Fiscal Year:	FY 2018	**Task Last Updated:**	FY 08/20/2018
PI Name:	Dinges, David F. Ph.D.		
Project Title:	Standardized Behavioral Measures for Detecting Behavioral Health Risks during Exploration Missions		
Division Name:	Human Research		
Program/Discipline:			
Program/Discipline--Element/Subdiscipline:	HUMAN RESEARCH--Behavior and performance		
Joint Agency Name:		**TechPort:**	No
Human Research Program Elements:	(1) **HFBP**:Human Factors & Behavioral Performance (IRP Rev H)		
Human Research Program Risks:	(1) **BMed**:Risk of Adverse Cognitive or Behavioral Conditions and Psychiatric Disorders (2) **Team**:Risk of Performance and Behavioral Health Decrements Due to Inadequate Cooperation, Coordination, Communication, and Psychosocial Adaptation within a Team		
Space Biology Element:	None		
Space Biology Cross-Element Discipline:	None		
Space Biology Special Category:	None		
PI Email:	dinges@pennmedicine.upenn.edu	**Fax:**	FY
PI Organization Type:	UNIVERSITY	**Phone:**	215-898-9949
Organization Name:	University of Pennsylvania		
PI Address 1:	Department of Psychiatry		
PI Address 2:	423 Service Dr., 1013 Blockley Hall		
PI Web Page:			
City:	Philadelphia	**State:**	PA
Zip Code:	19104-4209	**Congressional District:**	2
Comments:			
Project Type:	Flight,Ground	**Solicitation / Funding Source:**	2013-14 HERO NNJ13ZSA002N-BMED Behavioral Health & Performance
Start Date:	07/21/2015	**End Date:**	07/20/2019
No. of Post Docs:	0	**No. of PhD Degrees:**	0
No. of PhD Candidates:	1	**No. of Master' Degrees:**	0
No. of Master's Candidates:	0	**No. of Bachelor's Degrees:**	0
No. of Bachelor's Candidates:	0	**Monitoring Center:**	NASA JSC
Contact Monitor:	Williams, Thomas	**Contact Phone:**	281-483-8773
Contact Email:	thomas.j.will1@nasa.gov		
Flight Program:	ISS		
Flight Assignment:	NOTE: End date changed to 7/20/2019 per NSSC information (Ed., 8/10/18) NOTE: Element change to Human Factors & Behavioral Performance; previously Behavioral Health & Performance (Ed., 1/18/17)		
Key Personnel Changes/Previous PI:	August 2018: Sarah McGuire, Co-Investigator, left the University of Pennsylvania faculty to accept another position, and is no longer on the project.		

COI Name (Institution):	Basner, Mathias M.D. (University of Pennsylvania) Mollicone, Daniel Ph.D. (Pulsar Informatics, Inc.) Stuster, Jack Ph.D. (Anacapa Sciences, Inc.) Strangman, Gary Ph.D. (Harvard Medical School) Stahn, Alexander Ph.D. (University of Pennsylvania) Roma, Peter Ph.D. (Johns Hopkins University School of Medicine) Gur, Ruben Ph.D. (University of Pennsylvania)
Grant/Contract No.:	NNX15AK76A
Performance Goal No.:	
Performance Goal Text:	
Task Description:	Isolated and confined environments anticipated during exploration missions will include stressors such as small teams living and working in extreme conditions for prolonged periods separated from family, friends; loss of the day/light cycle; loss or delay of communications with ground; partial gravity; and limited space, privacy, and food selection. NASA's Human Factors and Behavioral Performance Element seeks to maintain and enhance behavioral health and performance in such environments. The behavioral risk (Risk of Adverse Cognitive or Behavioral Conditions and Psychiatric Disorders) is a high priority within the NASA Human Research Program (HRP) because it has face validity, but lacks sufficient evidence due to a deficiency in measurement of the risk. Thus, there is concern that the behavioral health of the crew will be challenged in a Mars mission; however, there is no standardized method to detect and quantify the magnitude of the risk or its likelihood. The overarching goal of this project is to build on a successful record of software-based measurement of behavioral health indicators (e.g., mood, cognitive function, performance, physical and mental fatigue, sleep quality) to develop a complementary standardized suite of behavioral core measures (BCM) that would be quite feasible to implement within the constraints of spaceflight research, ground-based analogs (both short- and long-duration), and prolonged missions in isolated, confined, extreme environments lasting up to 12 months or longer. Achievement of this goal would permit a more rapid and reliable assessment and quantification of the Risk of Adverse Cognitive or Behavioral Conditions Psychiatric Outcomes for exploration class missions. The standardized behavioral medicine measures we are developing will be similar to the guidelines for standardization for bed rest studies. Defining standardized measures will not only allow for the systematic collection of data across multiple analogs, but it will also facilitate risk characterization for the Behavioral Medicine (BMed) risk. Without a standardized suite of behavioral health measures, the unknown BMed risk for exploration-class missions will continue to be estimated based on anecdote and conjecture. The elements of the BCM include the following: (1) The Cognition test battery (a suite of 10 brief neuropsychological tests specifically designed for astronauts), (2) actigraphy sleep/wake data, (3) several visual analog scales and brief questionnaires with proven validity and utility in space and space analog environments, (4) Journals (Audio/Video and/or typed), (5) a robotic arm track-and-capture grappling task, (6) Team Measure Questionnaires designed to measure aspects of crew performance, team processes, team climate, and group living. Data acquisition feasibility and flexibility, and user acceptability of BCM has been assessed in a short duration analog (4 HERA (Human Exploration Research Analog) N=16 subjects), is currently being assessed in a long-duration analog (10 months of data collection during a 12-14 month winter-over in the Antarctic Neumayer station, N=7 subjects), and will be further assessed on the International Space Station (ISS) for feasibility and acceptability (during a 6 month mission, N=2 astronauts). Mission controllers completed a team performance questionnaire during HERA missions, and flight directors will complete this brief questionnaire during the ISS study.
Rationale for HRP Directed Research:	
Research Impact/Earth Benefits:	This project will deliver a Behavioral Core Measures Tool (BCM) that will be tested for its feasibility, flexibility, and acceptability in research studies in both short and long duration space analog environments and on the ISS. With the BCM, it will be possible for NASA's HFBP (Human Factors & Behavioral Performance) program to much better assess and quantify the Risk of Adverse Cognitive or Behavioral Conditions and Psychiatric Outcomes for exploration class missions. With the proposed work we will relevantly contribute to HRP's goal to provide human health and performance countermeasures, knowledge, technologies, and tools to enable safe, reliable, and productive human space exploration. More specifically, the BCM will constitute an important technology to provide mission planners and system developers with strategies for monitoring and mitigating crew health and performance risks. Additionally, Behavioral Core Measures could also be beneficial for monitoring behavioral health during Earth-based operations, especially those involving isolated, confined and extreme environments (e.g., Antarctic research expeditions).
	Cognition: Data acquisition in NASA's Human Exploration Research Analog (HERA) facility was finalized on 10/26/2016. A total of N=288 full Cognition test bouts (100% of expected) were successfully collected as well as N=270 surveys (93.75% of expected) in N=16 crewmembers. During the current reporting period we have deployed and collected data on N=8 crewmembers during two out of the four Campaign 3 missions in NASA's Human Exploration Research Analog (HERA) facility. In HERA research participants perform the Cognition test battery on the Apple iPad; however, through discussions with the International Space Station Medical Program (ISSMP) it was determined that the iPad is not a feasible platform for data collection on ISS. Thus a Windows PC version of the Cognition software will be used on-board the ISS and, in the past year, the Windows version of Cognition was deployed in the Antarctic Neumayer station. Data acquisition on Neumayer station was finalized in November 2017. We received 82 full Cognition batteries from N=7 crewmembers in total (Cognition was performed on a monthly basis). We also received N=59 Psychomotor Vigilance Task (PVT) tests (the PVT plus survey only was performed 2 weeks after each full Cognition battery). Post-mission data acquisition at Charité Berlin was performed on 3/22/2018. Self Report and Visual Analog Scale Measures: A set of visual analog scales and brief questionnaires with proven validity and utility in space and space analog environments were chosen in order to evaluate several key aspects of behavioral health and crew interaction. These questionnaires include (a) Visual Analog Scales (VAS) to evaluate

perceived mental (mentally sharp—mentally fatigued) and physical (energetic—physically exhausted) exhaustion, fatigue (tired—fresh, ready to go), sleepiness (not sleepy at all—very sleepy), stress (not stressed at all—very stressed), and workload (very low—very high); (b) the Social Desirability Scale (SDS-17) to measure self-desirability bias; (c) sleep diaries to evaluate sleep quality and duration; (d) the Profile of Mood States, Short Form (POMS-SF) to evaluate mood; (e) the Beck Depression Inventory (BDI-II) to evaluate depression; and (f) the Conflict Scale (CS) to evaluate perceived conflict among crewmembers and between crewmembers and mission control. The Visual Analog Scales and sleep diaries have already proven useful for measuring neurobehavioral health during 6-month missions on ISS and these data informed the need for a one-year mission.

During the 14-month Neumayer missions these surveys were deployed using the University of Pennsylvania's Qualtrics electronic web-based survey tool in order to accommodate offline data collection and storage. Data were successfully collected on N=7 Neumayer crewmembers, and self-report measures were successfully collected with 88.4% response adherence. A total of N=501 BCM questionnaires were completed: n=82 nighttime Visual Analog Scales; n=82 Conflict Scales; n=165 morning Visual Analog Scales and sleep diaries; n=84 POMS-SF, n=81 BDI-II, and n=7 SDS-17 questionnaires.

ISS survey data collection is currently scheduled to begin in 2018 in ISS increment 57/58 using the on-board Data Collection Tool (DCT) software.

Journals: At the end of the reporting period, journal entries from all four of the HERA simulated asteroid rendezvous missions (N=16 crewmembers) had been transcribed and analyzed. Only one crewmember of the Neumayer III Antarctic station completed a confidential journal, which allowed limited analysis to be performed.

Category Analysis: The numbers of parsed journal statements assigned to the 25 major topical categories were tabulated for each subject's journal and then combined to calculate the overall total for each category. The subjects' rankings are remarkably similar to the category ranking derived from the journals of ISS astronauts during the 13-year Journals Flight Experiment.

Subcategory Analysis: Adjustment

Statements extracted from the HERA A/V journals were assembled according to their category assignments and the mission quarter during which they were recorded. Only the primary, secondary, and tertiary assignments for the Adjustment category were subjected to a subcategory analysis in this study. Statements were assigned to subcategories based on similar content. A total of 20 subcategories emerged from the data. Journal statements assigned to the Adjustment category range from references to the positive effects of exercise to the importance of remaining busy with meaningful work.

Net Positivity/Negativity: Each parsed journal statement was assigned a code to indicate whether the statement was positive, negative, or neutral in its tone or content. A metric called Net Positivity/Negativity (NPN) was derived by subtracting the proportion of negative entries from the proportion of positive entries. This metric was calculated for all category assignments by quarter for each expedition. NPN analyses were performed by journal with data from all categories combined and by journal focusing exclusively on statements assigned to the Adjustment category. NPN analyses were conducted to test hypotheses concerning a "third quarter phenomenon," a decline in affect, general attitude, or "morale" during the third quarters of missions, regardless of duration. The third quarter phenomenon was suggested by anecdotal accounts and evidence from previous content analyses of expedition journals.

Conclusion

Results of the Journals Component of the BCM study that was conducted during four HERA simulations in 2016 and the Neumayer Antarctic expedition in 2017 confirm previous experiences with French polar explorers and NASA astronauts that analysis of confidential journals can be used to assess the behavioral health of personnel in remote-duty environments. However, certain conditions must be met for the method to be effective: 1) The personnel must make at least weekly journal entries; 2) Audio entries must be at least eight minutes in duration, or one page of typed text; 3) Participants must describe events, experiences, and observations candidly; 4) Journal entries must be transmitted to an analyst who is outside of the participants' organization soon after they are made for immediate review and analysis; and, 5) A mechanism must be in place to enable the analyst to intervene when a serious behavioral issue is detected without jeopardizing the confidentiality of participants.

ROBoT: In the past year of the ROBoT project, we worked on four main efforts.

Task Progress:

First, we completed running of our control subjects for the HERA portion of the project. This included N=14 subjects who were matched on age, sex, and education to the HERA participants.

Second, we conducted additional analyses on the data from the four HERA missions of Campaign 3. Our new analyses focused in particular on comparing the healthy controls to HERA subjects. Participants in both groups steadily improved on the task over time, and at the group level there was no significant difference in learning rates. Interestingly, however, HERA participants performed better (lower overall scores and higher percent success) yet significantly slower as compared to controls. This suggests the HERA participants adopted a different strategy as compared to controls. Specifically, HERA participants appeared to focus on more careful and accurate performance of the ROBoT-r task, trading this off against the time used to complete each run.

The data show a clear learning curve, where task scores improve, they perform the tasks faster, and the capture success is higher. Also, controls perform significantly worse but also significantly faster than HERA participants, suggesting a strategy difference between the two groups. These strategy differences were evident across all difficulty levels, although the groups were more similar in performance and duration on the most-difficult trials. The performance effects were largely driven by errors in angular alignment at the point of contact, and not linear displacement error.

Third, we completed data collection from crewmembers at Neumayer station and conducted preliminary analysis of the data. Data included monthly ROBoT-r sessions monthly from Feb through November 2017 (10 months) from each of N=7 crewmembers. Data collection from these subjects was >96% complete, expecting 12 runs from each of 7 subjects (84 total) per month.

Preliminary analyses showed significant differences in angle error at grapple as a function of difficulty, no significant differences in distance error as a function of difficulty, and no significant differences in overall score as function of difficulty. A notable performance improvement in overall score appeared in May (month=6), but this was not significant

due to substantial between-subject variability.

Trials come in groups of 3 (trials 1-3=25%, of maximum ROBoT-r task difficulty, whereas trials 4-6=50%, trials 7-9=75% and trials 10-12=100%). Angle error exhibits the largest effect of difficulty, as found in prior studies.

Thus, while there were generally clear effects of trial difficulty, there was no significant group effect of time in mission. That result, however, masks the substantial intra-individual variability (from month to month) as well as between-subject variability. In the coming year, we will investigate optimal approaches to characterize this type of within- and between-subject variability and seek to develop methods for individualized performance prediction.

Subject N6-17 showed steady performance early, followed by steady improvement (scale=0-10, with 0=best). Subject N7-17 exhibited one particularly good month (June), with poorer performance most of the rest of the year.

Fourth, and finally, we continued working towards the feasibility study of ROBoT-r testing aboard the ISS. This involved IRB (Institutional Review Board) approvals, revising the ROBoT-r scoring system to better align with astronaut expectations, modifying the user feedback screens to avoid conflicts with ROBoT trainer feedback, developing training materials, informed consent briefings, team member training for astronaut instruction, and developing baseline data collection plans suitable for both US-based and Russia-based crews. All such steps are essentially complete and two crewmembers have agreed to participate in the in-flight BCM feasibility study, including ROBoT-r.

Team Measure Questionnaires: Data from the Team Measures battery was successfully acquired from all four 30-day HERA C3 missions (N=16 individuals) with 100% compliance. At the end of the previous reporting period, data from the initial battery of Team Measures was collected in two four-person crews (N=8) throughout two 30-day missions in the HERA facility. During the current reporting period Team Measures data was collected in the remaining two 30-day missions (N=8). Preliminary analyses of the HERA data, thematic analysis of the initial Team Measures battery, and assessment of operational acceptability led to recommendations for a reduced and modified Team Measures battery and data collection schedule for long-duration missions in operational environments. This iteration of the Team Measures battery was deployed in a long-duration mission at Neumayer Station in Antarctica (N=7 individuals). Data processing is currently in progress. The pre-mission demographics and IPIP-NEO-120 measures were completed by all participants (100% compliance). The in-mission measures included a combined Team Performance, Team Processes, and Team Climate session scheduled for completion every two weeks, as well as the BCM's newly developed Group Living assessment scheduled for completion every month. Compliance for the bi-weekly team measures was generally high, with 75% of sessions completed. Notably, non-compliance was not distributed evenly across the mission. Specifically, no data were recorded for January or February 2018, although it is uncertain why this lapse occurred. Interestingly, and despite the less frequent administration schedule, compliance for the Group Living measure was low, with only 32% of sessions completed (2017 February, July, November, and 2018 March).

Lessons learned from this assessment have been incorporated into the development of Standard Measures for spaceflight. Specifically, the Group Living measure is no longer a 360 assessment where every crewmember rates every other crewmember, instead opting for a format where each item refers to the entire team. Although this provides lower resolution data, it is more operationally acceptable and still provides insight on this critical competency for long-duration exploration missions.

Discussion: Data collected during the 14-month Neumayer mission have shown that the standardized suite of Behavioral Core Measures is feasible, acceptable, and reliable for tracking cognitive performance and behavioral health during a long-duration mission in extreme conditions (12-14 months). The possible exception of Journals (only 23% of expected data were acquired) may reflect a cultural bias of the German crew.

We are currently working with ISSMP on finalizing implementation of the BCM measures for deployment on-board the ISS during which feasibility and acceptability will be determined for spaceflight conditions.

Bibliography Type:	Description: (Last Updated: 05/08/2025)
Articles in Peer-reviewed Journals	Roma PG, Reed DD, DiGennaro Reed, FD, Hursh SR. "Progress of and prospects for hypothetical purchase task questionnaires in consumer behavior analysis and public policy." The Behavior Analyst. 2017 Nov;40(2):329-42. Review. https://doi.org/10.1007/s40614-017-0100-2 , Nov-2017
Books/Book Chapters	Roma PG, Bedwell WL. "Key factors and threats to team dynamics in long-duration extreme environments." in "Team Dynamics Over Time: Advances in Theory, Methods, and Practice, vol. 18." Ed. E. Salas, W.B. Vessey, L.B. Landon. Bingley, UK: Emerald Publishing, 2017. p. 155-187. https://www.emeraldinsight.com/doi/full/10.1108/S1534-085620160000018007 , Aug-2017