Fiscal Year:	FY 2011	**Task Last Updated:**	FY 10/25/2010
PI Name:	Sandor, Aniko Ph.D.		
Project Title:	Usability evaluation		
Division Name:	Human Research		
Program/Discipline:	HUMAN RESEARCH		
Program/Discipline--Element/Subdiscipline:	HUMAN RESEARCH--Space Human Factors Engineering		
Joint Agency Name:		**TechPort:**	No
Human Research Program Elements:	(1) **SHFH**:Space Human Factors & Habitability (archival in 2017)		
Human Research Program Risks:	(1) **HSIA**:Risk of Adverse Outcomes Due to Inadequate Human Systems Integration Architecture		
Space Biology Element:	None		
Space Biology Cross-Element Discipline:	None		
Space Biology Special Category:	None		
PI Email:	Aniko.Sandor-1@nasa.gov	**Fax:**	FY
PI Organization Type:	NASA CENTER	**Phone:**	281.483.9726
Organization Name:	Lockheed-Martin/NASA Johnson Space Center		
PI Address 1:	2101 Nasa Parkway		
PI Address 2:	Mail Code: C46		
PI Web Page:			
City:	Houston	**State:**	TX
Zip Code:	77058	**Congressional District:**	22
Comments:			
Project Type:	Ground	**Solicitation / Funding Source:**	Directed Research
Start Date:	10/01/2008	**End Date:**	10/01/2010
No. of Post Docs:		**No. of PhD Degrees:**	
No. of PhD Candidates:		**No. of Master' Degrees:**	
No. of Master's Candidates:		**No. of Bachelor's Degrees:**	
No. of Bachelor's Candidates:		**Monitoring Center:**	NASA JSC
Contact Monitor:	Woolford, Barbara	**Contact Phone:**	218-483-3701
Contact Email:	barbara.j.woolford@nasa.gov		
Flight Program:			
Flight Assignment:	NOTE: Project ended early per E. Connell/JSC; previous end date was 9/30/2011 (10/20/2010)		
Key Personnel Changes/Previous PI:			
COI Name (Institution):	Holden, Kritina (Lockheed-Martin/ NASA Johnson Space Center) Archer, Ronald (Lockheed-Martin/ NASA Johnson Space Center)		
Grant/Contract No.:	Directed Research		
Performance Goal No.:			
Performance Goal Text:			
Task Description:	This proposal addresses the need for research in the area of metrics and methodologies used in hardware and software usability testing in order to define quantifiable and verifiable usability requirements. A usability test is a human-in-the-loop evaluation where a participant works through a realistic set of representative tasks using the hardware/software under investigation. The purpose of this research is to define metrics and methodologies for measuring and verifying usability in the aerospace domain in accordance with FY09 focus on errors, consistency, and mobility/maneuverability. Usability metrics must be predictive of success with the interfaces, must be easy to obtain and/or calculate, and must meet the intent of current Human Systems Integration Requirements (HSIR). Methodologies must work within the constraints of the aerospace domain, be cost and time efficient, and be able to be applied without extensive specialized training.		

The key driver for this directed research project (DRP) is the desire to promote and facilitate the development of usable Constellation vehicles and habitats. In past programs, usability has often been an afterthought – with human factors activities coming far too late in the development lifecycle to make a difference. It is the goal of this DRP to provide research-based methodologies and metrics early enough in the Orion program to positively impact development.

Once new methodologies and metrics are developed, they will be field tested in real-world design efforts, iterated based on results, and finally described in reports and guidelines manuals, along with their application to requirements

Rationale for HRP Directed Research:

Research Impact/Earth Benefits:

The Usability Evaluation DRP team provided the research, the proposed methodology, the case study examples and wording for the requirement and verification of both the usability and legibility requirements, as well as coordinated with all stakeholders from the Human-Systems Integration Group (HSIG), crew office, and Prime contractor.
The literature review on usability metrics related to efficiency, effectiveness, and satisfaction resulted in standards that were included in NASA-STD-3001. The process for collecting and analyzing error rates developed as part of this DRP provided the basis for the Usability Process in CCT-1002 Commercial Human-Systems Integration Process document.

A maneuverability scale was developed and tested for use in evaluating maneuverability in space suits and unsuited in confined spaces such as crew quarters. The Maneuverability Assessment Scale (MAS) is a 5-point scale ranging from 1 – Excellent to 5 – Very Poor that measures the ability to move in any direction with the desired pace and accuracy.

Legibility:
Legibility is defined as the ability of an observer to discriminate visual stimulus details to such a degree that it can be recognized. Legibility refers to the perceptual clarity of visual objects. It is influenced by the method of display generation, application of human factors guidelines for correct depiction of the object in relation to the task requirements, the environmental conditions, and eyesight standards. Legibility of text is often defined in terms of readability. Legibility of alphanumeric information, symbols, and icons on interfaces is a major part of system usability. In general, there are guidelines and standards that need to be followed to insure good legibility in all environmental conditions in which information need to be read off the interfaces. In FY09 a literature review was conducted on legibility methodologies for software labels in order to find a method that can be proposed for the verification statement of the legibility requirement in the Human Systems Integration Requirements (HSIR) along with a criterion for successful verification. In FY10 we tested the proposed software legibility methodology. A study was conducted to evaluate the methodology on an Orion display with Monotype, Monotype Italics, Verdana and Verdana Italics using the 0.17” font size and 25” viewing distance that is used by Orion. The methodology used was based on rapid serial visual presentation and verbal identification by subjects of the labels tested. The study showed that the 98% accuracy required in the HSIR Rev E (NASA, in review)and in ISO 9241-11 (1998) is attainable: all 5 subjects in the study reached an accuracy of 99.6 and higher. Furthermore, a literature review was conducted to find and recommend a methodology for hardware labels as well. The results of this line of research within the Usability Evaluation DRP provided the methodology, wording, and criterion for the current HSIR Rev E legibility requirement and verification.

Consistency:

Onboard space vehicles astronauts work with a large variety of hardware and software that are designed and built by various groups within NASA or external to NASA. The outcome of having multiple developer groups is sometimes a serious lack of consistency among the user interfaces, resulting in increased training requirements, errors, and frustration for crewmembers. Thus, a special area of concern within the NASA human factors community is consistency of design. Consistent design is commonly listed as a usability guideline, but it has been proven difficult to measure and quantify it. Consistency is an important factor of usability of user interfaces: consistent interfaces can reduce time spent on training and can improve task completion times. In spite of its importance, there is no standard method or evaluation tool to measure consistency. As part of the Usability Evaluation DRP, in FY09 a general system consistency scale has been developed and evaluated on a website. The System Consistency Scale is composed of 3-point rating scales (1 being very inconsistent and 3 being very consistent) for interface elements in the areas of text, navigation, icons, symbols, hardware, and virtual elements. In FY10 the general System Consistency Scale was adapted to a case study: Orion display formats and needed only minor modifications. The customized display format consistency scale was evaluated on the Orion display formats to see how well the scale works. Inter-rater reliability was also evaluation for the scale.

Maneuverability:

To properly design the hardware to be used by the crew, current human factors evaluations collect various types of objective and subjective data to determine the usability of the hardware. Objective data (i.e., Range of Motion, Torque) have been used to quantify the mobility of space suits; however, there is also a need to collect subjective ratings on the mobility/maneuverability of hardware while completing a specific task. Subjective data can provide a different point of view on maneuverability as noticed from comments during evaluations. However, none of the subjective scales used during these evaluations provide a clear subjective measurement of the ease of movement while conducting the tasks. In FY09 a maneuverability scale was developed that can be used to evaluate maneuverability in space suits and confined spaces such as crew quarters. The definition used for maneuverability was “the ability to move in the direction and at the desired pace required to complete the task.” Although this definition proved to be appropriate based on previous evaluations, it is possible that maneuverability is affected by factors other than direction and desired pace and successful task completion. Therefore, in FY10 the purpose of the Usability Evaluation DRP was to refine the definition for maneuverability and to evaluate factors affecting maneuverability by looking at factors such as cognitive and physical effort, compensation, and fatigue besides desired direction and pace. The study consisted of participants completing a full body (donning and doffing of a flight suit) task in free space and confined space, as well as a fine motor task gloved and ungloved. The hypothesis of the study was that the conditions for the two tasks lead to differences in maneuverability. The collected metrics looked at all factors that may affect maneuverability. A multiple regression analysis was conducted to look at which factors are good predictors of maneuverability. Based on the results the maneuverability scale was refined. Future plans include conducting a reliability and validity study of the scale.

Efficiency, Effectiveness and Satisfaction:

Efficiency, effectiveness, and satisfaction are the three major components of usability and all three should be measured

Task Progress:	for a system to get a good idea of the usability of the system. Efficiency is defined as the relation between 1) the accuracy and completeness with which users achieve certain goals and 2) the resources expended in achieving them. Effectiveness is the accuracy and completeness with which users achieve certain goals. Satisfaction is the users' comfort with and positive attitudes towards the use of the system. Research has shown that these factors are independent of each other with very low correlations among them (less than 0.15) (Hornbæk & Law, 2007; Sauro & Lewis, 2009). A literature review was conducted on measures of efficiency, effectiveness, and satisfaction that can be adapted to crew interfaces. This line of research from the Usability Evaluation DRP provided wording for the NASA STD 3001 and the Commercial Human Systems Integration Requirements. References Apple Computer, I. (1992). Macintosh Human Interface Guidelines. : Reading, MA: Addison-Wesley Publishing Co. Bias, R. G., & Mayhew, D. J. (2005). Cost-Justifying Usability: An Update for the Internet Age. San Francisco, CA: Morgan Kaufmann. Crandall, B., Klein, G., & Hoffman, R. R. (2006). Working Minds: A Practitioner's Guide to Cognitive Task Analysis.: MIT Press. Hackos, J. T., & Redish, J. C. (1998). User and Task Analysis for Interface Design: John Wiley & Sons, Inc. Hornbæk, K., & Law, E. L.-C. (2007). Meta-analysis of correlations among usability measures. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, April 28-May 03, 2007, San Jose, CA, USA ISO-9126-2. (2003). ISO/IEC TR 9126-2 Software engineering product quality, Part 2: External metrics. Geneva, Switzerland: International Organization for Standardization. ISO-9241-11. (1998). ISO/IEC 9241-11:1998. Ergonomic requirements for office work with visual display terminals (VDTs), Part 11: Guidance on usability. ISO-9241-304. (2008). ISO 9241-304 Ergonomics of human-system interaction Part 304: User performance test methods for electronic visual displays. Geneva, Switzerland: International Standards Organization. Kirwan, B., & Ainsworth, L. K. (1992). A Guide To Task Analysis: The Task Analysis Working Group. USA: CRC Press. Kuniavsky, M. (2003). Observing the user experience. San Francisco, CA: Morgan Kaufmann Publishers Inc. Microsoft. (1995). The Windows Interface Guidelines for Software Design. : Redmond, WA: Microsoft Press. Microsoft. (1999). The Microsoft Windows User Experience: Redmond, WA: Microsoft Press. MIL-STD-1472D. (1989). Military Standard: Human Engineering Design Criteria for Military Systems, Equipment and Facilities MIL-STD-1472D. NASA. (2008). Human Systems Integration Requirements Revision C (HSIR Rev. C.), NASA-CXP70024. Houston, Texas: Lyndon B. Johnson Space Center. NASA. (in review). Human Systems Integration Requirements Revision E (HSIR Rev. E.). Houston, Texas: Lyndon B. Johnson Space Center. Nielsen, J. (1993). Usability Engineering. San Francisco, CA: Morgan Kauffman Publishers Inc. Nielsen, J. (1994). Usability Engineering. San Francisco, CA: Morgan Kauffman Publishers Inc. Rosson, M. B., & Carroll, J. M. (2001). Usability engineering: scenario-based development of human-computer interaction San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. Salvendy, G. (1997). Handbook of Human Factors.: John Wiley & Sons, New York. Sauro, J., & Lewis, J. R. (2009). Correlations among Prototypical Usability Metrics: Evidence for the Construct of Usability. Paper presented at the Computer Human Interaction (CHI), Boston, MA. Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. Psychological Review, 84(1), 1-66. Seffah, A., Donyaee, M., Kline, R. B., & Padda, H. K. (2006). Usability measurement and metrics: A consolidated model. Software Quality Journal, 14, 159-178. Tullis, T., & Albert, B. (2008). Measuring the user experience: Collecting, analyzing, and presenting usability metrics. Burlington, MA: Morgan Kaufmann Publishers Inc.
Bibliography Type:	Description: (Last Updated: 03/03/2016)
Abstracts for Journals and Proceedings	Sandor A, Holden KL. "Legislating Usability: A Brief History of the Usability Requirement at NASA." Presented at the Houston Human Factors and Ergonomic Society Meeting, Houston, TX, May 21, 2010. Houston Human Factors and Ergonomic Society Meeting, May 21, 2010. http://www.houstonhfes.org/conferences/conference2010/program.html , May-2010