

IAC-22-A1-1-3-x69152

Isolation Standard Measures: a set of validated and feasible measurements ensuring comparability across isolation and confinement studies

Angelique Van Ombergen^{a*}, Didier Chaput^b, Elena Fomina^c, Valérie Gil^d, Michaela Girgenrath^e, Natalie Hirsch^d, Natsuhiko Inoue^f, Perry Johnson-Green^d, Thu Jennifer Ngo-Anh^g, Jancy C. McPhee^h, Alain Maillietⁱ, Keiji Murakami^j, Katrin Stang^e, Melanie von der Wiesche^k, Suzanne T. Bell^l, Erin E. Flynn-Evans^m, Sara E. Whitingⁿ, Lauren B. Landonⁿ, Brian E. Crucian^o, Sara Zwart^p, William B. Vessey^q

^a *SciSpacE Team, Research and Payloads Group, Directorate of Human and Robotic Exploration Programmes, European Space Agency, Keplerlaan 1, P.O. Box 299, 2200AG Noordwijk, The Netherlands, angelique.van.ombergen@esa.int*

^b *CADMOS- CNES (DOA/EVH/VHD), Centre National d'Etudes Spatiales, BPi 2221,18 Avenue Edouard Belin, 31401 Toulouse Cedex 9 – France, Didier.Chaput@cnes.fr*

^c *Laboratory of Countermeasures to Negative Effects of Microgravity, Department of Physiology of Human Gravitational Resistance, State Scientific Center of Russian Federation, Institute of Biomedical Problems, Russian Academy of Sciences, Khoroshevskoe Shosse, 76 A, 123007, Moscow, Russia, fomin-fomin@yandex.ru*

^d *Operational Space Medicine, Canadian Space Agency, 6767 route de l'Aéroport, Saint-Hubert, QC J3Y8Y9, Canada, natalie.hirsch@asc-csa.gc.ca; Health and Life Sciences, Canadian Space Agency, perry.johnson-green@asc-csa.gc.ca and valerie.gil@asc-csa.gc.ca*

^e *Research and Exploration, German Aerospace Center, DLR, Koenigswinterer Strasse 522-524, 53227 Bonn, Germany, katrin.stang@dlr.de and michaela.girgenrath@dlr.de*

^f *Space Biomedical Research Group, Astronaut and Operation Control Unit, Japan Aerospace Exploration Agency (JAXA), 2-1-1, Sengen, Tsukuba, Ibaraki, 305-8505, Japan, inoue.natsuhiko@jaxa.jp*

^g *Research and Payloads Group, Directorate of Human and Robotic Exploration Programmes, European Space Agency, Keplerlaan 1, P.O. Box 299, 2200AG Noordwijk, The Netherlands, jennifer.ngo-ahn@esa.int*

^h *Chief Scientist's Office, Human Research Program, Johnson Space Center, National Aeronautics and Space Administration, 2101 E. NASA Parkway, Houston, TX, 77058, United States of America, jancy.c.mcphee@nasa.gov*

ⁱ *MEDES-IMPS for CNES DS/DAP/SUE, Bpi2221, 18 Avenue E. Belin, 31401 Toulouse Cedex 9, France, alain.mailliet@cnes.fr*

^j *Human Spaceflight Technology Directorate, Japan Aerospace Exploration Agency (JAXA), 2-1-1, Sengen, Tsukuba, Ibaraki, 305-8505, Japan, murakami.keiji@jaxa.jp*

^k *Lead Site Management West, DLR - German Aerospace Center, Linder Hoehe, 51147 Koeln, Germany, melanie.vonderwiesche@gmx.de*

^l *Behavioral Health and Performance Laboratory, Biomedical Research and Environmental Sciences Division, Johnson Space Center, National Aeronautics and Space Administration, 2101 E. NASA Parkway, Houston, TX, 77058, United States of America, suzanne.t.bell@nasa.gov*

^m *Fatigue Countermeasures Laboratory, Human Systems Integration Division, Ames Research Center, National Aeronautics and Space Administration, Moffett Field, CA, 94035, United States of America, erin.e.flynn-evans@nasa.gov*

ⁿ *Human Factors and Behavioral Performance Element, Human Research Program, Johnson Space Center, National Aeronautics and Space Administration, 2101 NASA Parkway, B21, Houston, TX, 77058, United States of America, lauren.landon@nasa.gov*

^o *Immunology/Virology Laboratory, Environmental Sciences Branch, Johnson Space Center, National Aeronautics and Space Administration, 2101 E. NASA Parkway, Houston, TX, 77058, United States of America, brian.crucian-1@nasa.gov*

^p *University of Texas Medical Branch, Nutritional Biochemistry Laboratory, Johnson Space Center, 2101 E. NASA Parkway, Houston, TX, 77058, United States of America, sara.zwart-1@nasa.gov*

^q *Research Operations and Integration, Human Research Program, Johnson Space Center, National Aeronautics and Space Administration, 2101 E. NASA Parkway, Houston, TX, 77058, United States of America, william.b.vessey@nasa.gov*

* Corresponding Author

Abstract

Isolation and confinement studies have been essential to the preparation of crewed long-duration space missions, acting as analogues that facilitate the study of psychological and physiological responses to isolation and confinement. They also serve as opportunities for the development, testing, and validation of countermeasures and coping methods to handle the challenges that arise in such scenarios. Due to the small sample sizes in combination with high inter-individual variability, single isolation campaigns are not always sufficient for obtaining statistically significant scientific findings. To address this issue and improve comparability between different studies, the need for standardized measures to be collected in all future isolation and confinement studies was identified. Additionally, these measures should also be shown to be generally valid, reliable, feasible and acceptable in analogue and spaceflight environments. Standard measures in isolation and confinement studies allow for more direct comparisons of results and synthesis of data across isolation and confinement studies as well as provide an important step toward standard measures in spaceflight. An international expert group with representatives from different space agencies worldwide was brought together to define a core set of standard measures for isolation and confinement studies. This paper provides an overview of the expert group's recommendations for international standard measures for future isolation and confinement studies, along with subsequent updates coordinated by the International Countermeasures Working Group (ICMWG), which was established as a sub-Working Group under the International Space Life Sciences Working Group (ISLSWG). Additional experts were consulted by the ICMWG partner agencies as required.

The collection of the described set of isolation standard measures will provide data on the following parameters: sleep, mood, psychological state, psychophysiology, cognitive performance, stress and the immune system, general health and well-being, team measures, nutritional measures, and environmental conditions. For each measure, recommendations were made about duration and frequency of administration, along with specific implementation recommendations in relation to the duration of the isolation study. The set of isolation standard measures will be re-assessed every two years at a minimum to ensure they are up to date and reflect the current state-of-the-art.

Keywords: Isolation, Confinement, Standard Measures, Spaceflight, Space Analog

Acronyms/Abbreviations

Appraisal of Life Events Scale (ALE), Bio-electrical Impedance Analysis (BIA), Center for Epidemiologic Studies Depression Scale (CES-D), Cognition Test Battery (CBT), Council on Nutrition and Appetite Questionnaire (CNAQ), dehydroepiandrosterone (DHEA), Dual Energy X-ray absorptiometry (DXA), Food Craving Questionnaire (FCQ), Hypothalamic-pituitary-adrenal axis (HPA axis), General Health Questionnaire (GHQ), Global Physical Activity Questionnaire (GPAQ), International Countermeasures Working Group (ICMWG), International Space Life Sciences Working Group (ISLSWG), Isolation Standard Measures (ISM), neuropeptide Y (NPY), Positive and Negative Affect Schedule (PANAS), Profile of Mood States (POMS), Pittsburgh Sleep Quality Index (PSQI), Social Desirability Scale (SDS), Team Climate Inventory (TCI), Visual Analogue Scale (VAS), World Health Organization (WHO).

1. Introduction

Isolation and confinement studies have been essential to the preparation of crewed long-duration space missions, acting as analogues that facilitate the study of psychological and physiological responses to isolation and confinement. They also serve as opportunities for the development, testing, and validation of countermeasures and coping methods to handle the

challenges that arise in such scenarios. Due to the small sample sizes and resulting high inter-individual variability that they provide, single isolation campaigns are not sufficient for obtaining statistically significant scientific findings. To tackle this issue and improve comparability between studies, the need for validated standardized measures was formulated to be collected in all future isolation and confinement studies with the aim of significantly improving comparability between studies. Additionally, these measures should also be shown to be generally feasible and acceptable in analogue environments.

The general concept of Standard Measures is that the measure has been shown to be valid, reliable, feasible, & acceptable in the spaceflight environment. As a general rule we would recommend that new technologies could be evaluated as a potential standard measure in the future (possibly carried for consideration in the 'Future Measures to Be Considered' section of the document; Annex 1), but these new technologies as they currently stand would not meet the definition of a Standard Measure.

An international expert group was brought together to define international standard measures to act as a core set of measures for standard data collection. This report provides an overview of the expert group's recommendations for international standard measures

for future isolation and confinement studies along with subsequent updates coordinated by the International Countermeasures Working Group (ICMWG). Additional experts were consulted by the different partner agencies as required.

It should be noted that this report is a dynamic document and will be adapted as different technologies and measures change and become validated over time. It should be emphasized that the Isolation Standard Measures are not meant to be exhaustive. Additional and experimental measurements can be added as applicable to the objective of a given isolation study.

2. Overview of Isolation Standard Measures

Please note that where possible, the participant should take the version of the measure validated in their preferred language. In the absence of a validated measure in a language in which the participant is fluent, best practices for conducting assessment in a different language should be followed to the extent possible. Translation and back-translation by two separate fluent readers is a suggested minimum. Sufficient documentation of the measure is required to ensure appropriate interpretation of the data.

2.1 Sleep

2.1.1 Sleep Diary

Concurrent with actigraphy, to improve interpretability of the actigraphy data, a sleep diary should be implemented. The sleep diary will help to comprehensively document and quantify sleep and waking behaviours and provide context for the data provided by the actigraphy watches. These sleep diaries typically include questions regarding a participant's perception of time they went to sleep, time they woke up, number of awakenings and what happened during those awakenings, and other events that could influence sleep such as exercise, naps, and food intake.

2.1.2 Sleep Diary

The PSQI is a well-validated measure of subjective sleep quality (Backhaus et al. 2002; Grandner et al. 2006) [1-2]. It is a self-report questionnaire, made up of 19 individual items and 7 component scores on: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, and daytime dysfunction. Unlike the sleep diary, the PSQI captures an overall subjective measure of sleep over the course of the previous month.

2.2 Mood, Psychological State, and Psychophysiology

2.2.1 Profile of Mood States (POMS)

The Profile of Mood States is a well-validated and commonly used measure of mood state (McNair et

al. 1971) [3]. The questionnaire is short, simple to administer and measures how six different dimensions of mood (tension, depression, anger, vigour, fatigue, confusion) vary over time. The questionnaire should be filled out in identical conditions each time to the maximum extent practicable. These should include the same place and time of day, no or minimal distractions, privacy and the person alone (or positioned so others cannot see answers), the person alone and timed to avoid major scheduled events that might arouse particular emotional reactions. In particular, it is recommended to be filled on a weekday, mid-morning, and to avoid administration immediately following exercise, meals, or procedures involving pain (e.g., blood draws, etc.). If constraints to avoid certain mood-altering activities prior to administering cannot be met, the prior activity should be captured (e.g., What activity did you do just prior to completing this survey? Eating, sleeping, etc.).

2.2.2 Positive and Negative Affect Schedule (PANAS)

In order to obtain a comprehensive understanding of the effect of isolation and confinement on both positive and negative affective state, and to avoid overburdening the participants with the same questionnaire, it is recommended that the PANAS (Crawford and Henry 2004) [4] also be administered during isolation studies, unless the study's science team requests only one of these questionnaires. The PANAS, unlike the POMS, provides a more balanced approach to assessing negative versus positive states. By alternating the PANAS and POMS (See Section 3), the monotony of completing questionnaires can be reduced, and a fuller view on mood state variations can be captured.

2.2.3 Appraisal of Life Events Scale (ALE)

The ALE is recommended to be administered only pre- and post- future isolation studies as a measure of the effect of isolation and confinement on the participant's assessment of stressful life events. The ALE scale measures three factors of appraisal: threat, challenge and loss. It has been found to be both valid and reliable (Ferguson et al. 1999) [5].

2.2.4 Five Factor Model Personality Inventories

The Five Factor Model Personality Inventories such as the NEO Personality Inventory (NEO-PI-3) and the International Personality Item Pool (IPIP-NEO) are well-established, reliable and valid (McCrae et al. 2011) [6] personality questionnaires, which should be administered prior to a study for all study durations, as well as after a study if the isolation period is greater than 6 months. The questionnaires measure a participant's Big Five personality traits: openness to experience, conscientiousness, extraversion,

agreeableness, and neuroticism, as well as six subcategories of each trait.

2.2.5 Social Desirability Scale (SDS)

The SDS (only to be administered pre- isolation period) measures the extent to which a participant is influenced by the desire to respond to questionnaires in a socially approved manner. This questionnaire (Reynolds 1982) [7] is recommended to allow scientists to interpret the other questionnaires with an understanding of the extent that social desirability bias may play a role in the responses.

2.3 Cognitive Performance

2.3.1 Cognition Test Battery

The Cognition battery is a neurocognitive assessment tool developed for spaceflight (already part of NASA's Spaceflight Standard Measures) which is composed of 10 tests designed to measure the main cognitive domains: executive, episodic memory, complex cognition, social cognition, and sensorimotor speed (Basner et al. 2015) [8]. The tests were developed specifically for testing high-performing astronauts and astronaut-like individuals. These tests are administered within the Cognition software, with the most recent versions capable of running on iPads and computers running Windows 7 or higher. The battery starts with a brief questionnaire, followed by administration of the 10 cognitive tests. The total administration time for one battery (questionnaire plus all 10 tests) is between 20 and 30 minutes. Typically, the first session is preceded by a 60-minute familiarization session, which includes a standardized video introduction, brief slides with technical notes, and a 20-30 minute practice administration that is not used for data analysis.

2.3.2 Subjective Assessment of Performance Visual Analogue Scale (VAS)

The VAS is recommended as a simple method for measuring a characteristic along a continuum and has been found to have good validity and reliability (de Boer et al. 2004) [9]. It has been used in measuring several different factors, such as stress, pain, and anxiety (Williams et al. 2010; Lesage and Berjot 2011; Lesage et al. 2012) [10-12] and can be completed rapidly and with minimal participant burden. For relevant items to include in isolation studies, see (Basner et al. 2015) [8].

2.4 Stress and Immune System

2.4.1 Salivary Stress Markers

Salivary cortisol is a widely used biomarker of psychological stress (Bozovic et al. 2013) [13]. Assessment of salivary dehydroepiandrosterone (DHEA), Neuropeptide Y (NPY), and alpha-amylase will complement the cortisol measurement by

monitoring HPA axis and sympathetic mediated alterations.

2.4.2 Salivary Markers of Latent Herpesvirus Reactivation

Most adults are previously infected with one or more herpesviruses (EBV, VZV, HSV, etc.). These viruses establish 'latency' in a body location and are generally well controlled in immunocompetent individuals. Certain stressors that compromise immune function may result in the reactivation of these viruses, after which virus DNA will be detectable in saliva. Latent virus reactivation is therefore an excellent biomarker for clinically relevant immune dysregulation (Mehta et al. 2017) [14].

2.4.3 Plasma Cytokines

Blood soluble mediators (cytokines) of immune activity are readily detectable and relevant to identification of stress associated immune dysregulation. Multiplex array affords analysis of broad categories of cytokines (pro/anti-inflammatory, Th1/Th2 regulatory, chemokines, etc.) in a minimal plasma sample, allowing a comprehensive physiological assessment. (Crucian et al. 2014) [15]. Samples can be stored frozen until analysis.

2.4.4 Stress VAS

Please see subsection 2.3.2. This stress VAS is typically a single item indicator of stress.

2.5 General Health and Well-Being

2.5.1 Actigraphy

Wrist-worn actigraphy is a well-validated measure of rest/activity cycles (Weiss et al. 2010; Martin and Hakim 2011; Marino et al. 2013) [16-18]. While it does not provide a direct measure of sleep, it has been shown as an effective, non-invasive, low-burden method for inferring key sleep metrics (including the measurement of parameters such as total sleep time, sleep efficiency, sleep latency etc.). Continuous measurement throughout the studies is recommended.

2.5.2 General Health Questionnaire (GHQ)

The GHQ is a tool for identifying psychological distress and was part of the MARS500 standard measures. It is widely used and has been translated into dozens of different languages and is considered to be high in reliability and validity (Jackson 2006) [19].

2.5.3 Global Physical Activity Questionnaire (GPAQ)

The Global Physical Activity Questionnaire (GPAQ) was developed by the World Health Organization (WHO) in 2002. GPAQ comprises 19 questions grouped to capture physical activity undertaken in different behavioral domains: work, transport and discretionary

(also known as leisure or recreation). One additional item collected is time spent in sedentary activities. It is recommended to administer the GPAQ once pre-isolation.

2.5.4 Demographic Measures

The following demographic measures are recommended: Age, Gender, Biological Sex, Years of prior analogue experience, Years of military experience or Comparable experience, Nationality, Education, Profession, Languages spoken and level of proficiency. A question regarding social constructs such as race and ethnicity is recommended, following local research practices. It is recommended to use an open question, such as ‘In your own words, how would you describe your race or ethnicity?’.

2.5.5 Anthropometry

The following measures should be collected: Weight, Height, Waist-Hip Ratio, Weight Changes.

2.5.6 Body Composition

Body Composition could be measured with e.g., Bio-electrical Impedance Analysis (BIA) or with Dual Energy X-ray absorptiometry (DXA) scans.

2.5.7 Center for Epidemiologic Studies Depression Scale (CES-D)

CES-D is a 20-item clinically validated self-report of depression symptoms. It is widely used in clinical research, is in the public domain, and is available in multiple languages (English, German, Russian, Japanese, etc.). It includes reverse-coded items that allow for validity checks of careless responding. This measure is intended to supplement and accounts for some inconsistencies in quality of data produced by the GHQ Severe Depression subscale. The recommended use of this scale is as a monitoring tool, planned for use once around the mid-point of isolation missions and with additional use on an as-needed basis in instances when responses to the GHQ and/or POMS depression subscales indicate concerning trends. It does not include the highly sensitive suicidal ideation questions that have led to missing GHQ data in the past

2.5.8 Critical Incidents Log

A critical event is any unusually stressful event that may lead to psychological or psychosocial damage/problems, and where, therefore, crisis intervention of some sort is required. The critical incidents log is only completed when a critical event occurs. The log of critical incidents is completed by the monitoring psychologist. Observations of significant events that affect participant’s psychological state are described. Entries may require an interview with the

participant by the psychologist. This log is useful to help explain issues that cannot be quantified.

2.5.9 Physical Performance Test

A test for aerobic functional capacity will provide an objective measure of the participants’ cardiorespiratory fitness. An incremental-load cycle ergometer test is recommended pre-isolation to establish maximum heart rate and maximal aerobic capacity and post-isolation to identify any changes in cardiorespiratory fitness.

2.6 Team Measures

2.6.1 Team Conflict

Team conflict has a long-standing history of analog data collection via measures of various forms. Based on Basner et al. (2014) [20], this team conflict measure includes items asking participants whether they have had a conflict with other analog crew members, with similar items addressing different roles (e.g., conflict with mission control). If participants indicate a conflict has occurred, they are asked to indicate whether it was resolved.

2.6.2 Team Process

Team processes are grouped into transition (mission analysis, goal specification, strategy formation and planning), action (monitoring progress toward goals, systems monitoring, team monitoring and backup, coordination), and interpersonal processes (conflict management, motivating and confidence building, affect management). (Mathieu et al. 2020) [21] recently created a shortened measure of the key team processes, which demonstrated validity and reliability relative to longer measures of those constructs. This version of the measure is heavily based on the Mathieu et al. (2020) [21] version, but with some minor adaptations for analogs and is currently undergoing validation in an isolation analog setting.

2.6.3 Team Cohesion

Measurement of two key facets of team cohesion (i.e., task cohesion and social cohesion) are important when predicting team outcomes (Salas et al. 2015) [22]. While there are several methods and measures of cohesion, the measure recommended (Kozłowski 2015) [23] captures both task and social cohesion, is short and appropriate for an operational environment, and has shown good reliability and convergent/discriminant validity in preliminary studies (Powers, 2012) [24].

2.6.4 Team Climate

The Team Climate Inventory (TCI) (14-item version) is a structured self-report measure designed to assess the “facet-specific climate for innovation within groups at work” based on four dimensions: vision, participant

safety, task orientation, and support for innovation (Anderson & West 1998, p. 235) [25].

2.6.5 Team Performance

Team Performance Tasks may vary from analog to analog. Metrics however could be collected in a standardized way to provide an objective assessment of team performance, given the team task abides by some basic parameters. As an example, *number of task objectives complete on a team task* (relative to a baseline), could provide a meaningful assessment of team performance across an isolation mission, while not restricting an analog to a specific measure. An example team performance task is provided below.

2.6.5.1 Problem-Solving Task: Survival Analysis Task

This task presents participants with survival scenarios (e.g., moon, desert, winter). Each participant is provided an explanation of the scenario and a list of items which they rank in order of their importance to crew survival. Crewmembers complete the task individually, then work together to create an overall crew consensus ranking. Task performance is scored by computing the difference between the crew ranking of the item and the experts' rankings for each scenario. 30min per session with a recommended minimum of 3x in-isolation (early, mid, late). The measure is widely available in the literature and has been iterated to examine performance over time in analogs (Larson et al. 2019) [26].

2.7 Nutrition Measures

2.7.1 Council on Nutrition and Appetite Questionnaire

Changes in appetite should be monitored. The Council on Nutrition and Appetite Questionnaire (CNAQ) (Wilson et al. 2005) [27] could be used.

2.7.2 Food Craving Questionnaire

The incidence of food cravings should be monitored. The Food Craving Questionnaire (FCQ-T) questionnaire could be used.

2.7.3 Food Frequency Questionnaire/Food Record

Links between nutrition and mood and other parameters are known (Cook et al. 2020) [28] and patterns in food consumption should be documented. This could be done through the use of food frequency questionnaires or digital food records, to estimate energy and nutrient intake (Boushey et al. 2017) [29], or through a structured food diary.

2.8 Environmental Conditions

Temperature and humidity should be monitored and shall be documented throughout the whole study. It is assumed that atmosphere is Earth normal at sea level. In instances in which a facility is pressurised, O₂ and CO₂ should be monitored.

To achieve standardized study conditions, it is very important to control the environmental conditions such as room temperature and humidity. Room temperature and humidity should be kept at comfortable levels. It would be recommended to ensure values in the range of 21°C to 25°C for temperature and 50% to 70% relative humidity for most of the day so that sweating or freezing shall be avoided as both would affect study results.

3. Baselines, Duration, And Frequency of the Measurements

The recommended measures, along with the minimum frequency at which the measures should be administered throughout studies, details about the baseline data collection and the implementation time for each measure can be found in the following table, with in-isolation implementation defined by total isolation time in the analog. Note that all suggested measurement frequencies are minimums and additional sampling, particularly around key milestones or events, may be useful (see Table 1 at the end).

4. Recommendations for Future Actions

To keep the standard measures up to date, it is also recommended that an expert group reconvenes through teleconferences every two years at minimum to reassess the literature and determine whether or not updates are needed. After concluding an isolation mission, it is recommended to reassess and evaluate before collecting the ISM at a future opportunity. Objective measures of behavior and performance, at the Team and Individual level, and other future work items captured in the original report, should continue to be pursued.

5. Conclusions

The purpose was to provide an overview of the measures that this expert group recommends for future isolation and confinement studies. The measures described here will allow the collection of a set of isolation standard measures (ISM), which will provide data on the following parameters: sleep, mood, psychological state, and psychophysiology, performance, stress and immune system, general health and well-being, team measures, nutritional measures, and environmental conditions.

Acknowledgements

The authors would like to thank the ESA Space Isolation Focus Team (Prof. Dietrich Manzey, Prof. Denise Ferravante, Dr. Adrianos Golemis, Prof. em. Peter Suedfeld, Prof. Elisabeth Rosnet, Prof. Margarete Boos, Prof. John Drury and Dr. Sabrina Moerk),

Andrea Rossiter, Dr. Laetitia Vega, Dr. Phyllis Johnson, Prof. Bernd Johannes, Prof. Alexander Stahn, Prof. Elisabeth Rosnet, Dr. Alexandra Whitmire, Dr. Go Suzuki and Prof. Gro Sandal for their valued contributions.

References

- [1] J. Backhaus, K. Junghanns, A. Broocks, D. Riemann D, F. Hohagen. Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. *J Psychosom Res* 53 (2002) 737–740.
- [2] M.A. Gradner, D.F. Kripke, I-Y. Yoon, S.D. Youngstedt. Criterion validity of the Pittsburgh Sleep Quality Index: Investigation in a non-clinical sample. *Sleep Biol Rhythms* 4: 129–136, 2006.
- [3] D.M. McNair, M. Lorr, L.F. Droppelmann. *Manual for the Profile of Mood States*. San Diego, CA: Educational and Industrial Testing Services, 1971.
- [4] J.R. Crawford, J.D. Henry. The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *Br J Clin Psychol* 43: 245–265, 2004.
- [5] E. Ferguson, G. Matthews, T. Cox. The Appraisal of Life Events (ALE) scale: Reliability and validity. *Br J Health Psychol* 4: 97–116, 1999.
- [6] R.R. McCrae, J.E. Kurtz, S. Yamagata, A. Terracciano. Internal Consistency, Retest Reliability, and Their Implications for Personality Scale Validity. *Personal Soc Psychol Rev* 15: 28–50, 2011.
- [7] W.M. Reynolds. Development of reliable and valid short forms of the marlowe-crowne social desirability scale. *J Clin Psychol* 38: 119–125, 1982.
- [8] M. Basner et al. Development and Validation of the Cognition Test Battery for Spaceflight. *Aerosp Med Hum Perform* 86: 942–52, 2015.
- [9] A.G.E.M. de Boer et al. Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? *Qual Life Res* 13: 311–320, 2004.
- [10] V.S. Williams, R.J. Morlock, D. Feltner. Psychometric evaluation of a visual analog scale for the assessment of anxiety. *Health Qual Life Outcomes* 8: 57, 2010.
- [11] F.X. Lesage, S. Berjot. Validity of occupational stress assessment using a visual analogue scale. *Occup Med (Chic Ill)* 61: 434–436, 2011.
- [12] F.X. Lesage, S. Berjot, F. Deschamps. Clinical stress assessment using a visual analogue scale. *Occup Med (Chic Ill)* 62: 600–605, 2012.
- [13] D. Bozovic, M. Racic, N. Ivkovic. Salivary Cortisol Levels as a Biological Marker of Stress Reaction. *Med Arch* 67: 374, 2013.
- [14] S.K. Mehta et al. Latent virus reactivation in astronauts on the international space station. *npj Microgravity* 3: 11, 2017.
- [15] B.E. Crucian et al. Plasma Cytokine Concentrations Indicate That In Vivo Hormonal Regulation of Immunity Is Altered During Long-Duration Spaceflight. *J Interf Cytokine Res* 34: 778–786, 2014.
- [16] A.R. Weiss et al. Validity of Activity-Based Devices to Estimate Sleep. *J Clin Sleep Med* 06: 336–342, 2010.
- [17] J.L. Martin, A.D. Hakim. Wrist Actigraphy. *Chest* 139: 1514–1527, 2011.
- [18] M. Marino et al. Measuring Sleep: Accuracy, Sensitivity, and Specificity of Wrist Actigraphy Compared to Polysomnography. *Sleep* 36: 1747–1755, 2013.
- [19] C. Jackson. The General Health Questionnaire. *Occup Med (Chic Ill)* 57: 79–79, 2006.
- [20] M. Basner et al. Psychological and behavioral changes during confinement in a 520-day simulated interplanetary mission to mars. *PloS one*, 9(3), e93298, 2014.
- [21] J.E. Mathieu et al. The Development and Construct Validity of a Team Processes Survey Measure. *Organ Res Methods* 23: 399–431, 2020.
- [22] E. Salas et al. Measuring Team Cohesion. *Hum Factors J Hum Factors Ergon Soc* 57: 365–374, 2015.
- [23] S.W.J. Kozlowski. Advancing research on team process dynamics. *Organ Psychol Rev* 5: 270–299, 2015.
- [24] C.L. Powers. What does cohesion capture? An empirical and conceptual analysis [Poster presentation]. 27th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA, United States, April, 2012
- [25] N.R. Anderson and M.A. West. Measuring Climate for Work Group Innovation: Development and Validation of the Team Climate Inventory. *Journal of Organizational Behavior*, 19, 235-258, 1998.
- [26] L. Larson et al. Team performance in space crews: Houston, we have a teamwork problem. *Acta Astronaut* 161: 108–114, 2019.
- [27] M-M.G. Wilson et al. Appetite assessment: simple appetite questionnaire predicts weight loss in community-dwelling adults and nursing home residents. *Am J Clin Nutr* 82: 1074–1081, 2005.
- [28] N.R. Cook. Sodium and health—concordance and controversy. *BMJ* (June 26, 2020).
- [29] C.J. Boushey et al. New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods. *Proc Nutr Soc* 76: 283–294, 2017.

Table 1: Overview of the Isolation Standard Measures and their baseline, duration, and frequency.

Parameter	Measure	Required or Recommended	Duration	Pre-Isolation	In-Isolation ≤ 3 Months	In-Isolation 4-6 Months	In-Isolation >6 Months	Post-Isolation
Sleep	Sleep Diary	Required	7 consecutive nights (5-10 min)	During same weeks as actigraphy	Monthly	Monthly	Monthly	During same week as actigraphy
	Pittsburgh Sleep Quality Index (PSQI)	Required	5-10 min	Once	Monthly	Monthly	Monthly	Once
Mood, Psychological State and Psychophysiology	Profile of Mood States (POMS)	Required*	5 consecutive days, 5 min	Daily for five consecutive days pre-mission	Daily for five consecutive days, upon ingress and bi-weekly	Daily for five consecutive days, upon ingress and once per month	Daily for five consecutive days, upon ingress and once per month	Daily for five consecutive days post-mission
	Positive and Negative Affect Schedule (PANAS)	Required*	5 consecutive days, 5 min	Daily for five consecutive days pre-mission	Daily for five consecutive days, upon ingress and bi-weekly	Daily for five consecutive days, upon ingress and once per month	Daily for five consecutive days, upon ingress and once per month	Daily for five consecutive days post-mission
	Appraisal of Life Events Scale	Required	5 -10 min	Once	N/A	N/A	N/A	Once
	Five Factor Model Personality Inventories	Required	20-25 min	Once	N/A	N/A	N/A	Once; only if study lasts longer than 6 months
	Social Desirability Scale	Required	< 5 min	Once	N/A	N/A	N/A	Once
Cognitive Performance	Cognition Test Battery	Required	20-30 min	5 sessions to address learning effects	Weekly	Bi-Weekly	Monthly	Once

	Subjective assessment of performance (VAS)	Required	<5 min	Daily during same weeks as actigraphy	Weekly	Bi-Weekly	Monthly	Daily during same week as actigraphy
Stress/Immunity**	Salivary Stress Markers (Cortisol/DHEA/NPY/alpha-amylase)	Required	< 5 min morning and evening	During same weeks as actigraphy	Weekly	Bi-Weekly	Monthly	During same week as actigraphy
	Salivary Markers of Latent Herpesvirus Reactivation	Required	< 5 min, 3 consecutive days	During same weeks as actigraphy	Monthly	Monthly	Bi-Monthly	During same week as actigraphy
	Plasma Cytokine Concentrations	Required	<5 min	During same weeks as actigraphy	Monthly	Bi-Monthly	Quarterly	During same week as actigraphy
	Stress Visual Analog Scale (VAS)	Required	<5 min	Daily during same weeks as actigraphy	Weekly	Bi-Weekly	Monthly	Daily during same week as actigraphy
General Health and Well-Being	Actigraphy	Required	Continuous	Continuously for one week in early and one week in late pre-mission	Continuous	Continuous	Continuous	Continuously for one week
	General Health Questionnaire (GHQ)	Required	7-8 min	Once	Monthly	Monthly	Monthly	Once
	Global Physical Activity Questionnaire (GPAQ)	Required	15min	Once	N/A	N/A	N/A	N/A
	Demographic measures	Required	10 min	Once	N/A	N/A	N/A	N/A
	Anthropometry	Required	30 min (full measures) <5 min (weight only)	Once	Weekly (weight only)	Bi-Weekly (weight only)	Bi-Weekly (weight only)	Once
	Body Composition	Recommended	Variable depending on technique	Once	N/A	N/A	N/A	Once
	Center for Epidemiological Studies	Recommended	10 min	N/A	Once mid-mission	Once mid-mission	Once mid-mission	N/A

	Depression Scale (CES-D)				and as needed based on POMS/GHQ responses	and as needed based on POMS/GHQ responses	and as needed based on POMS/GHQ responses	
	Critical Incidents Log	Required	10 min	N/A	When Applicable	When Applicable	When Applicable	N/A
	Physical Performance Tests	Recommended	60 min	Once	N/A	N/A	N/A	Once
Team	Team Conflict	Required	<5 min	Weekly through BDC/training	Every 3 days	Weekly	Weekly	Once
	Team Process	Required	<5 min	Weekly through BDC/training	Weekly	Bi-Weekly	Monthly	Once
	Team Cohesion	Required	<5 min	Weekly through BDC/training	Weekly	Bi-Weekly	Monthly	Once
	Team Climate	Recommended	<5 min	Weekly through BDC/training	Weekly	Bi-Weekly	Monthly	Once
	Team Performance	Required	30 min	Once	Quarterly	Quarterly	Quarterly	Once
Nutrition	Council on Nutrition and Appetite Questionnaire	Recommended	<10 min	Once	Monthly	Monthly	Monthly	Once
	Food Craving Questionnaire	Recommended	<10 min	Once	Monthly	Monthly	Monthly	Once
	Food Consumption Record (e.g., photo before and after each meal)	Recommended	<5 min	Daily during BDC/training	Daily	Daily	Daily	N/A
	Food Frequency Questionnaire	Recommended	<10 min	Once	Weekly	Weekly	Weekly	N/A
Environmental Conditions	Environmental Measures (temperature, humidity, atmosphere)	Required	Passive habitat measure	N/A	Continuous	Continuous	Continuous	N/A

* It is required to collect either the POMS or PANAS, but it is recommended to collect both.