meeting. Santa Monica, CA: Human Factors and Ergonomics Society.*

de Visser, E. & Parasuraman, R. (2007). Effects of imperfect automation and task load on human supervision of multiple uninhibited vehicles. In *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society. (pp. 1081-1085).

de Visser, E. & Parasuraman, R. (2011). Adaptive Aiding of Human-Robot Teaming: Effects of Imperfect Automation on Performance, Trust, and Workload. *Journal of Cognitive Engineering and Decision Making*, 5 (2), 209-231.

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.

Dudenhoeffer, D., Hallbert, B, Miller, D., Quinn, T., Arndt, S., Bond, L., O'Hara, J., Garcia, H., Holcomb, D., Wood, R., & Naser, J. (2007). *Technology Roadmap: Instrumentation, Control, and Human Machine Interface to Support DOE Advanced Nuclear Power Plant Programs* (INL/EXT-06-11862). Washington, DC: Department of Energy.

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.

Endsley, M., & Strauch, B. (1997). Automation and situation awareness: The accident at Cali, Columbia. In R. Jensen & L. Rakovan (Eds.), *Proceedings of the Ninth International Symposium on Aviation Psychology* (pp. 877-881). Columbus, OH: Ohio State University.

Ezer, N., Fisk, A., & Rogers W. (2005). The Effect of Perceived Consequence on Automation Reliance in a Human-automation Collaborative System. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society.

Funk, K., & Lyall, B. (2000). *A Comparative Analysis of Flightdecks with Varying Levels of Automation.* Washington, DC: Federal Aviation Administration.

Ghazizadeh, M., Lee, J. & Ng Boyle, L. (2012). Extending the Technology Acceptance Model to assess automation. *Cognition, Technology & Work*, 14 (1), 39-49.

Goh, J., Weigmann, D., & Madhavan, P. (2005). Effects of Automation Failure in a Luggage Screening Task: A Comparison between Direct and Indirect Cueing. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society.

Hollnagel, E. (1999). From Function Allocation to Function Congruence. In S. Dekker & E. Hollnagel (Eds.). *Coping with Computers in the Cockpit*. Aldershot, UK: Ashgate.

Hollnagel, E., & Woods, D. (2005). *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering*. Boca Raton, FL: Taylor & Francis.

IAEA (1988). Basic Safety Principles for Nuclear Power Plants (Safety Series No. 75-INSAG-3). Vienna, Austria: International Atomic Energy Agency.

IAEA (2005). *Innovative Small and Medium Sized Reactors: Design Features, Safety Approaches and R&D trends* (IAEA-TECDOC-1451). Vienna, Austria: International Atomic Energy Agency.

IAEA (2006). *Status of Innovative Small and Medium Sized Reactor Designs 2005: Reactors with Conventional Refueling Schemes* (IAEA-TECDOC-1485). Vienna, Austria: International Atomic Energy Agency.

IAEA (2009). *Design Features to Achieve Defense in Depth in Small and Medium Sized Reactors* (IAEA Nuclear Energy Series Technical Report No. NP-T-2.2). Vienna, Austria: International Atomic Energy Agency.

Inagaki, T. & Sheridan, T. (2012). Authority and responsibility in human–machine systems: probability theoretic validation of machine-initiated trading of authority. *Cognition, Technology & Work*, 14 (1), 29-37.

Itoh, M. (2012). Toward overtrust-free advanced driver assistance systems. *Cognition, Technology & Work*, 14 (1), 51-60.

Jamieson, G., & Vicente, K. (2005). Designing Effective Human-Automation-Plant Interfaces: A Control-Theoretic Perspective. *Human Factors,* 47 (1), 12-34.

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

Jou, Y. T., Yenn, T. C., Lin, C. J., Tsai, W. S., & Hsieh, T. L. (2011). The research on extracting the information of human errors in the main control room of nuclear power plants by using Performance Evaluation Matrix. *Safety Science, 49,* 236-242.

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. (2012). *Human Factors/Ergonomics Research on Adaptive Automaton and Automaton Modeling Approaches*. HFES Webinar.

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., Stoll, N., Thurow, K., Green, R. Kim, S., & Mosaly, P. (2009). Human–Automation Interaction Strategies and Models. *Human Factors and Ergonomics in Manufacturing*, *19* (6), 601–621.

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.

Lacson, F., Weigmann, D., & Madhaven, P. (2005). Effects of Attribute and Goal Framing on Automation Reliance and Compliance. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society.

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.

Linegang, M., Stoner, H., Patterson, M., Seppelt, B., Hoffman, J., & Crittendon, Z. (2006). Human-automation Collaboration in Dynamic Mission Planning: A Challenge Requiring an Ecological Approach. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society.

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

Lyall, B., & Funk, K. (1998). Flight Deck Automation Issues. In M. Scerbo & M. Mouloua (Eds.) *Proceedings of the Third Conference on Automation Technology and Human Performance*. (pp. 288-292). Mahwah, NJ: Lawrence Erlbaum Associates.

Madhaven, P., Wiegmann, D., & Lacson, F. (2006). Automation Failures on Tasks Easily Performed by Operators Undermine Trust in Automated Aids. *Human Factors, 48* (2), 241-256.

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.

McGuirl J., & Sarter, N. (2006). Supporting Trust Calibration and the Effective Use of Decision Aids by Presenting Dynamic System Confidence Information. Human Factors, 48 (4), 656–665.

Metzger, U., & Parasuraman, R. (2005). Automation in Future Air Traffic Management: Effects of Decision Aid Reliability on Controller Performance and Mental Workload. *Human Factors, 47* (1), 35-49.

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.

Moray, N., & Inagaki, T. (2000). Attention and complacency, *Theoretical Issues in Ergonomic Science*, 1(4), 354-365.

Nicholls, D. R. (2001). The Pebble Bed Modular Reactor. *Transactions of the Royal Society of South Africa*, 56(2), 125-130.

Norman, D. A., & Draper, S. W. (1986). *User centered system design: new perspectives on human-computer interaction.* Lawrence Erlbaum Associates, Inc

NRC (2012). *Report to Congress: Advanced Reactor Licensing*. Washington, D.C.: U. S. Nuclear Regulatory Commission.

Oduor, K., & Wiebe E. (2008). The Effects of Automated Decision Algorithm Modality and Transparency on Reported Trust and Task Performance. In *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

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.

O'Hara, J, & Higgins, J. (2010). *Human-System Interfaces to Automatic Systems: Review Guidance and Technical Basis* (BNL Technical Report 91017-2010). Upton, NY: Brookhaven National Laboratory.

O'Hara, J. (2012). *Review of the Guidance for Automation Contained in the Department of Defense Design Criteria Standard: Human Engineering* (MIL-STD-1472G). Washington, D.C: NASA Engineering and Safety Center.

O'Hara, J., Higgins, J. & Pena, M. (2012). *Human-Performance Issues Related to the Design and Operation of Small Modular Reactors* (NUREG/CR-7126). Washington, D.C.: U. S. Nuclear Regulatory Commission.

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., & Manzey, D. H. (2010). Complacency and bias in human use of automation: an attentional integration. *Human Factors, 52*(3), 381-410.

Parasuraman R., Mouloua M., & Molloy R. (1996). Effects of adaptive task allocation on monitoring of automated systems. *Human Factors, 38*(4), 665-679.

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).

Parasuraman, R. & Wickens, C. D. (2008). Humans: Still vital after all these years of automation. *Human Factors, 50*(3), 511-520. doi:10.1518/001872008X312198.

Petrovic, B., Ricotti, M., Monti, S., Cavlina, N., & Ninokata, H. (2012). Pioneering Role of IRIS in the Resurgence of Small Modular Reactors. *Nuclear technology*, *178*(2), 126-152.

Prevot, T., Homola, J. R., Martin, L. H., Mercer, J. S., & Cabrall, C. D. (2012). Toward Automated Air Traffic Control—Investigating a Fundamental Paradigm Shift in Human/Systems Interaction, I*nternational Journal of Human-Computer Interaction, 28*(2), 77-98.

Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics, smc-13*(3), 257-266.

Reyes, J. N. (2012). NuScale Plant Safety in Response to Extreme Events. *Nuclear Technology*, *178*(2), 153-163.

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.

Riley, V. (1996). Operator reliance on automation: Theory and data. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and Performance*. NJ: Lawrence Erlbaum Associates.

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.

Rovira, E., McGarry, K., & Parasuraman, R. (2007). Effect of Imperfect Automation on Decision Making in a Simulated Command and Control Task. *Human Factors,* 49 (1), 77-87.

Ruff, H., Calhoun, G., Draper, M., Fontejon, J., and Guilfoos, B. (2004). Exploring automation issues in supervisory control of multiple UAVs. *Proceedings of the Human Performance, Situation Awareness and Automation Conference*, March 22-25, Daytona Beach, FL, 408-412.

Salmon, P., Stanton, N., Walker, G. (2008). Representing situation awareness in collaborative systems. Ergonomics, *51*, 367-384.

Sanchez, J. (2009). The conceptual model of human-automation interaction. Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting, San Antonio, TX.

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 5-19.

Schneiderman, B. (1998). *Designing the User Interface: Strategies for Effective Human Computer Interaction*, 3rd Ed. Reading, MA: Addison Wesley.

Sheridan, T. (1992). *Telerobotics, Automation, and Human Supervisory Control*. Cambridge, Massachusetts: The MIT Press.

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

Sheridan, T. (2011). Adaptive Automation, Level of Automation, Allocation Authority, Supervisory Control, and Adaptive Control: Distinctions and Modes of Adaptation. *IEEE Transactions of Systems, Man and Cybernetics, Part A: Systems and Humans,* 41, 662-667.

Sheridan, T. & Parasuraman, R. (2005). Human-Automation Interaction. In R. Nickerson, (Ed.) *Reviews of Human Factors and Ergonomics* (Volume 1). Santa Monica, CA: Human Factors and Ergonomics Society.

Skitka, L. Mosier, K., & Burdick, M. (1999). Does automation bias decision-making? *Int. J. Human-Computer Studies, 51*, 991-1006.

Skjerve, A. B. M., & Skraaning, G. (2004). The quality of human-automation cooperation in human-system interface for nuclear power plants. *International Journal of Human-Computer Interaction Studies, 64,* 694-697.

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.

Smith, J. & Moore, R. (2009). Small Modular Reactor Issue Identification and Ranking Program Control Room Staffing - Final Report. In *Proceedings from the Eighth American Nuclear Society Meeting on Nuclear Plant Instrumentation, Controls and Human-Machine Interface Technologies.* LaGrange Park, Illinois: American Nuclear Society.

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

Spain, R., Bustamante, E., & Bliss, J. (2008). Toward an Empirically Developed Scale for System Trust: Take Two. In *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Strauch, B. (1997). Automation and Decision-Making – Lessons From the Cali Accident. In *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting*, pp 195-199.

Surowiecki, J. (2004). *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations,* New York, NY: Doubleday

Thomas, K. (2012). Business Drivers for Nuclear Plant Operations and Maintenance Automation. In *Proceedings from the Eighth American Nuclear Society Meeting on Nuclear Plant Instrumentation, Controls and Human-Machine Interface Technologies.* LaGrange Park, Illinois: American Nuclear Society.

Thurman, D., Brann, D., & Mitchell, C. (1977). An Architecture to Support Incremental Automation of Complex Systems. In *Proceedings of the 1997 IEEE International Conference on Systems, Man, and Cybernetics*, Orlando, FL.

Trikouros, N. G. (2012). A Perspective on Small Reactor Licensing and Implementation. *Nuclear Technology*, *178*(2), 233-239.

Tsuboi, Y., Arie, K., Ueda, N., Grenci, T., & Yacout, A. M. (2012). Design of the 4S Reactor. *Nuclear Technology*, *178*(2), 201-217.

U.S. Department of Commerce. (2011). *The Commercial Outlook for U.S. Small Modular Reactors*. Manufacturing and Services Competitiveness Report, February 2011. Washington, DC: U.S. Department of Commerce, International Trade Administration.

U.S. Department of Defense (2012). *Defense Acquisition Guidebook*. Washington, D.C: U.S. Department of Defense.

U.S. Department of Energy. (2010). *Nuclear Energy Research and Development Roadmap: Report to Congress.* Washington, D.C.: U.S. Department of Energy.

Uhrig, R., Gao, R., & Tsoukalas, L. (2004). Use of Multi-agents to Improve Efficiency and Safety of Nuclear Power Plants. In *Proceedings from the Fourth American Nuclear Society Meeting on Nuclear Plant Instrumentation, Controls and Human-Machine Interface Technologies*. LaGrange Park, Illinois: American Nuclear 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

Vedantam, S. (2009). Metrorail crash may exemplify automation paradox. *The Washington Post*, June 29, 2009, Section A, page A.9

Vicente, K. J., & Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. *IEEE Transactions on Systems, Man and Cybernetics, SMC-22, 589-606.*

Vicente, K. J., Moray, N., Lee, J. D., Hurecon, J. R., Jones, B. G., Brock, R., & Djemil, T. (1996). Evaluation of a Rankine cycle display for nuclear power plant monitoring and diagnosis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *38*(3), 506-521.

Wang, L., Jamieson, G., & Hollands, J. (2008b). Improving Reliability Awareness to Support Appropriate Trust and Reliance on Individual Combat Identification Systems. In *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Warm, J.S. Parasuraman, R. Matthews, G. (2008). Vigilance Requires Hard Mental Work and is Stressful. *Human Factors, 50* (3), 433-441.

Wickens, C., & Alexander, C. D. (2009). Attentional tunneling and task management in synthetic vision displays. *The International Journal of Aviation Psychology, 19*(2), 182-199.

Wickens, C. & Dixon, S. (2007). The benefits of imperfect diagnostic automation: A synthesis of the literature. *Theoretical Issues in Ergonomics Science*, 8(3), 201-212.

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

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.

Wiener, E., & Curry, R. (1980). Flight-deck Automation: Promises and Problems. *Ergonomics*, *23*(10), 995-1011.

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.

Woods, D. D. (1991). The cognitive engineering of problem representations. In J. Alty & G. Weir (Eds.), *Humancomputer interaction in complex systems* (pp. 169-188). London: Academic.

Woods, D., & Hollnagel, E. (2006). *Joint Cognitive Systems*. Boca Raton, FL: CRC Press, Taylor & Francis.

Wood, R. T. (2012). U.S. Department of Energy instrumentation and controls technology research for advanced small modular reactors. *Nuclear Safety and Simulation, 3*(4), 292-299.

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

Yoo, H., (2012). Framework for designing Adaptive Automation. *In Proceedings of Human Factors Society and Ergonomics Society, 56th Annual Meeting 56, 2133-2136.*. Santa Monica, CA: Human Factors and Ergonomics Society.

APPENDIX

In the process of reviewing and analyzing the HAC literature, the research team identified a number of research issues. Based on the expected aSMR operating context, and the relationships among the three DOE ICHMI aSMR research projects described in Section 1.1, the research team made a preliminary determination that some HAC research issues are of higher priority than others. That is not to say, however, that the other research issues identified are not important. This appendix documents these other important research issues here, rather than as part of Section 4.2, so that the content of that section is focused on the research issues that are of higher priority.

Additional Human-Automation Research Issues

The following subsections summarize additional HAC research issues. Note that within the constraints of available resources and scope authorization, the research team will prioritize and pursue these research topics using best available means, including but not limited to microworlds; full scope simulators available at the Idaho National Laboratory, Halden, and elsewhere; meta-analyses; engineering analyses; and subject matter expert approaches.

Generalization of the Findings from Academic Research to Complex Systems

Many of the studies examining human-automation interaction have limitations for generalizing the findings to the target operational context this research is interested in: commercial nuclear power plants, highly trained professional operators, and complex HSIs. The findings of many of the studies reviewed for this project were based on students performing fairly simple tasks, using simple desk-top HSIs. Research results are generalized most easily when the operational context is the same. Thus, research is needed to assess the extent to which generalization between these contexts is supported. This would involve replication of key (or selected) findings for aSMR operations.

NPP Operating Experience

Commercial NPPs have begun see an extension of automation to a wider range of plant operations and to decision support systems, as well as the use more interactive forms of automation. Little operating experience relative to NPP operations is readily available in the literature or in industry human performance databases. Additional research is can address this need using proactive information solicitation methods, such has been done in the industry for other I&C issues and in other industries such as aviation.

Detecting and Managing Automation's Degraded Condition and Failure

Even though automation systems typically are highly reliable, the potential for their degradation or failure can significantly jeopardize plant performance and safety. Research is needed addressing the operator's ability to detect and manage degraded automation conditions.

Models of Teamwork

For multi-agent systems, designing automation to be a good "team player" has typically modeled human-automation teams on human-human teams. This made sense since it was the lack of typical human teamwork characteristics that led to poorly designed automation. However, it may be that different models of teamwork are needed for multi-agent teams. It is beneficial to investigate alternative models of teamwork for them.

Performance Measures

A more complete set of performance measures that focuses on the relationship between humans and automation, such as trust and neglect, and those needed to depict multi-agent teamwork, are needed. The concepts/constructs that are important to the relationship need to be defined along with approach to operationalize those concepts for use in:

- Research on the effects of automation design on performance, such as in simulator studies
- Engineering procedures for determining the types of automaton that are appropriate to a given design
- Design evaluations and validations of implementations of design-specific HAC

Isolating the Effects of Confounded Dimensions

Functions and levels and of automation often are confounded in the literature, as are modes and levels. Studies more specifically isolating the effects of each are required, so researchers can better understand the independent effects on the operator's performance of these two independent dimensions.

Calibrating Trust - Personnel trust in automation is based on their perception of its reliability, which may or may not be consistent with reality. This relationship is called "trust calibration." When the operator's perceptions accurately match the automation's reliability and capabilities, trust is "well-calibrated" and operators use it appropriately. Miscalibrated trust leads to either an overreliance on automation (misuse) or its underutilization (disuse). Research is needed on how the trust can be calibrated so that operators properly use automation when they should and override it when they should. There are training and HSI design aspects to this issue.

Communication of Automation's Reliability to Plant Personnel

Reliability affects the operators' trust in automation and their decision to use it. Further, providing information about reliability in the HSI supports these decisions. But how reliability should be quantified and represented are not easy questions to answer, especially for automation supporting situation assessment and response planning.

Processes Used by Automation

Automation's process can range from simple to complex. Research is needed to develop a better understanding of the relationship between process complexity, operator trust, and automation usage. For example, some studies have indicated that if operators do not understand the process used by automation to arrive at a result, the result is more likely to be discounted. A better understanding of this relationship between processes and behavior is needed.

Triggering Mechanisms for Adaptive Automation

Additional research is required to identify the appropriate triggering mechanisms for automation changes, and how they should be implemented to minimize any disruptions to the operator's performance when the change occurs.