Human Research Program
Space Human Factors Engineering (SHFE)
Standing Review Panel (SRP)
Final Report

December 2009

I. Executive Summary & Overall Evaluation

The Space Human Factors Engineering (SHFE) Standing Review Panel (SRP) evaluated 22 gaps and 39 tasks in the three risk areas assigned to the SHFE Project. The area where tasks were best designed to close the gaps and the fewest gaps were left out was the *Risk of Reduced Safety and Efficiency due to Inadequate Design of Vehicle, Environment, Tools or Equipment*. The areas where there were more issues with gaps and tasks, including poor or inadequate fit of tasks to gaps and missing gaps, were *Risk of Errors due to Poor Task Design* and *Risk of Error due to Inadequate Information*.

One risk, the *Risk of Errors due to Inappropriate Levels of Trust in Automation*, should be added. If astronauts trust automation too much in areas where it should not be trusted, but rather tempered with human judgment and decision making, they will incur errors. Conversely, if they do not trust automation when it should be trusted, as in cases where it can sense aspects of the environment such as radiation levels or distances in space, they will also incur errors. This will be a larger risk when astronauts are less able to rely on human mission control experts and are out of touch, far away, and on their own. Aha et al. (2005) describe relevant Naval Research Laboratory studies on techniques for building trust in human computer interaction. See also Lee & See (2004) and Parasuraman & Riley (1997).

The SRP also identified 11 new gaps and five new tasks.

Although the SRP had an extremely large quantity of reading material prior to and during the meeting, we still did not feel we had an overview of the activities and tasks the astronauts would be performing in exploration missions. Without a detailed task analysis and taxonomy of activities the humans would be engaged in, we felt it was impossible to know whether the gaps and tasks were really sufficient to insure human safety, performance, and comfort in the exploration missions. The SRP had difficulty evaluating many of the gaps and tasks that were not as quantitative as those related to concrete physical danger such as excessive noise and vibration. Often the research tasks for cognitive risks that accompany poor task or information design addressed only part, but not all, of the gaps they were programmed to fill. In fact the tasks outlined will not close the gap but only scratch the surface in many cases. In other cases, the gap was written too broadly, and really should be restated in a more constrained way that can be addressed by a well-organized and complementary set of tasks. In many cases, the research results should be turned into guidelines for design. However, it was not clear whether the researchers or another group would construct and deliver these guidelines.

The SRP felt that researcher access to astronaut data, errors, critical incidents, and human factors issues presently sequestered within the NASA astronaut office was absolutely essential to make progress in future research. Such data would also facilitate the development of ground-based analogs, which is another major constraint of SHFE.

The SRP also felt that some superfluous legacy research projects were evident in the tasks. NASA might achieve a better cost-benefit ratio by ending a set percentage of projects every year; i.e., the least productive projects or those that were obsolescent based on current mission goals.

Regarding NASA's collaboration with outside organizations, the SRP felt that the National Space Biomedical Research Institute (NSBRI), a research task grantee, took a clinical (i.e., personnel screening) rather than human factors approach to certain gaps that was inappropriate (see p. 10, Advanced Displays for Efficient Training). NASA should recommend they use a human factors expert to confer on issues such as spatial disorientation.

NASA should consider the use of consultants or focused workshops to obtain greater efficiency in filling some of the gaps, as some do not require research but rather a systematic engineering approach to apply available technology (see p. 12, SHFE 3.1.2.3.2 for example).

NASA should communicate more with the Jet Propulsion Laboratory (JPL) concerning robotic research. See p. 7, Human Automation and Robotics.

A systematic approach, including considerations of both flight crew and ground crew, should be implemented throughout the research projects.

II. *RISK OF REDUCED SAFETY AND EFFICIENCY DUE TO AN INADEQUATELY DESIGNED VEHICLE, ENVIRONMENT, TOOLS OR EQUIPMENT*

GAPS AND TASKS:

SHFE 2.1.8.1: How can we determine the effects of combined vibration and acceleration on task performance?

- This gap is adequately scoped.
- Vibration can affect cognition, such as attention and working memory. The tasks to be performed by human participants should reflect this.
- The SRP recommends that the results of these studies be turned into guidelines, but was unsure who reviews these results or develops the guidelines.

Current Tasks:
Effects of Vibration and Acceleration on Performance

- This task is adequate to fill the gap.

Effects of Vibration and Acceleration on Performance

- This task should include some of the more complex cognitive tasks, attention, and working memory.

Effects of Vibration During Launch (SDBI 1904)

- The effects of vibration on oculomotor vs. cognitive perception during launch should be teased out.

SHFE 2.1.9.1: How do we develop validated tools and models to verify Constellation vehicle acoustic design environment requirements?

- This gap is adequately scoped.

Current Task:
Acoustics Modeling

- If humans are not elements included in the current model, they should be. This will change the shape of the acoustics envelope.

SHFE 2.1a: What are the effects of habitable volume and architecture on safety and performance and how can an integrated evaluation of those effects be performed?

- The gap is too broad. Unless it is part of the task to fill the gap, the aspects of safety and performance, and the tasks crew will be performing in the habitable space should be operationally defined.

Current Task:
Effects of Architecture on Safety and Performance

- NASA should continue to develop computational models of crew physiological and psychological changes that occur in long-term spaceflight and microgravity conditions (e.g., spinal elongation).
- Such human models should work within typical NASA Computer Aided Design (CAD) systems.
- These human models should be given access to as many properties as possible from the NASA STD-3000 man-system integration efforts.
- Changes to human performance brought on by mission conditions, such as spinal elongation, muscle weakness, and disruption of circadian rhythms, should be incorporated into the human models so that mission planners can use the integrated architecture and human models for safety and performance assessments and decision-making.
- Efficiency should have been defined quantitatively.
- The need for a decision support tool was not clearly stated.
- Boeing had a successful project to evaluate whether a physical maintenance task could be performed in a confined airliner space. It involved using a mannequin to put a human model into a CAD space and then making it do the movements of the task (Esposito et al., 1995).

SHFE 2.3.a (SBIR): How can crews easily document human factors related issues that occur on orbit?

- This gap should be re-scoped to include all mission phases, not just on-orbit.
- What does NASA do with the data once it is collected? There needs to be an easy means of access, such as the Aviation Safety Reporting System. Will this data be analyzed and used to refine guidelines for future designs?

- The reporting tool needs to be evaluated from both the input perspective and from the perspective of the human factors personnel who will be trying to analyze and use the data.
- Cross reference to Gap 1.1.2.2.1 (standardizing measure and metrics for user interface evaluation).
- One technique to minimize the overhead of entering data after the fact is to make it incidental to the performance of the task. Incidental data entry is when data recording can be accomplished with the same action as fixing the problem. When the task involves human-computer interaction, there are existing techniques for collecting behavioral data by instrumenting the software, i.e. embedding code in the application software to log and time stamp user commands and system responses.

<u>Current Task:</u>
Semantic Language and Tools for Reporting Human Factors Incidents

- The SRP was not sure how useful event data from the astronauts would be without any interpretation of such events, because astronauts are not human factors experts.
- This task only weakly fills part of the gap. There should be standards for building and searching such reports and relating their contents to visual images of the physical layout or current condition of the environment.

SHFE 2.3.1.1: What is the effect of microgravity on spinal elongation?

- The gap is adequately scoped.

<u>Current Task:</u>
Spinal Elongation and its Effects on Seated Height in a Microgravity Environment (Spinal Elongation)

- The task is adequate to fill the gap, assuming data is available from past missions on non-seated postures.
- Results of this task should feed Gap 2.1a (habitable volume).

SHFE 3.1.2.2.a: How do we ensure that the displays and control designs and technology developed for the operational environments of the Cx Program will improve performance and reduce errors?

- The SRP recommends rewording the gap to read: What are the optimal requirements for displays and controls for robust and consistent crew performance? (If it truly is about performance *improvement*, it needs to address current levels of performance and why they are not adequate.)
- There is no mention of use of acoustic alarms. These are also displays.
- It seems like there should be four subtasks:
    - Identify common tasks across vehicles. The tasks to be performed should drive display and control design.
    - Develop knowledge of best practices that is appropriate to information display for those tasks.
    - Develop guidelines and requirements for displays and controls.
    - Validate the guidelines using mockups or prototypes.

Current Tasks:

Information Presentation – Displays Development (Visual and Auditory) (Info Pres)

- There are other relevant reports by Zhang and Norman on external representation and Endsley, et al. (2003) that address situation awareness issues. See Zhang and Norman (1994) and Zhang (1997) for research on how to integrate external information with internal knowledge. See Butler, et al (2007) for a case study that applied these display principles to design a mixed initiative scheduling system.
- Deliverables need to include what tasks they are developing displays for, unless these are given by NASA or other research.

Information Presentation – Controls Technology Survey and Testing (Info Pres)

- Human participant testing should be included.

Information Presentation – Electronic Procedures and Fault Management (Info Pres)

- Evaluation of prototypes with human participants should be included.
- There should be a detailed task analysis of what the operator will be doing using the electronic procedures viewer (EPV).
- How do you operate the EPV when seated with arms constrained? Would there be a speech interface to supplement hand control?
- In terms of schedules, Gap 2.1.8.1 (vibration control) will not be closed when this task is going on, and those results will have an important effect on the outcome of this task.

Model Development and Applicability

- This task will fill the gap.
- Models also need to account for known changes in physiology and psychology due to long-duration missions.
- How can software modeling tools interface with known human data (e.g., impairments in strength over time due to microgravity)? What is the easiest way to import data or new information about human capabilities into these models?

SHFE 2.3.b: How can existing models be modified to adequately represent the specified user population (e.g. field of view, visibility) in reduced gravity and be portable to other simulations environments?

- This gap is adequately scoped.
- What models are they using?

<u>SUGGESTED NEW GAPS:</u>

- **What tasks will be performed in the exploration missions that may be subject to performance degradation due to a poorly designed vehicle, environment, tools, or equipment?**
- **What causes "cognitive fog" in spaceflight, and what are some mitigation approaches?**
- **What computer-readable databases could be used in modeling human performance, such as human factors information on physiological and psychological changes that may occur in long-term missions?**
- **What is the effect of long duration missions on human performance in space? (may be covered under Behavioral Health and Performance (BHP) or Human Health Countermeasures (HHC). How could databases be used to monitor changes?**

<u>SUGGESTED NEW TASKS:</u>

- **Develop guidelines for integrated information displays, to minimize the number of discrete, different user interfaces, while optimizing clarity in portraying the needed information necessary for each application.**
    - What are the tasks that need to be supported by integrated information displays?
    - What are the categories and types of information needed in the displays to ensure situation awareness, variation in level of detail, explicit indication of human-automation division of control and so forth?
    - These guidelines can be applied to the procedures viewer (EPV), robot controllers, system engineering diagram viewers, among others.

- **Develop a modeling tool to support higher level aspects of display design, like situation awareness, information visualization, and multimedia, multisensory displays.**
    - <u>Model Development and Applicability</u> could be extended to address this.

III. *RISK OF ERRORS DUE TO POOR TASK DESIGN*

<u>GAPS AND TASKS:</u>

SHFE 1.1.1.2.1: How can standard measures or tools be developed that can unobtrusively measure workload?

- This gap is adequately scoped.
- The importance of measures being unobtrusive is not clear. Should obtrusive as well as unobtrusive measures be considered, especially since there are standardized measures of workload?
- Is the distinction really invasive vs. non-invasive methods?
- The measures should also help localize parts of the task where workload could be reduced, such as non-essential overhead.

Current Tasks:
Workload Tools and Guidelines. (Workload)

- There should be a clear and consistent definition of what type of workload will be measured: physical, motor, cognitive, sensory? These may require different measures, and some may be obtrusive (e.g. galvanic skin response (GSR), electroencephalography (EEG)).
- Workload must be addressed for both the individual and teams over time.
- The impact of workload on resource utilization (i.e., efficiency) should be addressed.

Operator Performance Models. (Workload)

- The SRP assumes there is one task although there appears to be a typo on page 28 of the SHFE Integrated Research Plan (IRP) Update. (In the middle of p. 28 it says Directed Collaboration with HRP, BHP.)
- What are acceptable levels of workload?
- Existing standards of acceptable task workload should be updated for long-term missions.

SHFE 1.1.1.3.1: How do we design tasks to ensure adequate situational awareness?

- The SRP recommends rewording the gap to read: How do we ensure adequate situation awareness on tasks? (The tasks themselves cannot be designed to affect Situation Awareness (SA)).
- Is there a comprehensive list of crew tasks?

Current Task:
Situational Awareness Evaluation

- Individual and team SA should be considered.

SHFE 1.1.2.1.1: How can performance, efficiency, and safety guidelines be developed for appropriate task automation and the effective allocation of tasks between humans and automation?

- The gap is extremely broad, and requires task taxonomy and analysis to know what tasks will be performed in space, as a precursor to human-machine task allocation efforts.

Current Tasks:
Automation Interface Design Development. (AIDT)

- This task seeks to effectively create an automated user experience designer or expert system. User experience design is a combination of engineering, psychology, graphics and industrial design, and computer science. This goal is really not achievable, because it is difficult to automate a task which combines so many high level human cognitive capabilities, especially design creativity.
- This is not a solid technical approach to achieving high quality user experiences.

Human/Automation and Robotics Operations and Tasks Concepts Design

- This is more realistic than the previous task.

- By the time this task is completed, robotics technologies will have changed dramatically. This should be anticipated and considered in the design of operations and tasks.
- This task is redundant with other NASA work. The JPL Space Robotics project (Mitch Ingram) and NASA Ames Research Center (ARC) (Mike Feary and Dorrit Billman) are already doing work in this area.
- Closing this gap hinges on being able to specify the fundamental work requirements independent of any given design concept. Current techniques, like task analysis or process modeling, are not sufficient for concept design because they build in premature assumptions about key design decisions such as a function allocation and use strategy. Knowledge modeling techniques have shown good potential for declarative modeling of "what" independently of "how. Ontology modeling is used in biomedical informatics (Rosse & Mejino, 2003), and recently for modeling the entity of cognitive work for scheduling (Butler et al., 2007). Algorithm verification (i.e., model checking) has good potential for exploring design concepts much more economically (Clarke et al., 2009 Turing Award Lecture).
- Goal-directed task analysis (GDTA) has also been shown to be effective for formulating technology-independent requirements (the what, not how) in many domains (see Endsley, 1993, 1995, 2003). It derives a user goal hierarchy, from which user decisions and SA requirements are specified. This provides a way to drive the design process without being tied to how they do things today or limited by artificial ceiling effects on what information they need.

SHFE 1.1.2.2.1: How can we develop standard measurement techniques and metrics for evaluating the quality of user interfaces with specific attention to the usability of an interface?

ALL OF THIS IS DISCUSSED UNDER SHFE 3.1.2.2.1.2: How can we develop standard measurement techniques and metrics for evaluating the quality of user interfaces with specific attention to the usability of an interface? (See p. 11)

Current Tasks:
Usability Evaluation. (Usability)
Operator Performance Models. (Usability)

SHFE 3.1.2.2.a: How do we ensure that the displays and control designs and technology developed for the operational environments of the Cx Program will improve performance and reduce errors?

THIS IS COVERED ON P. 4 SECTION 3.1.2.2.a IN THE PREVIOUS RISK.

Current Tasks:
Information Presentation – Displays (Visual and Auditory). (Info Pres)
Information Presentation – Controls. (Info Pres)

SUGGESTED NEW GAPS:

- **What is the complete list of tasks that should be designed for the exploration missions? What are the steps to accomplish them?**
- **How can we support spatially distributed and temporally asynchronous teams in task performance and collaboration? (This could be related to SHFE 3.2b, p. 13).**

IV. *RISK OF ERRORS DUE TO INADEQUATE INFORMATION*

GAPS AND TASKS:

SHFE 3.1.1.a: How can we develop objective training measures to determine operator proficiency during and after ground training?

- Is the issue about training or about the ability to find information?
- It wasn't clear what tool would be used to develop the information.
- The description is very crew focused; should mission control be included in the objectives?

Current Tasks:
Spaceflight Resource Management Training. (Training- SFRM)

- This task does not address the gap. It is not clear how paper-based games can help develop objective training measures.
- The presentations and addendums sent to the SRP indicated this task was to train flight surgeons. Yet the IRP supplement seems to address mission control.
- The SRP does not have enough information to evaluate this task.

Training Proficiency Methods Development

- This task seems to be an extension of the previous task. It was unclear why they were separated.

SHFE 3.1.1.b: How do we develop training methods and tools for space medical application if time is minimal?

- The gap is adequately scoped.

Current Task:
Medical Proficiency Training. (Training – Medical)

- Dave Kieras (1998) has done research on how to model procedures, dubbed Natural Goals-Operators-Methods-Selection Language (NGOMSL) modeling in the human-computer interaction literature.

SHFE 3.1.1.c: How can onboard training systems be designed to address Just in Time (JIT) and recurrent training needs for nominal and off nominal scenarios?

- The gap is adequately scoped.

Current Task:

Just in Time Training Methods/Model

- How will the crew *find* the training content they need in time?

SHFE 3.1.2.a: How can a capability for semi-autonomous planning and dynamically replanning of crew schedules be developed?

- The gap is adequately scoped.

Current Task:
Science Planning Interface to Engineering Scheduling Tool (SPIFE Scheduling)

- The research approach is applicable, and the task is almost done.

SHFE 3.1.2.2.a: How do we ensure that the displays and control designs and technology developed for the operational environments of the Cx Program will improve performance and reduce errors?

- ALL OF THIS IS ADDRESSED IN SECTION 3.1.2.2.a ON P. 4. (However, the tasks are not the same in both places.)

Current Tasks:
Information Presentation – Displays (Visual and Auditory). (Info Pres)

Information Presentation – Controls Technology Survey and Testing (Info Pres)

Information Presentation – Human Performance Modeling (Info Pres)

Cognitive Performance Modeling for Display and Procedure Evaluation

- It is good that it adds human oculomotor aspect.

Sensorimotor Displays and Controls to Enhance the Safety of Human/Machine Cooperation During Lunar Landing. (NSBRI – Sensorimotor Adaptation Team)

- Efforts should be combined with projects of Duda[1] and Young[2] (both NSBRI), because these tasks have significant overlap.

Advanced Displays for Efficient Training and Operation of Robotic Systems (NSBRI Sensorimotor Adaptation Team)

- This task seems unrelated to advanced displays and should not be a part of the SHFE Project (it could be moved to BHP). This is not a crew selection problem, but a display problem and the task does not address displays, but rather its goal is to develop spatiomotor skills screening techniques to select the best robotic systems operators.

---

[1] Human-Automation Interactions and Performance Analysis of Lunar Lander Supervisory Control. Kevin R. Duda, Ph.D. The Charles Stark Draper Laboratory, Inc

[2] Sensorimotor Displays and Controls to Enhance the Safety of Human/Machine Cooperation During Lunar Landing. Laurence R. Young, Sc.D. Massachusetts Institute of Technology

- This task does not adequately address the gap. There should be a human-automation work allocation task to address this gap, and personnel selection screening should be addressed in a BHP task.

SHFE 3.1.2.2.1.2: How can we develop standard measurement techniques and metrics for evaluating the quality of user interfaces with specific attention to the usability of an interface?

- The SRP did not consider this to be a gap. Research has already been done and there are significant guidelines and standards for evaluations in the human computer interaction and usability engineering fields that can be applied.
    - International Organization for Standardization (ISO) 13407 discusses definitions of usability and usage quality which will be applicable here.
    - ISO 9241 is a multi-part standard that addresses design and evaluation of components of software user interfaces such as menus.
    - ISO 25062 specifies how to conduct a usability test to measure effectiveness, efficiency, and satisfaction. See http://zing.ncsl.nist.gov/iusr/

- How are these tasks going to be done?
- This gap is redundant with SHFE 1.1.2.2.1 (p. 8).

Current Tasks:

Usability Evaluation. (Usability)

- There needs to be a task taxonomy. Usability cannot be measured without defining the tasks and steps to completion.
- Metrics should include effectiveness (degree to which a goal is achieved), along with efficiency (work per unit time) and satisfaction (a subjective rating).
- Objectives of the task could be improved to address the needs of the gap. How much overhead (steps in the task that do not add value and are only present due to the way the system was implemented) is in the procedure?

Operator Performance Models. (Usability)

- The SRP assumes the previous task will provide the metrics for this task. If so, the value of the deliverables of this task depend on revising the previous task.
- The task description is cut off in the IRP supplement.
- What kinds of models are they considering validating?
- The participant pool is not mentioned in the IRP. There may be a need to compensate for a small pool usually used in modeling studies
- Compare with representational analysis (Zhang, 1996).
- Usability is associated with functionality. Teasing these apart would be interesting.

SHFE 3.1.2.2.2.1 (SM11): Can crewmember spatiomotor abilities be more accurately predicted and countermeasures and training techniques developed to mitigate spatial disorientation during spaceflight?

- The gap is adequately scoped.

Current Tasks:

Modeling and Mitigating Spatial Disorientation in Low-Gravity Environments. (NSBRI – Sensorimotor Adaptation Team)

- The task adequately fills the gap.

Enhancement of Spatial Orientation Capability of Astronauts on the Lunar Surface. (NSBRI – Sensorimotor Adaptation Team)

- The task adequately fills the gap.

SHFE 3.1.2.3.2: How do we develop computer interface requirements that will ensure the commonality of the interface designs across multiple Constellation vehicles?

- This was considered more of a development gap rather than a research gap. Task-based design patterns, and other program-wide user interface design standards could be developed that would provide all vendors of vehicle interface subsystems with common requirements for the user interface.
- Consider also the value of centralizing the user interface (UI) design across the various systems and vehicles, as is done in some systems of systems programs (e.g., Army Brigade Combat Team Modernization (BCTM) program).

<u>Current Task:</u>
Interface Design and Commonality

- This is not primarily a research problem but a user interface development goal.
- NASA should develop task flow-based design patterns, or goal-oriented design approaches that have been used successfully on other programs such as the Army BCTM program. There are many sources of existing design patterns on the World Wide Web (WWW), and also books such as van Duyne et al., (2003). See also Endsley et al., (2003).
- This task is critically important and involves many factors including setting requirements in the design process, and managing requirements across systems. This is done by enterprise software applications companies today when developing broad, multi-product, software suites, and handled using an engineering project management system.

SHFE 3.2.2.2.2: How can we integrate multiple types of information and prioritize it appropriately to ensure mission success?

- This gap refers to information presentation and data visualization, and overlaps with 3.1.2.2a (p. 10) on the tasks addressing the software portion of that gap.
- The SRP recommends exercising caution in automatic prioritization of information, so as to avoid limiting crew anticipation of upcoming/possible problems and degrading SA. They also need to consider issues of supporting the crew's understanding of information reliability/confidence level if information is fused (i.e., multiple data sources combined automatically).

<u>Current Task:</u>
Information Integration and Presentation

- This deals with appropriately integrating information and presenting it to people.
- This needs to reference the Information Presentation Tasks (3.1.2.2.a, p. 10).

- Deliverables should include: information presentation architecture (all hierarchical levels), and design patterns for information integration and visualization.
- In order to do this effectively, there should be a definition of applicable crew tasks and data models.
- The task needs to review other successful complex information presentation systems for useful lessons learned.

SHFE 3.2a: How do we design decision support systems to assist crews in obtaining the information they need?

- Literature review should include decision analysis research (Skinner, 1999) and human-automation trade-offs with decision support systems.
- It was not clear to the SRP how much the astronauts use decision support systems. Further justification should be provided.
- Need to investigate potential issues with decision biasing that have been found with many decision support systems (see Layton et al., 1994).

<u>Current Task:</u>
Information Processing with Automated Aids

- This seems like an appropriate task but it does not indicate what kind of support astronauts need.
- Does NASA know what all the tasks will be? This would be important in a decision-making research task.
- The research needs to address how to combine decision parameters (e.g., weighting model) when a decision needs to be made.

SHFE 3.2b: How do we optimize information sharing between humans and automated systems?

- The gap is scoped too broadly.
- The gap should recognize how information sharing between humans and automation changes over time.

<u>Current Task:</u>
Information Sharing between Humans and Automated Systems

- This task is too narrowly focused on natural language. This is not the only means of automation and humans sharing information. In fact, it is one of the least proven methods.
- The task should describe the information content to be addressed, and who needs it.
- Will there be a shared mental model of what humans and automation need? How would that change over time? How can that best be supported through the design of the human-system interaction and function allocation?
- What devices (auditory displays, etc.) would be used by each?

SUGGESTED NEW GAPS & TASKS:

- **What are the tasks that need to be supported with information to reduce errors on exploration missions?**
- **How do we ensure that all the information that we need to accomplish tasks is available in the information architecture (IA)? How does the IA enable users to discover content? E.g., what is the user's mental model, and how does the system address it?**
- **What information does Mission Control provide now that should be on board any new vehicle going to the moon or Mars?**
- **How do we predict the impact of the user interface and automation on operator and team situation awareness (SA)?**
  - **Task: Develop models of SA for individual operator roles.**
  - **Task: Integrate individual role models to model group SA.**
- **How should multimodal information sources be used in HCI interfaces?**
  - **How can multiple modalities be applied to improve performance in timesharing, multitasking, and emergency situations?**

V. Discussion on the strengths and weaknesses of the IRP

Strengths:

- The SRP commends NASA for opening up the research to review by outside sources.
- The structured approach NASA has taken will hopefully produce good visibility into the research goals and deliverables for all the stakeholders.

Weaknesses:

- Overall, there were physically too many units of information spread across non-integrated documents and presentations to read, understand, evaluate, cross-reference, and integrate. There were many unrelated research elements, which in and of themselves are interesting and important, but if they do not produce complementary results which fit together, the gaps will not be closed.
- There were too many documents that did not always state the same information for individual tasks. The presentations were also different from the written documents.
- Some of the documents were poorly edited with cut-off sentences and information left out.
- More details on tasks would have been helpful. Without this information it made it hard to determine if the tasks adequately addressed the gaps.
- The way the tasks were written did not show how they addressed the gaps.
- There should be guidelines developed on how to write the task description in the IRP, e.g., tasks at least as detailed as a journal abstract, schedules, dependencies on other tasks, deliverables, etc., so that they are the same format and each on a single page. Bullet the objectives.
- Across gaps, it was not clear how or whether these projects will share information.
- It was not possible to evaluate schedules across this large number of gaps and tasks manually.

VI. Discussion of element specific questions in addendum and/or any other issues or concerns the panel chooses to address.

1. Are there obviously unrealistic aspects in the IRP schedule?
   - This task is unrealistic: **SHFE 1.1.2.1.1:** Automation Interface Design Development. (AIDT) (SHFE – Directed) (not achievable to automate UI design)
   - **SHFE 2.1a**: **Effects of Architecture on Safety and Performance** (scope too broad)
   - **SHFE 3.2b . How to optimize information sharing between humans an automated systems** (too broad)
   - **SHFE 3.1.2.2.a How do we ensure that the display and control designs** (Gap is too broad and needs to be subdivided. Gap states to design displays and controls to *improve* performance, but not clear if performance standards have been set.)

2. Does the portfolio of gaps and tasks allow us to adequately understand the risk and its likelihood and consequence?
   - AS A WHOLE, NO:
     - Most of the projects were so ill-defined that this is hard to judge.
     - There was not enough depth and breadth in the tasks to cover all the aspects of the gaps adequately.
     - Task analysis was not provided, so it was not clear which crew activities the research was meant to cover.
     - Many of the projects cover a long time period during which technology will advance. It is unclear how changes in technology will be factored into vehicle and user interfaces.
   - Exceptions: the vibration and acoustics studies were solid and are likely to produce valuable results.
   - SPIFE was a good task but it addressed a small part of a large and difficult gap.

3. Is the portfolio of tasks developing appropriate deliverables, such as preventative measures, standards and requirements, tools and technologies?
   - Deliverables were in many cases clearer than the task descriptions.
   - Some of these tasks and models are good, but it is hard to determine if they will do the job required to fill the gaps.
   - There appears to be very little technology being developed.
   - Additional deliverables should be task analyses and design patterns for common user interfaces.

4. Is the portfolio well balanced between risk description, development of standards, requirements, or recommendations and technology development activities?
   - The portfolio is light on technology development, and heavier on standards and guidelines.

- It would be beneficial for researchers and the SRP to clarify what technical readiness level (TRL) they are at currently and where they would like to be by the end of the task.

5. Is the approach to sequencing gaps and tasks appropriate? Are activities that depend on some task or gap being resolved scheduled to begin before the necessary products are available?
   - We were not able to evaluate dependencies in schedule very effectively, as there were too many gaps and tasks. This would require project management software, which would also help eliminate redundant gaps.

6. Other concerns:
   - Risks-Gaps-Tasks paradigm does not seem to be an optimal way to achieve applicable research. It is not top down regarding what goals the missions are trying to achieve.
   - There is a huge gap omitted, in that no one seems to be doing any kind of overall, high-level task analysis or task taxonomy. Without knowledge of the tasks or functions to be done, much of this research will yield insufficient knowledge. User interface design and human factors engineering cannot be effectively conducted in an activity performance vacuum. Feary and Billman at NASA ARC are trying to address this with research on how to do needs analysis.
   - There was no mention of dependencies between research tasks.
   - Big projects should be divided into multiple tasks so they do not continue for years before producing useful deliverables.

VII. REFERENCES

Aha, D.W., McSherry, D., & Yang, Q. (2005). Advances in conversational case-based reasoning. *Knowledge Engineering Review*, *20*(3), 247-254.

Butler, K.A. & Zhang, J. (2009) Design models for interactive problem-solving: context & ontology, representation & routines. April 2009 . In: *CHI '09: Proceedings of the 27th international conference extended abstracts on Human factors in computing systems* http://portal.acm.org/citation.cfm?id=1520340.1520659&coll=GUIDE&dl=GUIDE&CFID=61924578&CFTOKEN=57696211

Butler, K.A.; Zhang, J.; Esposito, C.E.; Bahrami, A.; Hebron, R.; & Kieras, D. (2007). Work-centered design: a case study of a mixed-initiative scheduler In: *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems.*

Clarke, E.M.; Emerson, E.A. & Sifakis, J. (2009). Model Checking: Algorithmic Verification and Debugging. Lecture from the winners of the 2007 ACM Turing Award. *Communications of the ACM*, *52*(11), 75-84.

Endsley, M. R. (1993). A survey of situation awareness requirements in air-to-air combat fighters. *International Journal of Aviation Psychology*, *3*(2), 157-168.

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors, 37*(1), 65-84.

Endsley, M. R., Bolte, B., & Jones, D. G. (2003). *Designing for situation awareness: An approach to human-centered design.* London: Taylor & Francis.

Esposito, C.E.; Paley, W.B.; & Orn, J.C. (1995) Of mice and monkeys: a specialized input device for virtual body animation. Symposium on Interactive 3D Graphics archive. *Proceedings of the 1995 symposium on Interactive 3D graphics* Monterey, California. Table of contents http://portal.acm.org/citation.cfm?id=199424

Kieras, D. E. (1998). A Guide to GOMS model usability evaluation using NGOMSL. In M. Helander, T. Landauer, and P. Prabhu, (Eds) *Handbook of Human-Computer Interaction*,. 2nd ed. North-Holland, Amsterdam.

Klein, G. A. (1993). A recognition primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* , 138-147.. Norwood, NJ: Ablex.

Layton, C., Smith, P. J., & McCoy, C. E. (1994). Design of a cooperative problem-solving system for en-route flight planning: An empirical evaluation. *Human Factors*, *36*(1), 94-119.

Layton, C., Smith, P. J., & McCoy, C. E. (1994). Design of a cooperative problem-solving system for en-route flight planning: An empirical evaluation. *Human Factors*, *36*(1), 94-119.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46,* 50-80. http://web.ebscohost.com/ehost/detail?vid=1&hid=4&sid=0d47694c-97dc-43a8-b402-d0744134d4e0%40sessionmgr12&bdata=JnNpdGU9ZWhvc3QtbGl2ZQ%3d%3d#db=psyh&AN=2005-03160-004

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39,* 230-253. http://web.ebscohost.com/ehost/detail?vid=1&hid=4&sid=af90dd8c-21ba-4050-8430-4372b4c62cc1%40sessionmgr14&bdata=JnNpdGU9ZWhvc3QtbGl2ZQ%3d%3d#db=psyh&AN=1997-05566-005

Pritchett, A. R., & Hansman, R. J. (1997). Pilot non-conformance to alerting system commands. *Proceedings of the Ninth International Symposium on Aviation Psychology*. Ohio State University, Columbus, OH, 274-279

Rosse, C. & Mejino, J.L. (2003) A reference ontology for biomedical informatics: The Foundation Model of Anatomy. *J Biomedical Informatics*, *36*, 478-500.

Skinner, D.C. (1999) *Introduction to Decision Analysis, 2nd Edition*. Probabalistic Publishing, Gainesville, Florida.

Van Duyne, D.K., Landay, J.A., & Hong, J.I. (2003) *The design of sites*. Boston: Addison-Wesley..

Zhang J. (1996). A representational analysis of relational information displays. *International Journal of Human-Computer Studies, 45*, 59-74.

Zhang, J. (1997). Distributed representation as a principle for the analysis of cockpit information displays. *International Journal of Aviation Psychology, 7*(2), 105-121. http://acad88.sahs.uth.tmc.edu/research/publications/cockpit.pdf

Zhang, J. & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science, 18*, 87-122.

VIII. Space Human Factors Engineering (SHFE) SRP Charge

The SRP is chartered by the Human Research Program (HRP) Program Scientist at the NASA Johnson Space Center (JSC). The purpose of the SRP is to review and provide analysis on the status and progress of HRP Elements and Projects. Your report will be provided to the HRP Program Scientist and will also be given as a courtesy to the SHFH Element and SHFE Project at JSC.

The SRP should (to the fullest extent practicable):

1. Evaluate the ability of the Integrated Research Plan (IRP) to satisfactorily address the risks by answering the following questions:
   A. Have the proper Gaps have been identified to address the Risks?
      i) Are all the Gaps relevant?
      ii) Are any Gaps missing?
   B. Have the proper Tasks have been identified to fill the Gaps?
      i) Are the Tasks relevant?
      ii) Are any Tasks missing?

2. Identify the strengths and weaknesses of the IRP, *and* identify remedies for the weaknesses, including answering these questions:
   A. Are the risks addressed in a comprehensive manner?
   B. Are there obvious areas of potential integration across disciplines that are not addressed?

3. Address (as fully as possible) the questions provided in the charge addendum and to comment on any additional information provided to the Panel that is not addressed in #1 or 2 above.

4. Expect to receive review materials at least five weeks prior to the site visit.

5. Participate in a SRP teleconference to discuss any issues, concerns, and expectations of the review process approximately three weeks prior to the face-to-face meeting
   A. Discuss the SRP charge and address questions about the SRP process
   B. Identify any issues the SRP would like to have answered prior to the site visit

6. Attend the SRP meeting (and possible tour) at NASA/JSC
   A. Attend Element and Project presentations, question and answer session, and briefing
   B. Prepare a draft report including recommendations from the SRP that will be briefed to the Program Scientist by the SRP chairperson or panel. The report should address #1 and 2 above, the questions in the charge addendum, and any other information considered relevant by the SRP.

7. Prepare a final report (within one month of the site visit) that contains a detailed evaluation of the risks and provides specific recommendations that will optimize the scientific return to the HRP. The final report should provide a comprehensive review of Item #1 and 2 above,

address the questions in the addendum to the charge, and any additional information the SRP would like to provide.

8. Consider the possibility of serving on a non-advocate review panel of a directed research proposal or on a solicited research peer review panel; or otherwise advise the Program Scientist.

Addendum to charge (Element Specific Concerns):

1. Are there obviously unrealistic aspects in the IRP schedule?
2. Does the portfolio of gaps and tasks allow us to adequately understand the risk and its likelihood and consequence?
3. Is the portfolio of tasks developing appropriate deliverables, such as preventative measures, standards and requirements, tools and technologies?
4. Is the portfolio well balanced between risk description, development of standards, requirements, or recommendations and technology development activities?
5. Is the approach to sequencing gaps and tasks appropriate? Are activities that depend on some task or gap being resolved scheduled to begin before the necessary products are available?

IX. Space Human Factors Engineering (SHFE) SRP Roster

Panel Chair:

Anna Wichansky
Applications User Experience
Oracle
500 Oracle Parkway-MS20p2
Redwood Shores, CA 94065
Ph: 650-506-8007
Email: anna.wichansky@oracle.com

Panel Members:

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

Mary Cummings
Humans and Automation Laboratory
Massachusetts Institute of Technology
77 Massachusetts Ave, 33-311
Cambridge, MA 02139
Ph: 617-252-1512
Email: missyc@MIT.EDU

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

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