Fiscal Year:	FY 2014	**Task Last Updated:**	FY 08/21/2014
PI Name:	Duda, Kevin R Ph.D.		
Project Title:	Metrics and Methods for Real-Time Task Performance Assessment		
Division Name:	Human Research		
Program/Discipline:	NSBRI		
Program/Discipline--Element/Subdiscipline:	NSBRI--Human Factors and Performance Team		
Joint Agency Name:		**TechPort:**	No
Human Research Program Elements:	(1) **SHFH**:Space Human Factors & Habitability (archival in 2017)		
Human Research Program Risks:	(1) **BMed**:Risk of Adverse Cognitive or Behavioral Conditions and Psychiatric Disorders (2) **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:	kduda@draper.com	**Fax:**	FY 617-258-2772
PI Organization Type:	NON-PROFIT	**Phone:**	617-258-4385
Organization Name:	The Charles Stark Draper Laboratory, Inc.		
PI Address 1:	555 Technology Sq		
PI Address 2:	MS 27		
PI Web Page:			
City:	Cambridge	**State:**	MA
Zip Code:	02139-3539	**Congressional District:**	7
Comments:			
Project Type:	GROUND	**Solicitation / Funding Source:**	2012 Crew Health NNJ12ZSA002N
Start Date:	07/01/2013	**End Date:**	06/30/2016
No. of Post Docs:	0	**No. of PhD Degrees:**	0
No. of PhD Candidates:	0	**No. of Master' Degrees:**	1
No. of Master's Candidates:	1	**No. of Bachelor's Degrees:**	0
No. of Bachelor's Candidates:	0	**Monitoring Center:**	NSBRI
Contact Monitor:		**Contact Phone:**	
Contact Email:			
Flight Program:			
Flight Assignment:			
Key Personnel Changes/Previous PI:			
COI Name (Institution):	Robinson, Stephen (University of California, Davis)		
Grant/Contract No.:	NCC 9-58-HFP03401		
Performance Goal No.:			
Performance Goal Text:			

1. Original Project Aims/Objective: The project objective is to produce a configurable and portable simulation capability for developing and validating real-time metrics for assessing flight performance, workload, and situational awareness. There are three integrated specific aims: (1) Define the system architecture for integrating vehicle and environmental models with the simulation environment. (2) Perform a critical analysis of four piloted tasks: MPCV/Orion docking, MPCV/Orion entry, Lunar Landing, and EVA SAFER self-rescue. Simulator data will be analyzed to identify candidate metrics for performance, workload, and situational awareness as well as operationally relevant options for presenting feedback to the operator. (3) Conduct a series of experiments using the simulated spaceflight tasks and real-time metrics engine to baseline performance, workload, and situational awareness in each task in order to develop algorithms and methods for alerting the operator to deviations from nominal.
2. Key Findings: In project year 1, we completed the development of the system architecture for integrating the vehicle

Task Description:	models with our simulation framework, calculating objective workload metrics from flight performance data, performing speech recognition for situation awareness estimation, and rendering an out-the-window view using NASA's Engineering Dynamic On-board Ubiquitous Graphics (DOUG) Graphics for Exploration (EDGE) package. We developed a portable ground station, which includes a single PC, two joysticks (a rotational hand controller (RHC) and a translational hand controller (THC)), five monitors, speakers, and a microphone. Three prior lunar landing simulations – all developed under prior NSBRI projects (Project SA01604 Fuel Contour Display, HFP02001 Mode Transition and Failure Detection models) – and the SAFER EVA self-rescue scenario were integrated in the simulation. These simulations will provide the basis for the planned year 2 experimentation for developing and validating the implemented real-time performance metrics. Three real-time objective workload metrics were prototyped for implementation with the Mode Transition simulation (Hainley et al., 2013). These included the baseline implementation of a two-chose secondary task response time, an analysis of the entropy of the RHC inputs, and the root mean square error (RMSE) of the difference between the actual and guidance recommended attitude during piloted flight. The RHC entropy and attitude RMSE metrics were chosen based on post-hoc analysis of prior experiment data. In addition, we implemented a real-time speech recognition engine to analyze the accuracy of required key system state callouts as a measure of situation awareness (see Hainley et al., 2013). In this implementation, the utterances spoken by the user are recognized by the speech engine and then compared against the actual simulated vehicle state and are scored correct based on temporal (e.g., did they make the callout within 2 seconds?) and spatial (e.g., did they make the callout within 10 feet of the actual altitude?) boundaries. It is this feedback – both from the secondary task response time, mental workload, and situation awareness metrics – that we plan to further develop in the out years of the project to refine the presentation to the pilot for making operations more safe and efficient. 3. Impact of Key Findings on hypotheses, technology requirements, objectives, and specific aims of the original proposal: The development of the integrated simulation platform for running the vehicle models, recording/logging data, unobtrusively estimating workload and situation awareness, and providing visualizations and feedback to the pilot has significantly enhanced the capabilities for developing real-time performance metrics. By using typical spacecraft command and control tasks, such as piloted lunar landing, we have an initial operational scenario to test our metrics. The HRP Integrated Research Plan gap (SHFE-TASK-01) states, in part, that, …The successful management or evaluation of workload must include a consideration of the nature of individual tasks that operators must perform, the combinations of tasks that are performed during a work period, priorities among tasks, and individual differences among operators. The development and evaluation of real-time performance metrics in representative operational settings—which include task performance, workload, and situational awareness, and are measured objectively as well as subjectively—will provide valuable data for the validity assessment. 4. Proposed research plan for the coming year: In project year 2, we aim to begin conducting human-in-the-loop experiments with our lunar landing Mode Transition experiment to quantitatively compare the real-time performance metrics calculations with our baseline post-hoc measures of flight technical error, workload, and situation awareness. The real-time metrics that are analyzed and collected on-the-fly will be quantitatively compared against the prior approach to analyze the data after-the-fact. This will be done in close collaboration with our team mates at the University of California, Davis. In addition, an operationally relevant approach for fusing the vehicle flight data, workload, and situation awareness data will be developed, as well as an approach for providing feedback will be prototyped and evaluated. In project year 2, we also aim to develop and integrate Orion/MPCV rendezvous and docking and atmospheric entry models with our simulation. Real-time performance metrics will also be developed and integrated with this simulation, with the goal of having consistent metrics between classes of simulations. Lastly, we propose to deliver and install a copy of our simulation station at NSBRI Headquarters to provide a foundational capability for subsequent test and evaluation with operators and subject matter experts.
Rationale for HRP Directed Research:	
Research Impact/Earth Benefits:	This proposed project aims to deliver a research capability for evaluating the applicability and robustness of metrics for quantifying operator performance in real-time. Although our case studies are specific to piloted spacecraft, the innovations and implementation approach are generally applicable to any vehicle that requires a human in the loop. This re-configurable, portable simulation and test station provides a capability for integrating and testing real-time performance metrics for assessing operator effectiveness continually throughout a trial, as opposed to a single mission effectiveness metric. In addition, temporal operator performance can then be assessed against system-level metrics such as fuel consumption vs. time. Regardless of the domain, the interaction between vehicle/operation performance, operator workload, and operator situation awareness is complicated. Prior approaches to quantify these metrics have relied on post-hoc analyses or measurement approaches that affect the parameter of interest. This project aims to reduce to practice in-situ real-time performance, workload, and situation awareness metrics that can be objectively and unobtrusively collected. We are doing this through a flexible and module architecture that allows researchers to develop their own modules (either vehicle/system models or metrics modules) that can be integrated with our simulation framework. Through rigorous testing and integration with operationally-relevant tasks and scenarios, our goal is that this platform be adopted by the human-system integration and research community as the gold standard in crew performance benchmarking through open-source integration of algorithms for metrics development and validation.
	In project year 1, we completed the development of the system architecture for integrating the vehicle models with our simulation framework, calculating objective workload metrics from flight performance data, performing speech recognition for situation awareness estimation, and rendering an out-the-window view using NASA's Engineering Dynamic On-board Ubiquitous Graphics (DOUG) Graphics for Exploration (EDGE) package. We developed a portable ground station, which includes a single PC, two joysticks (a rotational hand controller (RHC) and a translational hand controller (THC)), five monitors, speakers, and a microphone. (This ground control station leverages development that was performed for NASA's Game Changing Development Program.) Three prior lunar landing simulations – all developed under prior NSBRI projects (Project SA01604 Fuel Contour Display, HFP02001 Mode Transition and Failure Detection models) – as well as the SAFER EVA rescue scenario were integrated in the simulation, which included the vehicle/system dynamics models, flight displays, landing/return metrics, and data logging capability. Three real-time objective workload metrics were prototyped for implementation with the lunar landing Mode Transition simulation (Hainley et al., 2013). These included the baseline implementation of a two-chose secondary task response time, an analysis of the entropy of the RHC inputs, and the root mean square error (RMSE) of the difference between the actual and guidance recommended attitude during piloted flight. In addition, we implemented a real-time speech

Task Progress:	recognition engine with our simulation to analyze the accuracy of required key system state callouts as a measure of situation awareness (see Hainley et al., 2013). In this implementation, the utterances spoken by the user are recognized by the speech engine and then compared against the actual simulated vehicle state and are scored correct based on temporal and spatial boundaries. It is this feedback – both from the secondary task response time, mental workload, and situation awareness metrics – that we plan to further develop in the out years of the project to refine the presentation to the pilot for making operations more safe and efficient. Planning for the start of human subject experimentation in year 2 was initiated. This included the submission and approval of an experimental protocol to the University of California, Davis IRB, the shipment of a copy of the simulation control station with integrated lunar landing simulations, and the identification of research objectives and testable hypotheses for the investigation and validation of in-situ real-time flight performance, workload, and situation awareness metrics. In addition, the framework for fusing these metrics for providing feedback to the pilot and/or control system to make operations more safe and efficient was also completed. REFERENCE (reported in NSBRI "Human-Automation Interactions and Performance Analysis of Lunar Lander Supervisory Control" project): Hainley CJ Jr, Duda KR, Oman CM, Natapoff A. "Pilot performance, workload, and situation awareness during lunar landing mode transitions." Journal of Spacecraft and Rockets. 2013 Jul;50(4):793-801. http://dx.doi.org/
Bibliography Type:	Description: (Last Updated: 09/04/2023)
Abstracts for Journals and Proceedings	Duda KR, Robinson SK, Handley P, Tinch JD, West JJ. "Metrics and Methods for Real-Time Task Performance Assessment." 2014 NASA Human Research Program Investigators' Workshop, Galveston, TX, February 12-13, 2014. 2014 NASA Human Research Program Investigators' Workshop, Galveston, TX, February 12-13, 2014. http://www.hou.usra.edu/meetings/hrp2014/pdf/3053.pdf , Feb-2014