International Journal of Applied Technology & Leadership
ISSN 2720-5215
Volume 3, Issue 1, January 2024
ijatl@org

Space Habitat Human Factors Testing Using Simulated Space Missions

Scott Van Hoy

Capitol Technical University (USA)

Michael Johnson, PhD

Capitol Technical University (USA)

Abstract

Usability, acceptability, and habitability questionnaires have previously been developed and implemented by the National Aeronautics and Space Administration (NASA) to collect human factors engineering feedback about both operational and prototype space habitats. During a twelve-day simulated space mission at the Integrated Lunar/Mars Analog Habitat at the University of North Dakota, three crew members submitted habitability feedback using an adapted NASA questionnaire to measure the types, novelty, and priorities of feedback collected using space simulation as a human-in-the-loop (HITL) testing method. Of the categories surveyed, habitat equipment, environment, interface, and crew health were the most frequently reported feedback, followed by operations, activities, and use errors. The crew reported no feedback for recreation or group interaction. Most feedback for habitability items with the priority 'must be addressed' occurred by the second mission day. However, novel feedback with the priority 'nice to have' sustained new reports throughout the simulation. This study showed that short-duration mission simulations can be a useful HITL testing method to solicit actionable engineering feedback.

Keywords: Evaluation, Testing, Design, Space, Habitability, Usability, Acceptability

1. Introduction

The design and functionality of future terrestrial space habitat systems will directly impact an astronaut crew's mission performance, and engineers must assess the habitability hazards associated with habitat design and operations before a space organization establishes a habitat on the moon or Mars [1]. The National Aeronautics and Space Administration (NASA) developed several questionnaires and human-in-the-loop (HITL) testing methods measuring the human factors and behavioral performance outcomes from space habitats' usability, acceptability, and habitability.

NASA tested four usability, acceptability, and habitability questionnaires paired with HITL testing to collect feedback on existing systems, establish engineering requirements for future habits, and provide input on commercial vehicle mockups [2],[3],[4]. This testing involved crew members living within an operational habitat (e.g., the International Space Station) or a habitat mockup in conjunction with simulated space mission operations (e.g., the Human Exploration Research Analog), where they completed questionnaires about their subjective judgment of the systems. These established testing methods can be adapted and utilized as part of the developmental testing of future space habitability systems.

This article reviews the established HITL testing methods specific to space habitats and evaluates to what extent future terrestrial space habitat prototype testing might benefit from soliciting crew feedback during a mission simulation that includes the habitation of a mockup or prototype. A two-week simulated space mission was conducted at the Integrated Lunar/Mars Analog Habitat (ILMAH) at the University of North Dakota, with habitability feedback solicited from the three-person crew. The hypothesis was that by conducting HITL testing during a mission simulation, the crew would provide novel human factors feedback over the mission's duration. This novel feedback during a simulated mission could provide valuable information to the engineering team during the mockup and prototype phases of terrestrial space habitats.

2. Background

NASA developed several assessment tools and programs to assess the human factors engineering of space habitats using HITL testing. These methods were tested in operational and simulated environments and are primarily used to collect crew members' usability, acceptability, and habitability feedback. This section reviews four human factors methodologies implemented by NASA to assess space habitats.

2.1. Interactive Space Habitability Reporting and Observation Tool

NASA's Human Factors and Behavioral Performance Element created a mobile application called Interactive Space Habitability Reporting and Observation Tool (iSHORT) that allows crew members to provide input regarding vehicle habitability. The description provided on the mobile application's download page, "iSHORT is an iPad application designed to collect habitability observations (including text, photos, video, and audio recordings) from research study participants" [5]. NASA tested this application during Human Exploration Research Analog (HERA) missions at Johnson Space Center, the NASA Extreme Environment Mission Operations (NEEMO) missions at Aquarius Reef Base, and on the International Space Station (ISS) [4],[6],[7]. iSHORT leveraged the crew members living within each habitat to assess the habitability of each facility, providing feedback for future designs.

iSHORT provides a method of collecting real-time human factors and habitability data while crew members are living in the habitats. Previous human factors data collection was accomplished through post-mission debriefs, relying on human memory for accurate feedback [6]. When crew members use iSHORT to provide habitability feedback, they can choose from text, audio, photo, and video multimedia options to share their thoughts and suggestions [4]. Crew members are encouraged to provide good and bad comments about the environment and working environment, examples of difficulties, things that worked well, impacts on task performance, performance of equipment and systems, and suggestions for improvements [5]. Topics reported include labeling, hygiene, windows, stowage, lighting, housekeeping activities, work volumes, environmental factors, scheduling, procedures, recreation, team dynamics, exercise, training, EVA, cable management, and trash management [7]. In addition to standard habitability and human factors observations, walk-throughs and narrated task videos were conducted on the ISS using iSHORT [7].

2.2. NASA Next Space Technologies for Exploration Partnership

NASA established the Next Space Technologies for Exploration Partnerships -2 (NextSTEP-2) program to stimulate the commercial space industry and help the agency achieve its future exploration goals. Under NextSTEP-2, NASA solicited commercial partners to develop space habitation systems, stimulating the design and prototyping of space habitats [8]. A ground test and analysis protocol created by NASA assessed the habitats for mission durations of up to sixty days. The ground test included four components: inspection, subsystem standalone tests, analysis, and human-in-the-loop (HITL) integrated tests [2],[9].

The HITL testing phase used NASA astronauts (N = 4 per test) to provide subjective evaluations of the habitat mockups while participating in a high-fidelity space simulation with a supporting mission control. This integrated testing phase focused on the functional requirements of the habitat related to habitability, human factors, and crew performance. The astronaut crews completed questionnaires consisting of Likert scales that measured the acceptability and capability of mission tasks and habitat configuration, crew workload during the task, and fatigue after task completion [2],[9].

The simulated missions were three days, with day one focusing on habitat-centric function, day two on distributed function allocation, and day three on extravehicular activity (EVA). Before the start of the simulation, the crew received six hours of training inside the mockup habitats [2]. The mission timeline and mission control integration were designed to simulate both habitation and operational tasks that would be required during a cis-lunar mission, such as sleep, hygiene, meals, exercise, systems monitoring, maintenance, robotic teaming, EVA, and various science experiments [2],[9]. The results of the questionnaires helped determine which habitat design characteristics were most beneficial to crew performance. They were published as habitat design guidelines, enhancing and enabling mission capabilities and improved habitability characteristics [2].

2.3. The Subjective Habitability and Acceptability Questionnaire

The Subjective Habitability and Acceptability Questionnaire (SHAQ) is an instrument designed with the core goal of quantifying the effects of habitat size and layout on behavioral and human performance (BHP) outcomes in operational environments using subjective individual perceptions [3]. The scale uses a bipolar visual analog scale from negative one hundred to positive one hundred with zero meaning no perceived effect. SHAQ assesses thirty-six metrics using a matrix-style visual scale with six BHP metrics and six habitability moderators. The BHP outcomes include the performance of individual activities, group activities, mood, psychological stress, sleep, and social interactions. The habitability moderators have privacy, social density, efficiency, control, comfort, and convenience [3].

SHAQ was tested during a campaign at NASA's HERA during mission simulations. Sixteen crew members across four forty-five-day HERA missions completed the survey seven times per mission for each functional habitat area. The results contributed to evidence-based recommendations for habitat size, layout, and design [3].

2.4. The Scale for Habitat Usability

The Scale for Habitat Usability (SHU) was an attempt to create a "gold standard" questionnaire for capturing users' subjective viewpoints regarding space habitats [10]. The SHU is administered after a task of interest to evaluate the task environment based on intuitiveness, labeling, layout, lighting, satisfaction, situational awareness, and workload. A panel of subject matter experts at NASA paired with NASA human test subject questionnaires established these categories [11]. The questionnaire assesses each category using a Likert scale with one meaning 'strongly disagree' and five meaning 'strongly agree' [10]. The SHU analyzes the specific tasks' environment and provides feedback for habitat and vehicle design.

3. Methodology

This study adapted iSHORT to collect habitability feedback during a twelve-day mission at the Integrated Lunar/Mars Analog Habitat at the University of North Dakota. ILMAH is a multi-purpose NASA-funded space habitat mockup with a plant production module, EVA and maintenance module, research module, exercise module, and habitation module. An exterior tunnel system connects each habitat module [12].

3.1. iSHORT Adaptation

iSHORT, SHAQ, SHU, and NextStep surveys were each designed to provide human factors feedback regarding habitability, usability, acceptability, and task performance for space habitats. During their initial uses, SHU was a task-specific questionnaire. SHAQ quantified the behavioral and human performance factors of habitat size and layout, NextSTEP solicited astronaut feedback on specific habitat capabilities, and iSHORT used multimedia and real-time human factors feedback for habitability data collection. Using numerical scales, SHAQ, SHU, and NextStep quantified the crew's environmental and task design perceptions. These scales can provide valuable information to engineers, especially early in habitat design during requirements development.

iSHORT is unique because it was used operationally and tested on the ISS, NEEMO, and HERA. iSHORT focuses on qualitative crew feedback and allows crew members to use various media to explain their perceptions of the habitat. The priority system also gives engineers a better sense of how to allocate assets best when improving designs. Due to the tangible engineering feedback the crew can provide using this format, this study selected iSHORT as the base for the habitability feedback form in this study.

The Habitability Feedback Form (HFF) is the derivative of iSHORT used during this study. Google Forms is the host software for the HFF. The instructions provided in the survey were the same as iSHORT, and the crew was encouraged to provide good and bad comments about the environment and working environment, examples of difficulties, things that worked well, impacts on task performance, the performance of equipment, and systems, and suggestions for improvements [5].

The HFF consisted of four questions. Question one was a category selection. The available categories and associated examples were compiled from each referenced survey, iSHORT, SHAQ, SHU, and NextSTEP, to encourage crew feedback across different habitability tasks and environments. Crew members were permitted to select multiple categories per entry.

Table 1. List of categories available on the Habitability Feedback Form and the examples of items that fall into each category.

HFF Category	**Examples Given to the Crew**
Habitat interface	Information systems, habitat systems use, etc.
Use error	Crew error because of inadequate engineering and/or design.
Environment	Noise, lighting, temperature, odor, etc.
Recreation	Personal time, group activities, etc.
Group interactions	With ground, leadership, the crew, cultural differences, etc.
Health	Hygiene, sleep, sanitation, medicine, food, exercise, psychological needs, etc.
Operations and activities	Timeline, schedules, training, procedures, etc.
Equipment	Hardware, clothing, stowage, etc.

Question two was a priority ranking. The crew ranked the feedback as either ‘must be addressed,’ ‘nice to have,’ or ‘no change needed.’ Question three was the crew’s observation as a long-form text entry. Question four allowed an optional upload of photos, videos, and audio files.

The HFF is an iSHORT adaptation designed to solicit crew feedback regarding habitability during a simulated space mission in the mockup and prototype phases of space habitats. The expanded categories prompt the crew on the types of feedback the engineering team might desire if using simulated space missions as part of the habitat’s developmental testing and evaluation program.

3.2. Simulated Mission Scenario

Three crew members conducted two days of training covering the ILMAH habitat systems and then lived within the habitat for twelve days during a simulated stay on the Martian surface. The primary mission objectives were safe habitation, education, and research. Each crew member had a personal research project in addition to supporting multiple research projects sponsored by the American Public University System Analog Research Group (AARG).

The crew was encouraged to complete the HFF whenever they identified feedback valuable to engineering improvements to the ILMAH habitability systems. The feedback started on the first training day, with two hours allocated to the crew obtaining familiarity with the HFF. After the first training day, mission timelines provided no dedicated time to complete habitability feedback forms. The crew was encouraged to complete at least one survey per crew member daily on a not-to-interfere basis with other mission objectives.

3.3. Analysis

Inclusion in this study’s final dataset depended on the quality of the HFF report. The included feedback was specific enough to provide actionable information to generate an engineering action item. Once passing the inclusion criteria, entries were categorized as a first-time entry, a repeat entry with no amplifying information, or a repeat entry with amplifying information. Amplifying information is additional information, video, photos, or other media that would help influence habitat engineering and design. Due to the crew not living within ILMAH during the two training days, these first two days serve as the control and a basis for comparing crew feedback categories.

4. Results

Over the two training days and twelve-day simulated mission, the crew completed 51 habitability feedback forms. Of the 51 completed forms, 48 met the inclusion criteria of providing habitability information that is specific and actionable. Each form was classified into nine categories with multiple categories assigned, if applicable. The average number of categories assigned per report was 1.84 (σ = 1.00).

Table 2. Overview of the submitted feedback.

Feedback Forms		**Media**	
Total Forms	51	Total # Photos	18
# Forms Included	48	Avg Photos per Survey (σ)	0.375 (0.73)
# Forms Not Included	3	Total # Videos	0
Categories		Avg Words per Report (σ)	34.94 (21.28)
Habitat Interface	15	**Report Novelty**	
Use Error	3	First Time Reported	42
Environment	16	Repeat Reports (No Amplifying Information)	0
Recreation	0	Repeat Reports (Amplifying Information)	6
Group Interaction	0	**Priority**	
Health	17	Must be Addressed	17
Ops and Activities	6	Nice to Have	27
Equipment	37	No Change	4

Reports were most likely to be related to habitat equipment (N = 37), followed by crew health (N = 17), habitat environment (N = 16), habitat interface (N = 15), operations and activities (N = 6), and then use error (N = 3). There was no feedback for the 'recreation' or 'group interaction' categories. The crew completed 17 of the 48 forms (35.4%) before ingressing the habitat, sixteen forms during the two training days, and one on the morning of mission day one before ingress. Twenty-nine forms were completed by mission day two (60.4%) and 41 by mission day seven (85.4%).

Equipment and habitat interface feedback were prevalent throughout the entire duration of the training days and the mission. Health and environmental feedback were most prevalent during the training days and the first two mission days. Operations, activities, and use errors had no related pre-mission feedback and occasional yet minimal feedback throughout the mission.

Figure 1. Cumulative number of novel, specific, and actionable feedback reports by category.

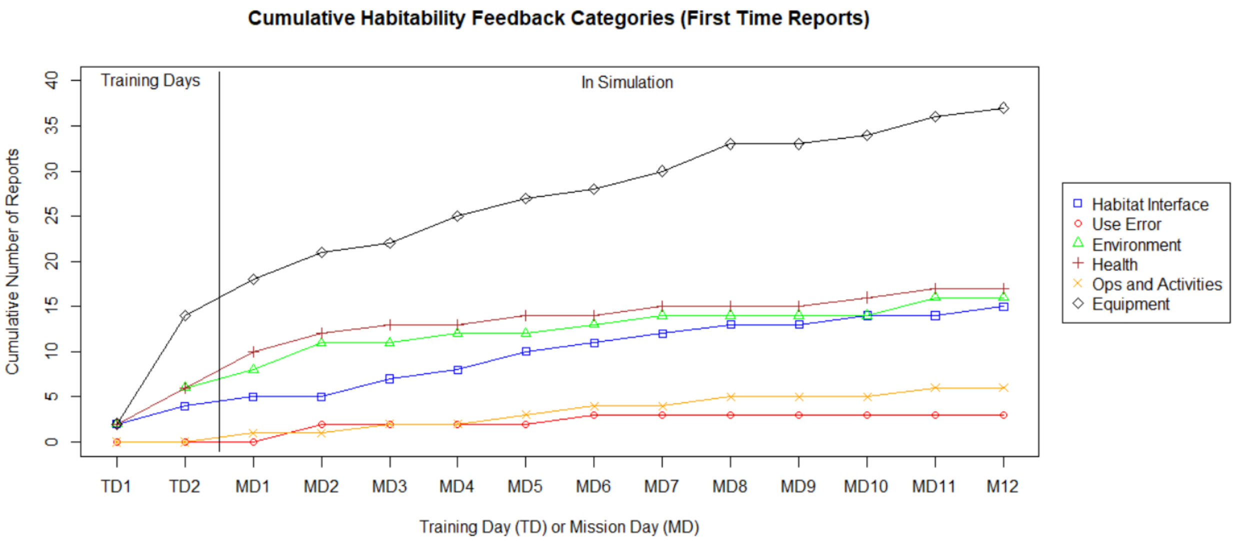

Figure 1 shows the categories associated with the habitability feedback throughout the mission, with repeated reports excluded. During the training days, the crew conducted training within the habitat but did not live within the habitat. Mission day one was the ingress day and the first night spent inside the habitat. While in the habitat, the crew lived under simulation protocol and could only leave if wearing the North Dakota Experimental -2 Analog Trainer (NDX-2AT) extravehicular activity suit.

Figure 2. Cumulative number of novel, specific, and actionable feedback reports by assigned priority.

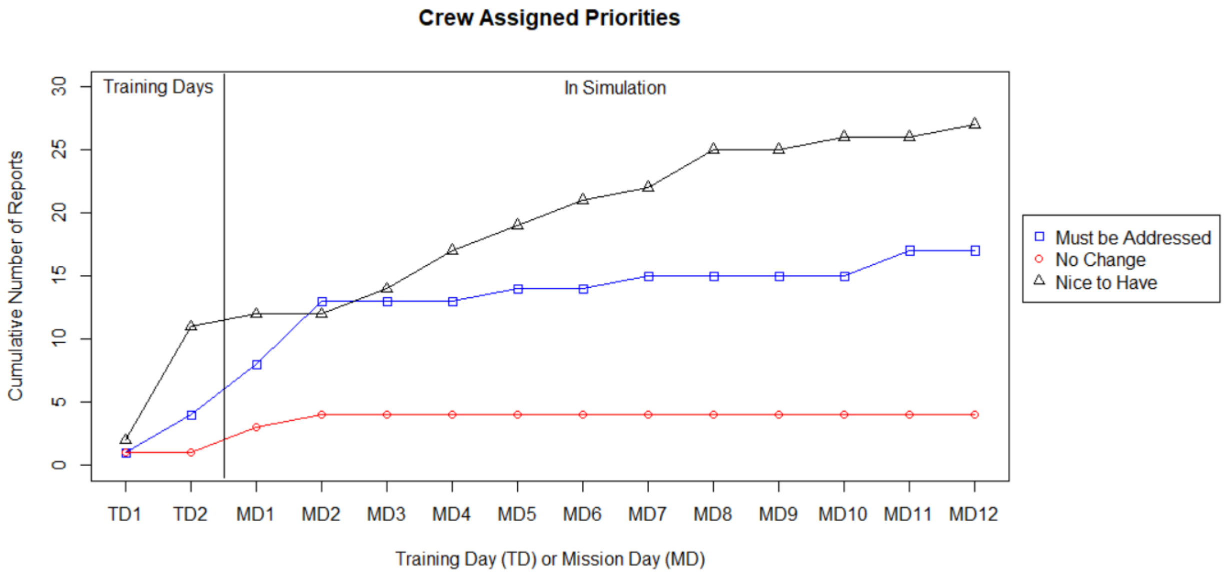

Figure 2 shows the cumulative crew assigned priorities for each feedback form. 'Nice to have' was the most common feedback (N = 27), then 'must be addressed' (N = 17), and the least common feedback was 'no change' (N = 4). Four of the 17 'must be addressed' priorities were submitted during the training days (23.5%), 13 of the 17 were submitted by mission day two (76.5%), and 15 were submitted by mission day seven (88.2%). 11 of the 27 ‘nice to have’ priorities were submitted during the training days (40.7%), 12 of the 27 were submitted by mission day two (44.4%), and 21 were submitted by mission day seven (77.8%). The crew submitted one ‘no change’ feedback form during the training days; the additional three forms were submitted during mission days one and two.

5. Discussion

This study hypothesized that by conducting HITL testing during a mission simulation, the crew would provide novel human factors feedback over the mission’s duration, providing valuable information to the engineering team during the mockup and prototype phases of terrestrial space habitats. Using an adapted version of iSHORT, the crew provided habitability feedback starting during the training days and continued feedback throughout the mission. iSHORT was selected for its extensive use in developing habitat requirements, multimedia feedback, and easy integration into mission operations, allowing for actionable feedback.

The crew submitted approximately one-third of the total feedback before the start of the simulation, and about two-thirds were completed by mission day two. Most of the input came within the first four days of collection, with only one night’s sleep in the habitat. Feedback, particularly feedback prioritized as ‘must be addressed,’ were identified without sustained operations in simulation. The feedback gathering during the short duration supports the relatively short three-day simulation times used during the NextSTEP simulations.

Two of the three ‘must be addressed’ feedback items between mission days three and thirteen were related to the external condition of the habitat and were discovered during EVA. ILMAH sustained weather damage over its lifespan, causing the identification of this habitat condition. These feedback items were specific and actionable and thus included in the dataset. However, ILMAH is designed as a low-fidelity space habitat mockup, thus making ‘must be addressed’ feedback of the habitat’s exterior less realistic than is to be expected during mockup or prototype testing of future habitability systems. Due to two of the three ‘must be addressed’ items coming from the EVA-identified weathering, there was only one practical ‘must be addressed’ habitability item after mission day three.

The categories of feedback received in the simulation became a relevant factor in the habitability assessment. The ‘operations and activities’ and ‘use error’ categories had no reports during the training days. Additionally, only 44.4% of the feedback categorized as ‘nice to have’ was submitted by mission day two. This delayed feedback between the training and mission days supports the hypothesis that simulation-generated habitability feedback produces novel data.

Diminished returns were apparent by mission day seven. The crew submitted 85.4% of the total feedback reports, 88.2% of the ‘must be addressed’ reports, and 77.8% of the ‘nice to have’ reports during the first week of the simulation.

These results indicate that non-simulation HITL testing or simulated missions with durations of approximately two to four days may be sufficient to openly solicit the majority of feedback regarding habitability concerns that astronauts recommend as ‘must be addressed.’ To capture factors of habitability that are ‘nice to have,’ conducting a mission simulation with a crew living and working within the habitat allows for feedback regarding mission-centered activities and tasks.

After one week of simulation, the diminished returns indicate a potential aimpoint for a simulated space mission focused on HITL testing of a habitat prototype. However, the two-week duration did not result in feedback regarding recreation and group interaction. Two-week mission durations may not be long enough for the crew to experience inconveniences like limited privacy, reduced crew comfort, and negative social interaction [13]. Missions longer than two weeks may be required to obtain feedback within these categories using simulation. Additional research is needed to determine whether simulated mission durations longer than two weeks would be beneficial during prototype testing. Future long-duration missions to the moon and Mars may benefit from identifying habitability hazards that may arise later in the mission due to the habitat environment.

In addition to using simulation and an adapted iSHORT methodology for HITL testing, other surveys that utilize Likert scales and specific task subjective feedback will undoubtedly be necessary during the developmental testing of habitat systems. Future studies could compare the use of the different questionnaires discussed in this study when soliciting actionable engineering feedback on habitat prototypes.

6. Conclusions

Space mission simulation and HITL testing have been critical components in developing space habitats’ engineering requirements. The methods used to develop habitat requirements can be adapted to aid in the human factors testing of future lunar and Martian habitats. This study showed that short-duration mission simulations can be a useful HITL testing method to solicit actionable crew feedback.

The methods, scope, and durations of human factors testing scheduled for future habitability systems will depend on funding, development timelines, and mission risk acceptability. Additionally, other methods of usability testing, task analysis, and acceptability testing should be used throughout the engineering process to create a holistic human factors testing program. Space mission simulation for habitat human factors testing is just one potential tool to generate feedback from the astronauts before a flight to contribute to optimizing mission performance and success.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge the American Public University System Analog Research Group, the University of North Dakota Human Spaceflight Laboratory, Dr. Kristen Miller, Dr. Pablo de León, and Travis Nelson for their support in collecting the data used in this study.

Funding: This work was supported by the American Public University System Graduate Studies and Research Office.

References

[1] M. Cohen, Mockups 101: Code and standard research for space habitat analogues. AIAA Space 2012 Conference & Exposition, Pasadena, CA, United States. (2012).

[2] M. Gernhardt, S. Chappell, K. Beaton, H. Litaker, O. Bekdash, C. Newton, J. Stoffel, Deep space habitability design guidelines based on the NASA NextSTEP phase 2 ground test program (NASA/TP-2020-220505). Johnson Space Center, National Aeronautics and Space Administration. (2019).

[3] P.G. Roma, L.L. Blackwell, C.A. Spencer, A.M. Whitmire, T.J. Williams, The Subjective Habitability & Acceptability Questionnaire (SHAQ): Development and validation. Human Factors. 65(6) (2022) 1074–1104. https://doi.org/10.1177/00187208211053462.

[4] S. Thaxton, Habitability assessment of International Space Station. Lead Increment Scientists Briefing, National Aeronautics and Space Administration. (2015).

[5] National Aeronautics and Space Administration, iSHORT. https://apps.apple.com/id/app/ishort/id1489223302, 2019 (accessed 22 September 2023).

[6] S. Thaxton, H. Litaker, K. Holden, J. Adolf, J. Pace, R. Morency, Developmental testing of habitability and human factors tools and methods during Neemo 15. Human Research Program Investigator's Workshop, Galveston, TX, United States. (2011).

[7] S. Thaxton, M. Greene, S. Schuh, T. Williams, R. Archer, K Vasser, ISS habitability data collection and preliminary findings. Human Research Program Investigator's Workshop, Galveston, TX, United States. (2017).

[8] National Aeronautics and Space Administration, Next space technologies for exploration partnerships -2 (NextSTEP-2) Broad Agency Announcement NNH16ZCQ001K. (2016).

[9] M. Gernhardt, K. Beaton, S. Chappell, O. Bekdash, A. Abercromby, Development of a ground test & analysis protocol for NASA's NextSTEP phase 2 habitation concepts. 2018 IEEE Aerospace Conference, Big Sky, MT, United States. (2018).

[10] I. Robertson, R. Amick, A summary of the development and validation of the Scale for Habitat Usability (SHU). NASA Technical Reports Server (No. 20230000377). (2023).

[11] I. Robertson, B. Munson, J. Stephenson, R.Z. Amick, Developing and validating measurement scales during pandemic conditions: a case study with the Scale for Habitat Usability. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 66(1) (2022) 677–681. https://doi.org/10.1177/1071181322661420.

[12] S. Van Hoy, B. O'Hara, D. Wallace, T. Trevino, R. Worku, K. Miller, American Public University System Analog Research Group mission 1: A space analog 11-day mission. Space Education & Strategic Applications. 3(1) (2022). https://doi.org/10.18278/sesa.3.1.3.

[13] C. Heinicke, M. Arnhof, A review of existing analog habitats and lessons for future lunar and Martian habitats. REACH. (2021) 21-22. https://doi.org/10.1016/j.reach.2021.100038.

