SPACE FLIGHT TASK CONTEXTS FOR LONG DISTANCE AND DURATION EXPLORATION MISSIONS: APPLICATION TO MEASUREMENT OF HUMAN AUTOMATION INTERACTION

Chelsea Iwig[1,2], James Oglesby[1,3], Misa Shimono[1,3], Kimberly Stowers[1,3], Kevin Leyva[1,2], Eduardo Salas[1,3]

Institute for Simulation and Training[1], Department of Modeling and Simulation[2], Department of Psychology[3], University of Central Florida

An effort is currently underway to determine methods for measuring safety and performance of human-automation systems to improve their functioning for future long duration space flights. However, an important step in system evaluation is understanding the contexts in which they operate. The identification of contexts will help in targeting what variables may be related to the overall system's effectiveness. A review of NASA documents and literature has resulted in the identification of four categories of task contexts that are believed to be important for future Long Distance and Duration Exploration Missions (LDDEM) to Mars and beyond. These four categories include (1) spacecraft navigation, (2) robotic/habitat operations, (3) systems monitoring, and (4) mission planning and scheduling. Within each of these four task categories there exist varying task demands and environmental conditions that impact the user's interaction with the automation and, subsequently, the types of measurement that are appropriate for analyzing performance and safety within the human-automation system.

INTRODUCTION

Human-Automation System (HAS) integration is an important topic in research and practice due to the common application of HAS in various work environments. Automation is beneficial to completing tasks as it provides an opportunity for increased productivity, efficiency, and safety (Sheridan & Parasuraman, 2005). Nevertheless, automation does not always lead to better performance. For example, in some instances when automation has not been effectively implemented, operator workload or error rates may actually increase due to inadequate human-system integration (Mosier, Skitka, Dunbar, & McDonnell, 2001). Thus, the success of human-automation systems is contingent upon the ability of the human operator and the automation agent to collaborate together on required tasks.

One domain that utilizes automation to a large degree is space flight, and automation is expected to play an even bigger role in the future. In the coming years, manned space flight will be venturing in completely new directions, with mission characteristics that are unlike any other manned space flight mission to date. These future missions will epitomize the term "exploration" beyond that seen in previous missions in that crewmembers will travel further into space than ever before, including to near earth objects (NEOs) such as asteroids, as well as on interplanetary expeditions to Mars. These missions are expected to involve longer durations than current missions to the International Space Station (ISS), where a Mars expedition would last for multiple years (Manzey, 2004). These exploration missions will entail communication delays that will interfere with the ability of Mission Control to provide the same level of oversight and communication with the crew that it currently affords the ISS crews. Long Distance and Duration Exploration Missions (LDDEM) crews will thus be far more autonomous than ISS crews, and crewmembers will need to place a heavier reliance on automation to maintain performance and habitability during future exploration (Sandal, Leon, & Palinkas, 2006).

Due to the importance of HAS for future space flight missions, it is necessary to ensure that implementation of these systems is safe and that they promote maximized performance. To accomplish this, assessment of HAS prior to implementation must occur, and must involve measurement of the safety and performance of the system using reliable and valid measurement methods. Applicable metrics are required to determine if developed systems meet the established criteria and standards associated with task performance and maintaining a high level of safety in the work setting. Such measurements can help inform the system designer about whether there is need for redesign or other countermeasures (e.g., supplemental training guides, rules for system operation) to optimize the effectiveness of the overall human-automation system. Despite the importance of measurement for assessing human-automation system safety and performance, there are a large degree of measures available that vary in effectiveness based on multiple conditions, including the user, the automation design, and the task environment. Specifically, the context of the HAS integration determines the criteria for assessing system effectiveness.

Currently, research is being conducted to determine ideal metrics and approaches to measurement for evaluating HAS safety and performance. Previous research from various HAS contexts has provided an extensive list of metrics for system inputs, processes, and outcomes. However, when attempting to deduce which measures are best for evaluating each of these constructs, it is necessary to recognize that measures validated in one context are not necessarily going to be valid in others. For instance, a certain measure's sensitivity to critical factors may change based on various conditions unique to the specific task context, such as skills required by the operator to interact with the automation, and environmental factors that can impact the functioning of an automation agent or how the entire system operates. Typically, these elements should be considered in the design and implementation process of the system's life cycle.

Therefore, the next step for advising the measurement of performance and safety within HAS for space flight is to understand the contexts of the types of systems that will be designed for future LDDEM. The identification of these contexts will allow for the tailoring of measurement methods to the unique conditions characteristic of exploration missions such as those to Mars or NEOs. The purpose of this paper is thus twofold: (1) to discuss the categorical contexts that exist within space flight and (2) to examine the application of metrics to tasks within these contexts.

SPACE FLIGHT TASK CONTEXTS

To understand the types of systems utilized in the space flight domain, we conducted a literature search to identify various efforts in designing, developing, and validating automation systems to complete different mission-related tasks. Based on this review of work associated with the National Aeronautics and Space Administration (NASA), we have conceptualized four contexts to help frame HAS safety and performance measurements for aiding system designers in developing effective automation for application to future space flight missions (Table 1). These four contexts have been categorized as: (1) Navigation of the Spacecraft, (2) Robotic/Habitat Operations, (3) System Monitoring, and (4) Mission Planning and Scheduling. The following will provide an overview of these four types of contexts, how automation can be implemented in these contexts, and a discussion of considerations when assessing systems within these types of tasks. We have attempted to categorize these contexts to minimize overlap between them, but also provide an inclusive overall view of the types of systems that would be important in current and future space flight missions.

Spacecraft Navigation

As with any space flight mission, the degree of control that flight crewmembers, Mission Control, and automated agents have in maneuvering space vessels will be an important consideration. According to the Jet Propulsion Laboratory Basics of Flight instructional guide (Doody & Stephan, 2013), spacecraft navigation comprises (1) orbital determination tasks, which include knowledge and prediction of spacecraft position and velocity, and (2) flight path control tasks, which include actually firing the spacecraft engines. Orbital determination tasks can be differentiated from route planning tasks, which are categorized in the Mission Planning and Scheduling context, in that orbital determination tasks involve the specific calculations necessary for predicting spacecraft position and velocity during the overall navigation process. These tasks are done after route planning tasks, which are more high level and focus on the path that should be followed rather than the more detailed calculations involved in orbital determination, such as calculations for position and velocity. Subtasks involved in these navigation procedures include docking with cargo vessels, habitat modules, and/or other vehicles; spacecraft launch and reentry; trajectory correction maneuvers; and orbit trim maneuvers (Doody & Stephan, 2013). Additionally, spacecraft navigation includes operation of pressurized ground vehicles during surface exploration

Table 1. Brief descriptions of tasks related to each context.

Context	Tasks Involved
Spacecraft Navigation	Launch and reentry procedures, orbital determination and maintenance burns, module/vessel docking and release, and pressurized rover operation (e.g. Crew Mobility Vehicle). Involves engaging in actual navigation procedures and orbital determination calculations.
Robotic/Habitat Operations	Teleoperated robots, remote rover operation, and habitat systems operation and maintenance tasks. Involves actual operation of habitat systems and robots or unmanned rovers.
System Monitoring	Monitoring/tracking information about automated systems and detecting and addressing errors or critical issues that may arise. Observation of system state information and reaction to changes in the information.
Mission Planning and Scheduling	Route planning (high-level, distinct from orbital determination and execution tasks), scheduling of activities and tasks, planning mission objectives, and rescheduling in response to mission change.

tasks. Specifically, NASA has designed a pressurized ground vehicle for such surface explorations which can support a crew of three for long range surface exploration missions (Drake & Watts, 2014). Therefore, with this information in mind, considerations about effective system design will involve the display of information to the operators, the controls utilized for the task, and the degree/level of automation utilized for maneuvering spacecraft.

When designing systems involving human-automation interaction within spacecraft navigation, it is important to consider the unique characteristics of future LDDEM to Mars or NEOs. The spacecraft navigation portion of these missions will be significantly longer than any previous manned space flight mission. During the time it takes the spacecraft to travel from Earth to the mission destination and then back again, the crew will experience variable levels of workload. During periods of low workload, or during an extended planetside stay without any navigation tasks, it is possible that skill decay will result (Garland, Endsley, Ellison, & Caldwell, 1999), which can be detrimental during the onset of high taskload after long periods of low workload (e.g., surface landing). Therefore, it is important to consider how to evaluate HAS to mitigate this potential for skill decay over time.

Performance assessment of navigation systems may include measuring aspects of the navigation tasks such as time to complete tasks, collision avoidance, error avoidance, reaction time to alarms, number of course corrections, and fuel consumption. Measurements of interface design should involve assessing mental workload and situation awareness during completion of the navigation task. Specifically, interface design measurement considerations within this context should address whether the alarms and displays are salient enough to alert operators of mode changes or occurrences of critical events during completion of the task.

Robotic/Habitat Operations

Crewmembers will interact with various system subcomponents during missions, such as manual manipulation of equipment, surface exploration, and habitat maintenance. En route or during the return trip, these tasks may include repair tasks inside or, in some contingency cases, outside of the spacecraft. Additionally, a science goal that NASA has detailed in the Human Exploration of Mars Design Reference Architecture (Drake & Watts, 2014) is telerobotic exploration of Mars' moons Phobos and Deimos and of the Martian surface. Therefore, it is important to know how to develop effective robotic systems for these contexts where operators may have certain constraints while performing the task, such as limited view of the robot, teleoperation, and incongruent spatial perspective of the task work.

For habitat missions, (e.g., Moon and Mars expeditions), crewmembers may be conducting scientific tasks (data collection, exploration) that require the use of automation systems. Specific surface exploration tasks involving robotic operation may involve crews remotely operating small robotic rovers and surface drills during explorations on the extraterrestrial surface (Drake & Watts, 2014). Other subtasks associated with the vehicles and the habitat operations include maintenance tasks and other basic system operation tasks.

Methods for evaluating the level of safety and efficiency in these types of robotic and habitat systems can be critical, especially when considering that these systems will be deployed in extreme environments with unique environmental conditions. Performance and efficiency of the system could be assessed by measuring time to complete tasks, amount of tasks completed, robot damage avoidance, and the number of inputs and retractions, which could indicate inefficiencies in operations. Additionally, there is the possibility of limited information being available to the operator regarding the robotic system in cases such as tele-operations tasks. In these cases, it is important to consider the unique characteristics of the task that should be measured to ensure safety and performance of the human-automation system.

Assessment of behavioral characteristics of the operator are also important within this context. Specifically, it is important to measure details such as how much time the operator interacts with the robot or if the operator overrides automated processing of the robot as this may indicate poor efficiency within the system resulting from disagreement between the user and the robot. Other aspects of HAS within this context that may be important to measure are operator cognitive processes such as workload and situation awareness. These factors are important for assessing user experience and ability to perform optimally when operating the robotic system.

System Monitoring

The long distance that space flight crewmembers must travel for future Mars missions will result in communication delays ranging from zero in low Earth orbit to 6-40 minutes in transit each way and upon arrival at Mars (Drake & Watts, 2014; Manzey, 2004). Thus, as the crew travels further from Earth, there will be less reliance on Mission Control to monitor and solve problems, along with an increased need for crew autonomy. This means that the flight crew will be conducting much more of the problem solving and decision making, and they must be able to quickly and effectively address potential problems that may arise without relying on the delayed turnaround of communication with the ground crew back on Earth. Automation systems that can help in detecting and diagnosing faults, as well as in addressing them, can be very beneficial for the crew's health and mission operation. However, developed systems need to ensure that the operators are accurately informed about the systems' operations, and can quickly respond to errors and issues as they develop. Design considerations can help with these types of tasks, and applied metrics that can be used during testing phases can help in guiding refinement in the system design and criteria for optimum performance.

Since fault detection is an important consideration for systems within this context, it is important to assess the reliability of the automation in these types of tasks by measuring the proportion of hits, misses, and false alarms produced by the system. Additionally, operator performance should be assessed through measuring the time it takes the operator to react to an alarm and address the problem. Also, it is important for the operator to be able to quickly diagnose problems when they occur; therefore, feedback from the system should be clear and unambiguous so as to allow the operator to have a clear picture of the problem. Since system monitoring requires vigilance, it is important to measure indications of whether the operator is attending to or ignoring the monitoring system. This may include measuring trust in the automation, situation awareness, fatigue, and response to rare alarm instances over long periods of time.

Mission Planning and Scheduling

For future LDDEM, one potential advancement in mission operations is to have automation agents involved in the planning and scheduling of tasks and objectives. Current efforts are being utilized to develop algorithms and functions for determining optimum timelines for space flight missions, as well as reacting to changes in the mission with the rescheduling of critical tasks. Specifically, this effort is currently focused on taking standalone tools that have been implemented in other operational settings, such as the Advanced Caution and Warning System (ACAWS), Scheduling and Planning Interface for Exploration (SPIFE), Extendable Uniform Remote Operations Planning Architecture

(EUROPA), and the execution system called Plan Execution Interchange Language (PLEXIL), and combining these tools into one integrated tool for providing automated planning, scheduling, and critical system feedback information to the crew (Morris, Schwabacher, Dalal, & Fry, 2013).

Implementation of such planning and scheduling systems can be optimized by taking into consideration the design principles that make information, organization, productivity, and decision aids most usable and effective. For example, steps should be taken to ensure that the user is informed about what events are happening when, as well as keeping track of changes made by the system, or those changes that should be made by the operator (Miller & Parasuraman, 2007). Metrics determining how well-perceived these systems are for field implementation, and the user's experience of interacting with mission planning and scheduling systems, can help in optimizing the design for these systems.

Future missions to Mars and/or NEOs will be much longer in duration; therefore, scheduling systems, like the ones mentioned above, will become even more important to mission success due to the complex operations that will be required to take place over the duration of the mission. These systems will be important for ensuring the timely completion of mission goals and the ability of the crew to work autonomously apart from Mission Control during periods of significant communication delays. Therefore, measures of interaction with such systems are important for optimizing performance of the HAS. Such measures may include assessing whether the user actually uses the system or if the user overrides automated scheduling changes. This is important for assessing user agreement with the system and the competency of the automation. Additionally, it is important to measure situation awareness of the users by assessing whether they notice automation-initiated scheduling changes.

DISCUSSION

To aid in developing effective automation systems for their intended purpose, system designers must be able to accurately assess system performance and safety throughout the system life cycle to optimize the development of such systems to meet mission requirements. Therefore, providing established, valid, and reliable measurements and techniques can help in efficiently developing systems, increasing safety, and optimizing the cost-effectiveness of system development and use.

The effort outlined in this paper provides a conceptualization of how space flight systems encompass a variety of purposes and contexts. As such, different types of automation systems place different demands and criteria for human-automation performance on how they interact and operate. Given these different demands, the assessment and evaluation of systems cannot be simply addressed with one set of measures. A major proposition of our efforts is that one must consider the performance characteristics of the system, how the user interacts with the automation, and the potential safety hazards and consequences that may arise during system operation. These considerations will provide a basis for determining what should be attended to in determining the effectiveness of prototypes and systems during testing and implementation phases.

The overall purpose of our project is to develop a metrics toolkit to guide space flight system designers and developers in selecting and accessing different approaches to measure automation system performance and safety. Part of this effort is to provide information about automation system types that are being developed to conduct a variety of tasks associated with space flight missions, and identifying the needs and requirements of the user to optimally utilize the automation. The categorization of these space flight contexts is an approach to delineate how various systems are unique in their purpose and operation, while also being inclusive of all such mission critical systems that are developed for future space flight missions.

FUTURE DIRECTIONS

The identification of these four space flight task contexts was done as part of a larger project aimed at developing a metrics toolkit for informing the design of HAS for future space flight missions. Currently, progress is being made on a qualitative literature review that categorizes the system contexts from numerous domains into the four categories described above. This literature review takes into account both studies focused on space flight contexts as well as studies focused on other domains and contexts. This will aid in determining whether there are other contexts that can be integrated into our current conceptualization of space flight system contexts.

Additionally, future efforts to validate this conceptualization of the relevant space flight task contexts will include subject matter expert (SME) interviews. These interviews will be used to obtain feedback on our current four categories of space flight task contexts as well as information as to whether this approach of applying measurements based on the task contexts would be useful during the development and design process of space flight systems. If feedback from the SME interviews is positive, we are looking at framing our metrics toolkit to structure information in this way.

REFERENCES

Doody, D., & Stephan, G. (2013). *Basics of Space Flight Section II. Space Flight Projects*. Retrieved from http://www2.jpl.nasa.gov/basics/bsf13-1.php.

Drake, B. G., & Watts, K. D. (2014). Human Exploration of Mars Design Reference Architecture 5.0 Addendum #2 (NASA/SP–2009-566-ADD2). Retrieved from National Aeronautic Space Administration (NASA) website: http://www.nasa.gov/sites/default/files/files/NASA-SP-2009-566-ADD2.pdf

Garland, D. J., Endsley, M. R., Ellison, J., Caldwell, B. S., & Mount, F. E. (1999, September). Space human factors engineering challenges in long duration space flight. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 43, No. 1, pp. 96-97). SAGE Publications.

Manzey, D. (2004). Human missions to Mars: New psychological challenges and research issues. *Acta Astronautica*, *55*(3), 781-790.

Morris, P., Schwabacher, M., Dalal, M., & Fry, C. (2013). Embedding Temporal Constraints For Coordinated Execution in Habitat Automation. Paper prepared for the 8th International Workshop on Planning & Scheduling for Space. Retrieved from https://c3.nasa.gov/dashlink/static/media/publication/msdc-iwpss-2013.pdf

Mosier, K.L., Skitka, L.J., Dunbar, M., & McDonnell, L. (2001). Aircrews and automation bias: The advantages of teamwork? *The International Journal of Aviation Psychology*, *11*(1), 1-14.

Sandal, G. M., Leon, G., Palinkas, L. A. (2007) Human challenges in polar and space environments. Life in Extreme Environments 399–414. doi: 10.1007/978-1-4020-6285-8_25

Sandal, G. M. Leon, G. R., & Palinkas, L. (2006). Human challenges in polar and space environments. *Reviews in Environmental Science and Biotechnology*, *5*, 281-296.

Sheridan, T. B., & Parasuraman, R. (2005). Human-automation interaction. *Reviews of Human Factors and Ergonomics*, *1*, 89-129.