2011 Space Human Factors Engineering Standing Review Panel

Evidence Review Final Report

I. Executive Summary and Overall Evaluation

The 2011 Space Human Factors Engineering (SHFE) Standing Review Panel (from here on referred to as the SRP) met for a site visit in Houston, TX on October 12, 2011 to review the evidence reports for the five new SHFE risks: Risk of an Incompatible Vehicle/Habitat Design (HAB), Risk of Inadequate Human-Computer Interaction (HCI), Risk of Inadequate Design of Human and Automation/Robotic Integration (HARI), Risk of Performance Errors Due to Training Deficiencies (TRAIN), and Risk of Poor Critical Task Design (TASK).

The SRP found that for the most part, the SHFE evidence reports provided justification for all the major space human factors risks anticipated. Some further work is needed in increasing the breadth of evidence, particularly for the HCI, HARI, and TASK risks. Interactions with other Human Research Program (HRP) risks were generally neglected topics in each of the five SHFE evidence reports, and specific non-SHFE risks which interact are called out for the risks reviewed. The interactions between the five SHFE risks and other HRP risks for long-term exploratory missions are detailed in Table 1 at the end of this report.

The level of expertise of the evidence book authors was generally adequate and minor exceptions are noted below.

The SRP does not recommend adding any additional risks at this time, although all SHFE risks' evidence would warrant adding additional gaps.

II. Review of the Evidence for the Risk of Incompatible Vehicle/Habitat Design (HAB)

1. Evaluate the Risk of Incompatible Vehicle/Habitat Design using the following criteria:

A. Does the SHFE Evidence Report provide sufficient evidence that the risk is relevant to long-term space missions?

Yes, the report provides sufficient evidence that the risk is relevant to long-term space missions.

B. Is the risk properly stated in the HRP Program Requirements Document (PRD Rev. E)?

Yes, the risk is properly stated in the HRP PRD Rev. E.

C. Is the text of the short description of the risk provided in the HRP PRD Rev. E clear?

Mostly, but a statement in the overview should be included that the SHFE risks apply to

habitats that may include the launch vehicle, a pressurized suit, or other occupied space (station, non-Earth outpost, re-entry capsule, etc.). Pressurized suit issues are presented in the HAB Evidence Report, but are not uniformly addressed in the subsequent gap analyses. Sometimes a suited crewmember is the subject of the SHFE gaps and sometimes not. For example, gap HAB-01 does not explicitly cite suited crew for acoustic analyses, gap HAB-02 does not cite extravehicular activity (EVA) or suited missions explicitly, gap HAB-03 does not address the suited crewmember (unless we infer "vehicle" can be identified as a pressurized suit), gap HAB-04 does not explicitly address suited HF models, spinal elongation for spacesuit design is explicit in gap HAB-05, and gap HAB-06 is not explicit on suited vs. unsuited issues.

D. *Does the evidence make the case for the knowledge gaps presented?*

There is some lack of explicit evidence for risk factors (or their frequency or severity) involving workload and specific habitation layouts. For example, what tasks had to be aborted due to inadequate strength or access? Were there situations where crews were too exhausted to complete tasks (safely)? Were there medically significant crew trauma issues caused by inadequate habitat design? The situations may not be reported publicly, but we should have indirect evidence that such problems have occurred and thereby individual and team workspace interactions and workloads are indeed significant long-duration mission risks. The main cases cited are bruises and chafing due to pressurized suit activities, but very implicitly; one imagines there are more severe cases for this to be such a major risk class.

E. *Are there any additional gaps in knowledge that should be considered for this specific risk?*

Yes. The SRP does not have an explicit list of tasks expected of crew on long-duration flights. There is little or no discussion of the state of the art in modeling and simulation of habitat designs or the processes by which a robust design is determined. The gaps include knowledge of the physical interior space layout (relative to tasks) and personal/and privacy requirements for crew. Evidence is also missing on how the crew actually navigates about the vehicle and uses hand/foot holds and restraints for task accomplishment, when and how they experience awkward task postures and manage the consequences. Such information would be useful in the design and evaluation of habitat configurations.

F. *Does the SHFE Evidence Report address relevant interactions between this risk and others in the HRP PRD Rev. E/IRP Rev. C (Integrated Research Plan, Rev. C)?*

Some but not all: gap HAB-01 through gap HAB-06 does not integrate with each other, especially the Digital Human Model gap HAB-04. Interactions within teams working in a shared space are not connected to the Behavioral Health and Performance gaps. The habitat influences and interacts with every physical aspect of space operations: work, exercise, nutrition, sleep, and team proximity.

G. *Is the expertise of the author(s) sufficient for the given risk?*

Most of the expertise is sufficient, but not entirely. The discussions of modeling and simulation are almost non-existent. There is no discussion of the current habitation design process that might be useful for exposing gaps in knowledge or process.

H. *Is there information from other disciplines that need to be included in the SHFE Evidence Report?*

Yes, such as, injury prediction methodologies, interior design principles, environmental psychology. For example, is there evidence for repetitive strain injuries other than during EVA suit use?

I. *Is the breadth of the cited literature sufficient?*

The literature cited is limited, often due to the proprietary and private nature of crew anecdotal data. Applicable citations to human modeling and simulation, acoustic modeling and simulation, and current habitat design principles and processes are lacking.

J. *What is the overall quality and readability?*

The overall quality and readability of the report is good, though a bit uneven. A few concepts should be better explained, such as the how the EVA spacesuit gets "positioned" and the usage of the cupola.

2. *Provide comments on any important issues that are not covered in #1.*

In general, the SRP thinks that the evidence is presented via anecdotal inference that a risk is posed by the existence of documented human factors issues. The document lacks any standing as a compendium of knowledge about habitat design for space operations.

There is evidence that stowage is a major problem and hence a HF risk, yet there is no gap that addresses mitigation or management of stowage and waste. Methods to mitigate this risk through conventional technologies ought to be discussed, e.g., via radio frequency identification (RFID) tags, image tags, or camera smart phones to track stored objects.

Is there any evidence that certain fit and task problems may be addressed by reducing the range of allowable crew anthropometry? Or that this is even possible?

III. Review of the Evidence for the Risk of Inadequate Human-Computer Interaction (HCI)

1. *Evaluate the Risk of Inadequate Human-Computer Interaction using the following criteria:*
   A. *Does the SHFE Evidence Report provide sufficient evidence that the risk is relevant to long-term space missions?*

Overall, yes there was sufficient evidence that the risk is relevant in long-term space missions. The evidence is convincing that expanded dependence on conventional HCI for information on long missions will jeopardize them with errors, unpredictable performance on tasks and crew fatigue. There is ample evidence that information deficits have contributed to serious accidents via unusable displays, unavailable information, misallocated attention or cognitive overload, and misperception or misinterpretation. Conventional levels of HCI usability and efficiency, that are acceptable for office work, were never intended for mission-critical/safety-critical domains. Better application of rigorous standards and some new tools are needed to avoid design-induced errors, unpredictable overhead, manual compensation for unmet information requirements, uncontrolled task times to control computer-based information resources, and excessive training.

B. *Is the risk properly stated in the HRP Program Requirements Document (PRD Rev. E)?*

Yes, the risk is properly stated in the HRP PRD Rev. E.

C. *Is the text of the short description of the risk provided in the HRP PRD Rev. E clear?*

No, the description needs clarification. The SRP explains why in comment #2 below.

D. *Does the evidence make the case for the knowledge gaps presented?*

Overall, the evidence does make a case for the knowledge gaps presented. Additional evidence for misperception and misunderstanding of information is available in Jones & Endsley (1996), who found that 20% of situation awareness (SA) errors in aviation fall into the HCI category. (Jones, D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, Space and Environmental Medicine, 67*(6), 507-512.)

The SRP has some additional comments in #2 below.

E. *Are there any additional gaps in knowledge that should be considered for this specific risk?*

Yes, and the SRP addresses these in Comment #2 below. In summary, these include:

- Task overhead from computer use
- Function allocation between humans and computers
- Situation awareness on an individual and team basis

F. *Does the SHFE Evidence Report address relevant interactions between this risk and others in the HRP PRD Rev. E /IRP (Integrated Research Plan, Rev. C)?*

No, the report does not address relevant interactions between this risk and others in the HRP PRD Rev. E /IRP Rev. C. Risk interactions should include: 1) how the design of tasks is constrained by the information architecture and user interface of available

information resources (TASK); 2) how HCI Training is constrained by the complexity of HCI design;
3) how the design of HCI is similar to but distinct from HARI. If the understanding is not clear during design, the confusion will likely propagate to space operations.

G. Is the expertise of the author(s) sufficient for the given risk?

The expertise is mixed, and not sufficient overall. More expertise in system design, evaluation and building HCI into the procurement process is needed. The HCI Evidence Report does not show expertise on the roles of information architecture, familiarity with key relevant aerospace accidents (KAL 007 below), or the Representation Effect (J. Zhang & D.A. Norman, Representations in distributed cognitive tasks. *Cogn Sci*, 18 (1994), pp. 87–122; J. Zhang, The nature of external representations in problem solving. *Cogn Sci*, 21 (1997), pp. 179–217).

H. Is there information from other disciplines that need to be included in the SHFE Evidence Report?

Yes, there are relevant concepts of efficiency and overhead in industrial engineering, the concept of information granularity in artificial intelligence (AI), the concept of the inherent complexity of a task from health informatics, the difference between information architecture and information presentation from software engineering.

I. Is the breadth of the cited literature sufficient?

No, some important literature is omitted, such as Paul Fitts' evidence on design-induced error (Fitts, P. M. & Jones, R. E. (1947) Analysis of Factors Contributing to 460 "Pilot-Error" Experiences in Operating Aircraft Controls. Memorandum Report TSEAA-694-12. Aero Medical Laboratory. Air Material Command. Wright-Patterson Air Force Base, Dayton, Ohio, July 1, 1947) and the investigation of KAL007 in which the autopilot strayed into Soviet airspace twice and was shot down in 1983 killing all 269 onboard (Degani, Asaf (18 September 2001). "Korean Air Lines Flight 007: Lessons From the Past and Insights for the Future". NASA Ames Research Center: NASA).

J. What is the overall quality and readability?

The overall quality and readability of the report is good, though a bit uneven.

2. Provide comments on any important issues that are not covered in #1.

The Department of Defense Human Factors Analysis and Classification System (DoD HFACS) taxonomy that was adapted resulted in considerable overlap in the content of the contributing factors, which suggests arbitrary categorization and undermines its credibility. Also, the eight contributing factors do not flow into an understanding for disposition for HCI very well in many places. A better framework for HCI based on a high-level task analysis is needed to map the risks. This framework should also account for what is common to HCI

and to HARI, and what is distinct to HCI. If uncorrected, the current confusion will likely propagate through systems' designs and surface to disrupt user tasks during space missions.

The introductory sections for the HCI Evidence Report (Risk Overview and Evidence) are somewhat vague and tautological. If there are humans in the loop, and their procedures require information, then that information can be considered a required resource. If a task is attempted without all its needed resources, then its correct performance is at risk. Although information is not technically a resource (because it is not finite to use), information plays a similar role in task performance. All of the contributing factors should be understood in this information work ecology: as factors that jeopardize the human operator's ability to work with the needed information at the right time in the needed situation.

Two factors that seem distinct to using computers that are not treated adequately anywhere in the HCI Evidence Report are uncontrolled overhead from computer use, and function allocation between human and computer.

HCI can accidentally reduce efficiency and undermine the added value of computer functionality by imposing unplanned overhead tasks on the user. Overhead tasks are required only because of the way the software works and are not intrinsic to the nature of the work. For example, if the information architecture does not organize information in a manner that matches a task's information needs, then the user must either perform the overhead task of re-organizing the information, or risk task performance errors. Overhead may also be imposed on users when: the presentation format does not match the correct cognitive strategy for the task; information from diverse resources must be integrated manually; or controlling functionality requires excessive attention. When overhead is imposed in this manner it is difficult to detect and measure (without additional user overhead tasks) and is thus uncontrolled. Uncontrolled overhead is more than just extra work for users; it can disrupt their understanding of the true nature and goal of their tasks. In this view, usability is some inverse function of HCI overhead, and good design should try to identify and eliminate the latter (Zhang, J., & Butler, K. (2007) UFuRT: A work-centered framework and process for design and evaluation of information systems. *Proceedings of HCI International 2007*; Zhang et al., *(2011) J Biomedical Informatics*).

Function allocation is fundamental to HCI design but is often dealt with by automating everything that time and budget will allow, then leaving the rest for the users. This "left-overs" approach to function allocation can impose disjoint tasks on the user that are incoherent or result in awkward workflows. Conversely, synchronizing information flow with workflow is a key aspect of usability. The Introduction of the HCI Evidence Report explains the importance of information architecture, which is the appropriate component to manage information flow, but never refers to it in the body of the evidence report.

The evidence report discusses numerous problems and shortcomings in situation awareness, but this is not addressed in the gaps. Gaps should be added to include tools, methods, and metrics to support appropriate allocation of attention to critical information and support for multi-tasking, for individuals and teams in space.

The section on Modeling and Simulation is too brief to be very informative. It seems more about a solution approach, not a gap, so it seems out of place in the evidence report.

IV. Review of the Evidence for the Risk of Inadequate Design of Human and Automation/Robotic Integration (HARI)

1. *Evaluate the Risk of Inadequate Design of Human and Automation/Robotic Integration using the following criteria:*

   A. *Does the SHFE Evidence Report provide sufficient evidence that the risk is relevant to long-term space missions?*

      Yes, both research studies and incidents are used. The trust in automation is probably the weakest section and should probably be expanded.

   B. *Is the risk properly stated in the HRP Program Requirements Document (PRD Rev. E)?*

      Yes, the risk is properly stated in the HRP PRD Rev. E.

   C. *Is the text of the short description of the risk provided in the HRP PRD Rev. E clear?*

      It should be more specific. A suggested wording: "The risk that systems will not be designed to support appropriate dynamic allocation of tasks and appropriate interactions between humans and automation/robotics maintaining the necessary situation awareness for successful mission completion." It should also be made clear how automation addressed in HARI is different from the normal HCI.

   D. *Does the evidence make the case for the knowledge gaps presented?*

      Yes, but there should be a clear definition of what NASA means when it uses the term "automation" as there are many different ways this could be interpreted. Also, it is necessary to clearly state what HCI and HARI have in common and what differences exist, as many of the gaps in knowledge in HCI apply to HARI as well, while this section specifies the risks that are specific to HARI.

   E. *Are there any additional gaps in knowledge that should be considered for this specific risk?*

      Yes. There are a number of gaps that need to be mitigated to ensure that the human – automation/robotics integration is robust enough to support long-term space missions. Given that this field is rapidly developing and that this risk is the newest in the SHFE portfolio, it is not surprising that the new potential for use of robotics and automation has more knowledge gaps that must be investigated to ensure the safe and effective use of these technologies in space exploration. The list below has been divided into two sections. The first lists knowledge gaps mentioned to some degree in the HARI Evidence Report that were not incorporated into tasks in the IRP Rev. C. The second are knowledge gaps that were not mentioned in the HARI Evidence report.

Knowledge gaps mentioned at least partially in the HARI Evidence Report:

1. Common standards for displays and controls for human - robot interaction, including the necessary information for humans to maintain tasks and robot SA.
2. Tools for evaluating HARI designs in a wide variety of tasks/conditions associated with space operations. Conditions to be addressed need to include normal operations, automation failures and unforeseen conditions, transition support between different levels of automation, levels of situation awareness and workload provided by the designs, suitability of resulting tasks for human performance, and human performance with automation at differing levels of reliability.
3. Guidelines for supporting human understanding of complex systems created by automation and robotic systems.
4. Identification of methods to support development of accurate mental models of automation and decision support tools.
5. Methods to eliminate modes from interaction designs.

Missing knowledge gaps:

1. Adaptive automation appropriate for space exploration.
2. Operating policies for teams of humans and robots.
3. Guidelines for HARI design for space operations.
4. Identification of methods for compensating for loss of SA associated with out-of the loop syndrome from automation.
5. Identification of methods for appropriately calibrating trust in automation and decision support tools based on an understanding of the data used by the automation and decision support tools and the reliability of that data.
6. Identification of operational tasks/conditions most suited to automation and appropriate levels of automation or support for those tasks.
7. Methods to support human understanding of automation level transitions and decisions on when to transition to manual/automated support.
8. Identification of types of robot-to-robot interaction and the necessary situation awareness for humans and robots.
9. Method for eliminating mode errors and supporting automation transparency.
10. Methods for the design of decision support systems to reduce known performance problems (e.g., over-reliance on advice when wrong).

F. Does the SHFE Evidence Report address relevant interactions between this risk and others in the HRP PRD Rev. E /IRP (Integrated Research Plan, Rev. C)?

No, the report does not address relevant interactions between this risk and others in the HRP PRD Rev. E /IRP Rev. C. The display issues and control issues in HCI are also relevant to HARI. Displays for HCI and displays for HARI also need to be consistent. In addition, some HAB issues are relevant. In particular, it will be necessary to identify where in the habitat the operators will work when performing teleoperations. They will need to be able to use any specific device for teleoperation while stabilizing themselves, and will possibly need to be able to monitor a display for situation awareness of task progress. Training and education of the capabilities of automation and robots, including

repair of robots and retention of teleoperation proficiency will be necessary. Other risks that are relevant include:

- Risk of Performance Decrements Due to Inadequate Cooperation, Coordination, Communication, and Psychosocial Adaptation within a Team (TEAM),
- Risk of Compromised EVA Performance and Crew Health Due to Inadequate EVA Suit Systems (EVA),
- Risk of Performance Errors Due to Fatigue Resulting from Sleep Loss, Circadian Desynchronization, Extended Wakefulness, and Work Overload (SLEEP),
- Risk of Adverse Behavioral Conditions and Psychiatric Disorders (BMED), the Risk of Microgravity-Induced Visual Impairment/Intracranial Pressure (VIIP),
- Risk of Impaired Control of Spacecraft, Associated Systems and Immediate Vehicle Egress Due to Vestibular/Sensorimotor Alterations Associated with Space Flight (SENSORIMOTOR),
- Risk of Adverse Health Effects of Exposure to Dust and Volatiles During Exploration of Celestial Bodies (DUST),
- TRAIN, and
- TASK.

G. Is the expertise of the author(s) sufficient for the given risk?

Yes, the expertise is sufficient. One of the authors did her dissertation on telerobotics and worked on Robonaut. Another author worked on automation design at NASA.

H. Is there information from other disciplines that need to be included in the SHFE Evidence Report?

Perception and cognitive psychology are needed for perspective-taking for coordination and teleoperation, decision support expertise, and automation expertise as well as all the disciplines involved in the related risks listed in section F above.

I. Is the breadth of the cited literature sufficient?

No, there is more needed on adaptive, dynamic task allocation, misalignment of the views of humans and robots, displays and controls for human-robot interaction, and teams of humans and robots. Research in a number of areas, such as trust of automation and human-robot teams is evolving and will need to be closely monitored in the coming years.

Additional suggestions for literature include:

1. Supervisory Control of Unmanned Vehicles; Chen, Barnes, & Harper-Sciarini, ARL TR 5136, April 2010.
2. A Preliminary Study of Peer to Peer Human Robot Interaction, Fong et al., Systems, Man and Cybernetics, 2006. SMC '06. IEEE International Conference on 8-11 Oct. 2006, pp. 3198 - 3203 ISBN: 1-4244-0099-6.
3. M. Barnes & F. Jentsch (2010) *Human-robot interaction in future military operations*. London: Ashgate.

4. Endsley, M. R., & Jones, D. G. (2011) Unmanned and remotely operated vehicles. In: *Designing for situation awareness: An approach to human-centered design (2nd ed.).* London: Taylor & Francis.
5. Kaber, D.B. & Endsley, M.R. The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science,* Vol 5(2), Mar-Apr 2004, 113-153.
6. Macedo, J. A., Kaber, D. B., Endsley, M. R., Powanusorn, P. A., & Myung, S. (1998). The effect of automated compensation for incongruent axes on teleoperator performance. *Human Factors*, 40(4), pp. 541-553.
7. Drury, J.L., Scholtz, J., & Yanco, H.A., Awareness in human-robot interaction. Systems, Man and Cybernetics, 2003, IEEE International Conference on Oct. 5-8, 2003 Vol. 1, 912-918 ISBN 0-7803-7952-7.

Frame-of-reference/misalignment issue in minimally-invasive surgery references include:

1. DeLucia, P. R. & Griswold, J. A. (2011). Effects of Camera Arrangement on Perceptual-Motor Performance in Minimally Invasive Surgery. *Journal of Experimental Psychology: Applied*, 17, 210-232.
2. Emam, T. A., Hanna, G., & Cuschieri, A. (2002a). Comparison of orthodox versus off-optical axis endoscopic manipulations. *Surgical Endoscopy*, 16, 401–405.
3. Holden, J. G., Flach, J. M., & Donchin, Y. (1999). Perceptual-motor coordination in an endoscopic surgery simulation. *Surgical Endoscopy,* 13, 127–132.

J. What is the overall quality and readability?

The overall quality was good for the gaps that are stated, but it needs to be augmented to support the additional knowledge gaps that need to be investigated to support safe and effective space exploration.

2. Provide comments on any important issues that are not covered in #1.

The SRP thinks that the issues below also need to be investigated. These have been suggested as additional tasks in the 2011 SHFE SRP Research Plan Review:

- Differences between stationery vs. locomotor robot control, and context of use issues
- Business or process rules needed for adaptive automation
- Need a policy to keep modes out of interaction design
- Need a common human-robot interaction standard and design patterns

V. Review of the Evidence for the Risk of Performance Errors Due to Training Deficiencies (TRAIN)

1. Evaluate the Risk of Performance Errors Due to Training Deficiencies using the following criteria:

A. Does the SHFE Evidence Report provide sufficient evidence that the risk is relevant to long-term space missions?

The evidence provided supports the risk identified; however, an additional gap should be considered to deal with the potential degradation of skilled learning due to reliance on the automation identified in HARI and HCI.

B. *Is the risk properly stated in the HRP Program Requirements Document (PRD Rev. E)?*

The risk identified in the PRD Rev. E is properly stated as presented; however, discussion during the review identified a number of additional risk areas that should be considered for inclusion. In particular, current risks deal exclusively with training to do specified tasks and procedures. For long-duration exploration, crews will be expected to: deal with emergent problems not specifically identified beforehand; be responsive to emergent opportunities; and be able to develop appropriate behaviors and procedures as the situation may demand. In order to achieve this, in addition to training for specific tasks, the crew will require more in-depth education in spacecraft systems and overall operation as well as more extensive cross-training among members.

C. *Is the text of the short description of the risk provided in the HRP PRD Rev. E clear?*

The short descriptions are clear enough, but could benefit from more specificity with regard to the additional training risks suggested above.

D. *Does the evidence make the case for the knowledge gaps presented?*

The evidence makes a good case for the knowledge gaps identified, but the gaps as stated are too vague to adequately guide project selection.

E. *Are there any additional gaps in knowledge that should be considered for this specific risk?*

There are clearly training risks associated with highly flexible and adaptive automation and robotics identified in HARI, TEAM, and HCI (e.g., human-automation teaming, trust and confidence). Direct consideration of these will generate a number of more focused gaps that should be addressed. Another potential gap relevant to long-duration missions is the need for more in depth education to augment specific skill training in order to deal with emergent situations in conditions of long communication delays. Additionally, gaps and training strategies for dealing with social, medical, and psychological issues that may emerge in long-duration missions need to be considered. Some examples to consider are: how to deal with the differences between expectations and emergent conditions in space; how to develop skill-based training to replace task rehearsal for enhancing crew flexibility; when to use classroom training vs. computer-based training or other media; methods that would permit crews to develop their own training; strategies to initially teach and later maintain skills in problem solving and decision making on long-term space missions.

F. *Does the SHFE Evidence Report address relevant interactions between this risk and others in the HRP PRD Rev. E /IRP (Integrated Research Plan, Rev. C)?*

A long-existing issue for training programs in general is that training is often considered late in development, with little involvement during system evolution, particularly in the early and middle stages. The TRAIN Evidence Report continues to promulgate this unfortunate strategy, the effect of which is much more likely to produce negative results in long-duration missions with reduced communications. It would be very beneficial to address the need for early and ongoing interaction with TRAIN during HRI, HARI, and TASK development. Similarly, earlier EVA involvement with TRAIN has the potential to provide significantly better understanding of suit capabilities. Not specifically addressed is training for strategies to help overcome physical and mental challenges likely to be associated with long-duration missions. In this regard, interactions with the TEAM, SLEEP, SENSORIMOTOR, BMED, and the Risk of Acute and Late Central Nervous System Effects from Radiation Exposure (CNS) would also be beneficial.

In summary: *training should be considered much earlier and in more areas for the kinds of missions envisioned.*

G. *Is the expertise of the author(s) sufficient for the given risk?*

The authors have the needed expertise to deal with the stated risks, but should the additional training issues previously mentioned be incorporated in the program, there may be a need to include additional expertise within the team.

H. *Is there information from other disciplines that need to be included in the SHFE Evidence Report?*

Assuming that some of the additional risks and gaps are included there will be a need to work with cognitive psychologists and team performance specialists as well as subject matter experts to give advice on cross-training for multiple skills. There may be useful information to be derived by understanding how military Special Forces deal with cross-training. Although the missions are significantly different in time span and venue, many of the training issues dealing with risk management, responding to emergent situations, cross training, and team behavior are relevant.

I. *Is the breadth of the cited literature sufficient?*

The literature cited is sufficiently broad for the current thrust of training. However, it would be very beneficial to consider the transfer of training, evaluation, and team performance literature if some of the suggested risks and gaps were to be adopted. Clearly additional literature will need to be included.

J. *What is the overall quality and readability?*

The overall quality of the writing and readability are good.

2. *Provide comments on any important issues that are not covered in #1.*

The SRP has no additional comments for this risk.

VI. Review of the Evidence for the Risk of Poor Critical Task Design (TASK)

1. *Evaluate the Risk of Poor Critical Task Design using the following criteria:*

A. *Does the SHFE Evidence Report provide sufficient evidence that the risk is relevant to long-term space missions?*

No, problems are described in each section below.

B. *Is the risk properly stated in the HRP Program Requirements Document (PRD Rev. E)?*

The risk as stated does not match the risk title or the evidence. The current title makes it impossible to separate this risk from HARI and HCI, which significantly impact the design of tasks, leading to confusion. The SRP recommends renaming it "Risk of inability to perform required work necessary to accomplish mission goals."

C. *Is the text of the short description of the risk provided in the HRP PRD REV. E clear?*

Yes, the short description of the risk is clear.

D. *Does the evidence make the case for the knowledge gaps presented?*

No. The evidence has a long section on state-of-the-art in workload measurement and good work done at NASA on procedures. It does not clearly lay out the actual risks being addressed in this risk area. This section should be rewritten and reorganized to clearly state: (a) why current workload models and metrics are insufficient for future space operations (e.g. changes needed to reflect workload capabilities over long-duration missions), (b) what additional work is needed to support future procedures and tools (e.g. changes in current approach needed to reflect increased human autonomy and ability in performing tasks in future operations), and (c) what information needs to be collected and included in human performance models to make them applicable to future missions (e.g. the specific perceptual, cognitive and physiological changes that need to be accounted for).

E. *Are there any additional gaps in knowledge that should be considered for this specific risk?*

The SRP thinks that there are two additional gaps in knowledge that should be considered for the risk:

1. Validation of workload metrics as predictors of ability to accomplish long-term mission tasks.
2. Possible tradeoffs between paper and electronic media (what are the criteria for media use in task aids to support task procedures).

F. *Does the SHFE Evidence Report address relevant interactions between this risk and others in the HRP PRD Rev. E /IRP (Integrated Research Plan, Rev. C)?*

The report needs to specifically discuss linkages to HAB (lighting), TRAIN, HCI, HARI, SLEEP, VIIP and DUST (ability to see task aids), EVA (suits), and BMED (cognitive decrements).

G. *Is the expertise of the author(s) sufficient for the given risk?*

The authors are strong in the area of procedures, but they may need more help in updating human performance models.

H. *Is there information from other disciplines that need to be included in the SHFE Evidence Report?*

Yes, all of the related risk areas listed in section F above.

I. *Is the breadth of the cited literature sufficient?*

No. Specific literature associated with changes due to new long-duration missions needs to be included.

J. *What is the overall quality and readability?*

Organization and readability was poor in that the evidence did not link well to the stated risks and those risks do not map well to the stated tasks.

2. *Provide comments on any important issues that are not covered in #1.*

- There is a missing gap to address how/whether workload models correlate with actual task performance and productivity. Good validation of the workload models (including how the workloads of individual tasks combine) is the Achilles heel of these models. It will be particularly important to extend the validation to address how the environmental and physiological changes associated with long-duration missions may change assumptions in the models.
- There is a missing gap to address task design for multiple operators, including mixed human-automation/robotic teams. Current models tend to be concentrated on the workload incurred by a given individual. The optimal means of task-shedding or distributing workload across multiple team members, including robotic team members, needs to be addressed.
- There is a missing gap to address the effect of longer and more variable time delays on task execution on longer missions. In longer-duration missions, greater time delays will occur in human-human communications and/or automated instructions relayed from Earth. These time delays will have an impact on human task performance and tendency to try to perform tasks or parts of tasks without waiting for delayed communications. These factors need to get worked into human performance models and procedures.

- There is a missing gap to address the effect of much greater reliance on automation (vs. human mission control) on task execution. Given that a critical factor for the design of tasks is that of function allocation between the human and the automation, this linkage with HARI is critical. As humans become more reliant on automation, there is also a problem of becoming less capable of performing the tasks manually (skill degradation), as well as out-of-the-loop problems in understanding when it will be important to do the task manually. These factors will affect task performance assumptions regarding task time, error rates and perceived workload.

TABLE 1: INTERACTIONS BETWEEN SHFE RISKS AND OTHER HRP RISKS

	HAB	HARI	HCI	TRAIN	TASK
BHP – Bmed	X	X	X	X	X
BHP – Sleep	X	X	X	X	X
BHP – Team	X	X	X	X	
HHC – Bone					
HHC – Cardio	X				
HHC - EVA	X	X	X	X	X
HHC – Immune					
HHC - IVD					
HHC – Muscle					
HHC – Nutrition	X				
HHC – Occupant Protection	X				
HHC - Pharm					
HHC – Sensorimotor	X	X	X	X	
HHC-VIIP	X				X
SHFH-HAB	N/A	X	X	X	X
SHFH-HARI	X	N/A	X	X	X
SHFH-HCI	X	X	N/A	X	X
SHFH-TRAIN	X	X	X	N/A	X
SHFH-TASK	X	X	X	X	N/A
SHFH - Dust	X	X			X
SHFH – Food					
SHFH - Microhost					
Space Radiation – Acute					
Space Radiation – Cancer					
Space Radiation – CNS				X	
Space Radiation – Degenerative					
Space Radiation – Acute					

VII. SHFE SRP Evidence Review Statement of Task

The 2011 Space Human Factors Engineering (SHFE) Standing Review Panel (SRP) is chartered by the Human Research Program (HRP) Program Scientist at the NASA Johnson Space Center (JSC) to review the evidence base for the five new SHFE risks. The 2011 SHFE SRP will generate a report of their analysis of the SHFE evidence, including any recommendations on how to improve the SHFE Evidence Reports, and submit it to the HRP Program Scientist.

In 2008, the Institute of Medicine reviewed NASA's Human Research Program Evidence in assessing the SHFE risks identified in NASA's Human Research Program Requirements Document (PRD). Since this review there was a major re-evaluation of the SHFE evidence. This SRP is being asked to review the latest version of the five SHFE Evidence Reports.

The 2011 SHFE SRP is charged to:

1. Evaluate each of the SHFE Evidence Reports based on the following criteria:
   A. Does the SHFE Evidence Report provide sufficient evidence that the risk is relevant to long-term space missions?
   B. Is the risk properly stated in the HRP Program Requirements Document (PRD)?
   C. Is the text of the short description of the risk provided in the HRP PRD REV. E clear?
   D. Does the evidence make the case for the knowledge gaps presented?
   E. Are there any additional gaps in knowledge that should be considered for this specific risk?
   F. Does the SHFE Evidence Report address relevant interactions between this risk and others in the HRP PRD REV. E /IRP (Integrated Research Plan, Rev. C)?
   G. Is the expertise of the author(s) sufficient for the given risk?
   H. Is there information from other disciplines that need to be included in the SHFE Evidence Report?
   I. Is the breadth of the cited literature sufficient?
   J. What is the overall quality and readability?

2. Provide comments on any important issues that are not covered in #1.

Additional Information regarding this review:

3. Once the 2011 SHFE SRP members have received the review materials and had the opportunity to look over the documents, the panel members will participate in a teleconference to discuss any issues, concerns, and expectations of the review process to start the review.
   A. Discuss the Statement of Task and answer questions about the process.
   B. Identify any issues/concerns the 2011 SHFE SRP would like to have addressed.

4. Attend a meeting at NASA JSC in October 2011 to discuss Evidence Reports with the SHFE Project. At this meeting, prepare a draft report, including any recommendations. Debrief the HRP Program Scientist on what will be included in final report and should address #1 and #2 above.

5. Finalize the panel report (within one month of the debrief) that contains a detailed evaluation of the risks organized by Items #1 and #2 above. The report will be sent to the HRP Program Scientist. A copy of the report will be provided to the Space Human Factors and Habitability Element at JSC that sponsors the SHFE discipline. Once the report is finalized it will be made available to the public.

VIII. SHFE SRP Roster

Panel Chair:

Anna M. Wichansky, Ph.D.
Applications User Experience
Oracle
500 Oracle Parkway-MS30p3
Redwood Shores, CA 94065
Ph: 650-506-8007
Email: anna.wichansky@oracle.com

Panel Members:

Norman I. Badler, Ph.D.
Department of Computer and Information Science
University of Pennsylvania
3330 Walnut Street
Philadelphia, PA 19104-6389
Ph: 215-888-5862
Email: badler@seas.upenn.edu

Keith A. Butler, Ph.D.
College of Engineering
University of Washington
Box 352315
Seattle, WA 98125-2315
Ph: 206-329-1981
Email: keith.a.butler@gmail.com

Patricia R. DeLucia, Ph.D.
Psychology Department
Texas Tech University
Boston Ave and 18th Street
Lubbock, TX 79409-2051
Ph: 806-742-3711 x259
Email: pat.delucia@ttu.edu

Mica R. Endsley, Ph.D.
SA Technologies
3750 Palladian Village Drive
Building 600
Marietta, GA 30066
Ph: 770-565-9859
Email: mica@satechnologies.com

Jean C. Scholtz, Ph.D.
National Security/Information Analysis
Battelle Memorial Institute
340 Northslope Way
Rockaway Bench, OR 97136
Ph: 503-355-2792
Email: jean.scholtz@mindspring.com

Randall P. Shumaker, Ph.D.
Institute for Simulation and Training
University of Central Florida
3100 Technology Parkway
Orlando, FL 32826
Ph: 407-882-1301
Email: Shumaker@ist.ucf.edu