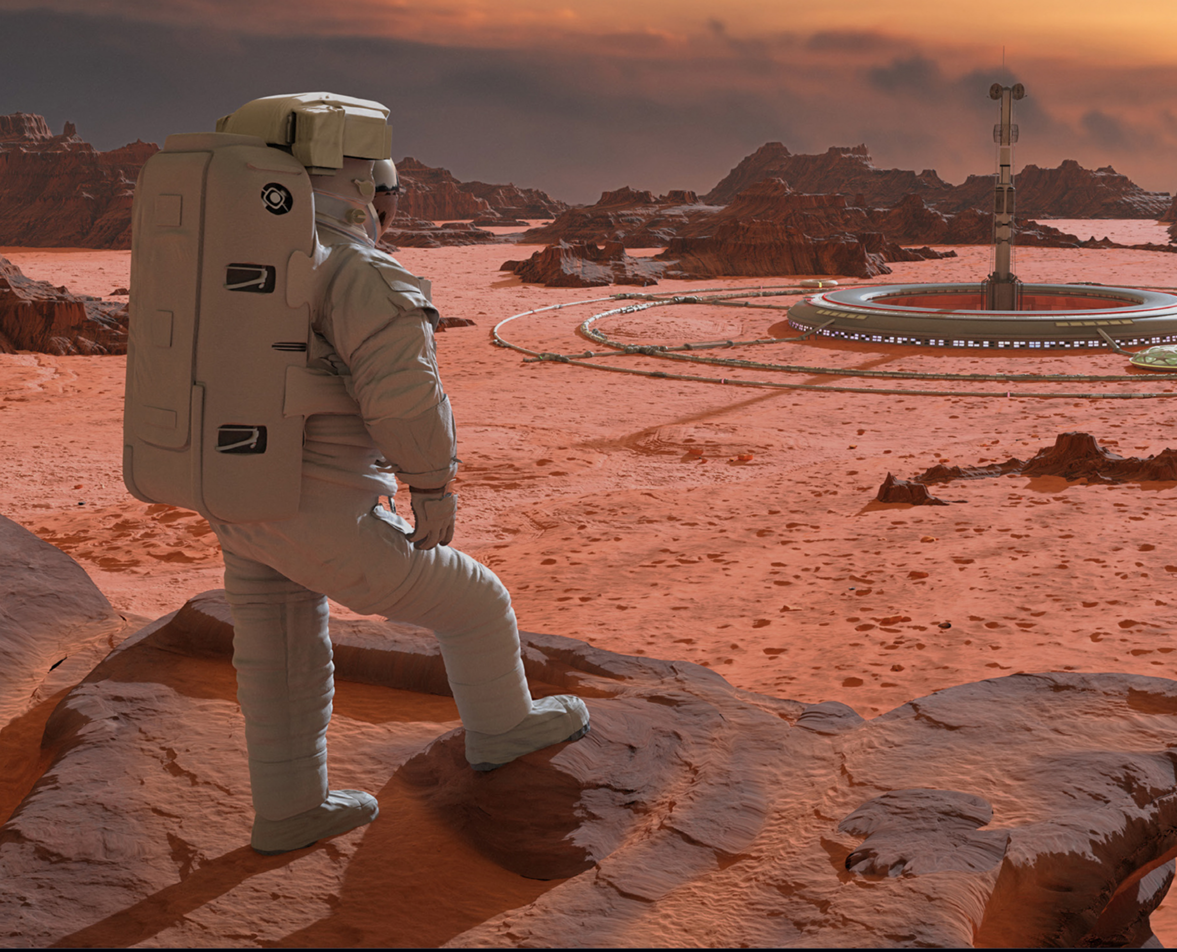

Psychology and Human Performance in Space Programs

Research at the Frontier

Edited by
Lauren Blackwell Landon
Kelley J. Slack
Eduardo Salas

CRC Press
Taylor & Francis Group

3 Special Considerations for Conducting Research in Mission-Simulation Analog Environments

Challenges, Solutions, and What Is Needed

Suzanne T. Bell
DePaul University
KBR/NASA's Lyndon B. Johnson Space Center

Peter G. Roma and Bryan J. Caldwell
KBR/NASA's Lyndon B. Johnson Space Center

CONTENTS

INTRODUCTION

Future space exploration, such as missions to the moon, asteroids, and Mars, present a uniquely challenging environment within which astronauts and crews must operate. Individuals and teams will work and live for extended durations under physical conditions of zero or low gravity, limited variety of food and exercise, and increased radiation exposure, combined with psychosocial stressors of confined space, limited privacy, variable workload, and isolation from family and friends. Although much is known about how individuals and teams operate in traditional work environments, the extent to which our understanding of individuals and teams generalizes to the final frontier is not always clear.

Research is conducted in analog environments, or settings that share one or more defining features of the target environments or populations to which the research is expected to generalize, to characterize the risks associated with spaceflight. As an example, bed rest studies, in which individuals remain in a head down tilted position for extended periods of time, are used to better understand the effects of microgravity including fluid shifts, bone and muscle loss, cardiac alterations, sensorimotor deficits, and visual impairments. Analog research also allows for problems to be solved or countermeasures to be developed and refined before deployment to an operational environment. For example, lighting changes to help regulate sleep/wake rhythms can be tested in analog environments before they are installed in the International Space Station (ISS). Although research in analog environments can provide critically important data on the risks of working and living in space with fewer constraints than the target operational setting, it has considerable challenges.

The purpose of this chapter is to communicate key challenges, current solutions, and necessary developments for conducting research in space analog environments. We narrow our focus to a few key issues that are fairly unique to conducting research in space analog settings. Challenges and issues that affect research more broadly also apply to analog research (e.g., ethics of risk-benefit to participants), but are beyond the scope of this chapter. We focus on issues related to identifying and selecting analogs; obtaining access to analogs; data collection challenges including the logistics of data collection in hard-to-access locations, lack of criterion data, and small sample sizes; and disseminating findings while maintaining confidentiality. For each of these, we describe the challenges, current solutions, and identify what is needed. Key ideas are summarized in Table 3.1.

CHALLENGE 1: IDENTIFYING AND SELECTING ANALOGS

Several analog environments are available for research. A primary challenge is identifying the most appropriate one(s) for specific research questions. There is no comprehensive list of analog environments, in part because analogs are defined by the research questions and populations to which they are expected to generalize. Almost all the major space agencies operate or have affiliations with analog environments around the world. For those unfamiliar with spaceflight analogs, most can be found via an Internet search. For example, the National Aeronautics and Space Administration (NASA) keeps an extensive list of analogs they operate, fund, or are affiliated with (https://www.nasa.gov/analogs/types-of-analogs). Literature searches in academic databases and space agency databases, technical manuscripts, evidence reports, key reviews, attending space-oriented research conferences, and Internet searches of moon and Mars nonprofit organizations also are a fruitful means of identifying potential analogs.

There are many risks to consider, and every analog or potential analog has its own unique set of physical, human factors, and operational features and limitations. Because analog environments are not the target environment (e.g., a long-duration

TABLE 3.1
Key Challenges, Current Solutions, and What Is Needed

Specific Challenges	Current Solutions	What Is Needed
	Challenge 1: Identifying and Selecting Analogs	
No comprehensive list of analogs Appropriateness of an analog is dependent on the research questions and populations of interest to which the data are expected to generalize	Frameworks for suitability of an analog presented in Keeton et al. (2011), and Landon et al. (2018) Systematic approaches to describe the context (Bell et al., 2018), which can facilitate a comparison between the target and analog environment	Research (e.g., meta-analysis and experimental manipulation) to determine which features of the context impinge on individual and team functioning Descriptions of commonly used analogs that capture key features of the context, as well as updates and manipulations over time
	Challenge 2: Obtaining Access to Analogs	
Limited access for researchers, with varying processes to obtain access across different analogs Time-consuming process	Agencies have eliminated some redundancies (e.g., dual application, review process) Researchers can create a "best-case" protocol and a "minimum acceptable data yield" protocol Research funding expenditures should account for data collection delays Knowledge transfer should be actively managed across the lifespan of the research	Improving the predictability of when parts of the research process are likely to occur Creating efficiencies so that the process is less cumbersome Orientation training to new researchers, so that the researchers and analog managers develop a shared understanding of the research process in the analog
	Challenge 3: Data Challenges	
Hard to access location A lack of criterion data Small sample sizes Inappropriateness of inferential statistics and the null hypothesis significance testing framework for small sample size research	Establishing good rapport and a shared understanding with a local point of contact Identifying unobtrusive and existing data that are practical, relevant, and sensitive Using alternative data analysis approaches and simulation where appropriate Collecting data in such a way so that it is comparable across multiple analogs (e.g., same measures for similar constructs; Standardized Measures for Bed Rest Studies)	Ensuring that appropriate inferences are being made from small sample research Improved understanding of the appropriate analysis of and inferences from small sample size research Development and adoption of international standardized measures if not already established
	Challenge 4: Disseminating Findings and Maintaining Participant Confidentiality	
Maintaining data and protocol confidentiality while disseminating findings	Scrutinizing tables, figures, and other descriptions for breeches of confidentiality Principle investigators proactively creating a coordinated outreach plan, key talking points, etc.	Development of a standardized media training, which includes special issues related to data and protocol integrity in analogs

space exploration mission), they are necessarily imperfect approximations. Because of this, the appropriateness of an analog depends on the research questions and populations of interest to which the data are expected to generalize. Any given analog may be more appropriate for some research questions than others. “Natural” analogs are typically professional or operational settings that exist for purposes outside of data collection. For example, Concordia Station in Antarctica often serves as an analog for biomedical and behavioral research, as well as countermeasure testing through science campaigns organized by the European Space Agency (ESA). A crew of approximately 13 researchers and support personnel stay at Concordia during each winter-over season (typically February through October). Small crews, the dangerous physical environment surrounding the base, and inaccessibility during the winter-over months (~February through October) create an isolated, confined, and autonomous operational environment analogous to what is similar to a long-distance space exploration mission. Such an environment may be appropriate for research questions focused on individual and team behavioral health, performance, and biopsychosocial adaptation over long periods of time. However, Concordia Station or other remote outposts may not be suitable for studying other spaceflight relevant topics such as multiteam systems because these bases often operate autonomously, with virtually no “Mission Control” and very little remote mission support. In this case, more controlled analogs or isolation chamber habitats that have a mission control, such as the Institute of Biomedical Problems (IBMP) sponsored ground-based experimental complex (abbreviated NEK for its name in Russian) in Moscow may be more appropriate. Further, isolated, confined, and controlled environments, such as chamber habitats, more readily allow for more experimental manipulations and environment alterations than do natural analogs.

Solutions

Ideally, a comprehensive and continuously updated listing of all analogs would be available, outlining strengths and limitations for modeling aspects of space exploration. Keeton and colleagues (2011) provide a framework for systematically assessing the suitability of different analogs for research, with some updates provided by Landon, Slack, and Barrett (2018). However, a truly comprehensive list does not exist. Given the wide variety of potential analogs, unpredictable availability of existing analogs, increasing emergence of new analogs, and multiple (and often changing) organizations responsible for managing any given analog, a singular resource that is current and regularly updated is unlikely. Instead, there are some general guidelines that researchers should incorporate into the decision-making process when identifying and selecting analog environments for research.

The appropriateness of an analog environment is determined by identifying the phenomena of interest, and then systematically identifying key contextual features including the characteristics of the population of the operational environment to which the research is intended to generalize that are likely to impinge on the phenomena of interest. Bell, Fisher, Brown, and Mann (2018) outline an approach for doing so in extreme team research. Extreme teams complete their tasks in "unconventional performance environments and have serious consequences for failure" (p. 2, Bell et al., 2018). Astronaut teams are an example of extreme teams. The approach outlined can also be applied to individual-level analog research.

Context is defined as the situational opportunities and constraints that affect the occurrence of and meaning of behavior, as well as the functional relationships between variables (Johns, 2006). Context operates at two levels: the omnibus context and the discrete context. The omnibus context is the context broadly considered as a "bird's eye" view of the context of interest. One approach to describing the omnibus context is to take a journalistic approach – identifying the who, what, when, and why of the context of interest (Bell et al., 2018; Johns, 2006). The discrete context is the particular variables (e.g., isolation, interdependence) that shape human affect, behavior, and cognition (Johns, 2006). Features of the discrete context that are likely to impinge on the phenomena of interest should be driven by theory, and prior empirical research where possible (e.g., autonomy was manipulated and showed an effect on communication patterns).

Importantly, context should be considered systematically, and features of the discrete context as relevant to the phenomenon of interest should be documented. Bell et al. (2018) provide tables of the omnibus context and discrete context of future long-distance spaceflight as it pertains to team composition research. Such summaries can be combined with the specific research questions of interest to identify key features necessary in the analog environment. These summaries also can be reported in later publications and technical reports to better lend the research to integration across studies, an important consideration for analog research because small sample sizes are common.

After systematic identification and description of contextual features, the psychological and physical fidelity of a potential analog environment and the intended population should be compared. Note that fidelity is not a feature of the analog *per se*; it is dependent on the research topic and population of interest. Once context has been described systematically, inductive reasoning will often be enough to determine the suitability of the analog in representing the target environment. This process, however, requires the researcher to clearly articulate their needs and the key contextual features likely to impact the phenomena of interest, and for those involved with the administration or operation of analog environments to effectively communicate the characteristics of their analog. Data on the physical and psychological fidelity of the analog can be collected from facility description documents, subject matter experts, and previous or current participants, as well as the extent to which specific features of the context are present (e.g., participant ratings of autonomy). Fidelity information can be communicated in several ways. As an example, in their review of team composition research conducted in analog environments, Bell, Brown, Outland, and Abben (2015) provide a numeric rating of the fidelity of the analog research to long-distance space exploration. Importantly, it will be common that an analog does not share all the contextual or sample characteristics likely to affect the phenomena of interest. When that occurs, the limitations to the generalization of the analog research to the target population should be clearly communicated in any report or publication.

What Is Needed

Two aspects of the approach noted above are particularly difficult: determining the key feature(s) of the context likely to impinge on the phenomena of interest and having an adequate understanding of the analog to determine the fidelity and generalizability of the research to the target environment or population. A better understanding of how features of the context affect individual and team behavior and performance is necessary to appropriately select a site for research and countermeasure testing. Documentation and communication of the context and populations used with the analogs also are needed.

We suggest two ways that research can provide insights into which contextual features affect the phenomena of interest. First, within one study features of the context can be varied (e.g., workload) so that the effect of the context on the focal construct can be estimated. Second, the influence of the context on the phenomena of interest can be investigated through meta-analysis or systematic review. Meta-analysis provides a systematic means of quantitatively aggregating multiple primary studies so that summary statements can be made regarding the phenomena of interest. One of the strengths of meta-analysis is that moderators (such as context) can vary across studies rather than within one. So, for example, a meta-analyst could examine the relationship of mood and instances of conflict with the habitable volume of a chamber as a moderator, even if habitable volume was not manipulated in the same primary study. The meta-analyst could code the effect size of the mood and conflict relationship, code the habitable volume of each of the analog environments from which they calculated an effect, and then determine if habitable volume moderated the effect using meta-analytic approaches. To make progress in understanding the influence of features of the context on relationships of interest requires systematic reporting of the context in primary studies; in this case, habitable volume of each of the analog environments would need to be reported in the primary study or located in other ways (e.g., documentation of the analog). Consistent systematic description of the context is necessary in analog research.

There are several analog environments that are consistently used in research (e.g., McMurdo and Concordia bases in Antarctica; NASA's Extreme Environment Mission Operations [NEEMO] and Human Exploration Research Analog [HERA], Hawai'i Space Exploration Analog and Simulation [HI-SEAS], the German space agency's DLR :envihab facility, IBMP's NEK). Detailed, systematic description of these analogs that are available to researchers is necessary. Features that are consistent across campaigns or missions can be provided in documentation or reports, and changes made or additional features for each campaign or mission should be reported as an appendix or similar for use by researchers. Such descriptions exist for some analogs but not others. They should be developed for those that do not have them. Researchers could use detailed documentation about analogs to design and propose better experiments with realistic expectations in mind.

Given that the important features will vary somewhat according to the phenomena of interest, identifying an appropriate analog requires an iterative top-down (those who run analogs communicating key features) and bottom-up (researchers using their expertise to estimate what might impinge on functioning for their phenomenon of interest) effort. This process can start with researchers providing requirements via tables describing the omnibus and discrete context as discussed earlier (see Bell et al., 2018 for examples), and the organization who sponsors the analog providing a systematic description of the analog. Then the circumstance a researcher opts to generalize and the analog of interest can be compared for fidelity for a specific phenomenon to identify appropriate analogs for research questions. In some cases, the appropriateness of the analog for the research question could be determined from existing documents; however, most often additional communication will be required. Even so the documentation can help develop a shared understanding between the researcher and those managing the analogs, and to guide conversations used to determine the appropriateness of the analog for research.

CHALLENGE 2: OBTAINING ACCESS TO ANALOGS

Admittedly, obtaining access to research analogs is difficult. From Antarctic research stations to isolation chamber facilities built and managed by space agencies and universities, analogs vary wildly in construction, location, operations, and management. Each analog has its own requirements for obtaining access; however, there are common challenges across many analogs. First, it is typically difficult to get approval to collect data in an analog – access is limited to a small group of researchers, and sometimes the process for accessing analogs is mysterious. Second, obtaining access to analogs is almost always time-consuming. Regardless of the location of the analog, duration of the study, and status of funding, an analog researcher should always expect an extensive planning effort, interdependence among multiple parties, and a multiyear process from proposal of a project to final return of data and equipment.

For analogs sponsored by space agencies or their research arms, the avenue is typically to propose research when requests for proposals are issued. Access is typically limited to investigators funded by the sponsoring agency or partner agencies. As an example, access to NASA-sponsored analogs currently happens by submitting proposed research to calls on the online NASA Solicitation and Proposal Integrated Review and Evaluation System (NSPIRES) system, clearing the multimonth peer-review and programmatic approval processes, and then conducting the awarded research. Most requests for proposals require the proposer to indicate their requirements from a NASA-sponsored analog. NASA then facilitates entry into the NASA-sponsored analog, or an analog of another agency whom with they may have a cooperative agreement such as the National Science Foundation (NSF), IBMP, or Deutsches Zentrum für Luft-und Raumfahrt (DLR). Likewise, analogs with other agencies have their own systems. The European Space Agency's (ESA) website provides opportunities for data collection at Concordia Station and posts Announcements of Opportunities (AOs). Additional requirements vary by agency and analog. For example, Concordia Station in Antarctica is jointly owned and operated by the French and Italian Polar Agencies (IPEV and PNRA, respectively);

ESA oversees all biomedical and behavioral research conducted there, and all principal investigators' home institutions must be based out of the European Union (EU).

Analogs that are not owned or directly managed by space agencies each have their own process. For example, HI-SEAS, which is owned by the Blue Planet Foundation and operated by the University of Hawai'i through grants, has included "opportunistic research" projects, where a researcher may propose their ideas directly to the primary project team. Analog environments run by independent nonprofit organizations (e.g., Mars Desert Research Station [MDRS] via The Mars Society) often let researchers apply directly to conduct studies and field tests, and information on how to apply can be found on their websites.

Many analogs will require what is similar to a grant application. This can create a double review process, because the funding is typically from the investigators' sponsoring agencies (so the merit was reviewed when applying for funding), then the merit is reviewed again for inclusion in an analog mission(s). Proposals are almost always subject to review for scientific merit, feasibility (including potential interference with other protocols), and probability of success. For example, for Concordia Station, ESA typically accepts four to five proposals per campaign. The proposals need to be either innovative science or a field test of a countermeasure or product ready for validation in an extreme environment or population. Almost all analogs, especially those sponsored by space agencies, will include multiple projects with potentially overlapping or competing protocol demands. Importantly, those who manage analogs look to reduce redundancies, and will often ask for researchers to combine parts of protocols and enter into data-sharing agreements.

Obtaining access to analogs is often a time-consuming process: analog research can span multiple years. An example project timeline may include: (a) year 1: a letter of intent or brief proposal submitted to the sponsoring agency, and if invited, development of a full proposal, review, and decision notification; (b) year 2: preparation, logistics, and data collection, or if a current campaign or mission is finishing, waiting months or another year to begin data collection; (d) years 3 and 4: data collection that often spans multiple missions to amass a meaningful sample size; and (3) years 5 and beyond: return shipment of equipment, biological samples, and raw data; data management and analysis; and writing final reports and manuscripts for publication. The time-consuming process is a significant barrier, as it would often seem ill-advised for tenure-track professors or professionals working in limited-term appointments (e.g., postdoctoral fellows) to commit to this process.

Given the extreme length of obtaining access to analogs combined with the longitudinal nature of research, the research project team is likely to be dynamic over the course of the project. Research assistants, students, and postdocs may come and go during a project's lifespan. Collaborators, consulting companies, subcontractors, and subteams may join or leave during a grant. For example, a consulting company may build technology in the first year or two that is used for data collection for the remainder of the project. There are often "follow-along" projects that are designed to build on previous research. Knowledge transfer must be actively managed.

Solutions

Several solutions are being implemented by agencies and organizations to varying degrees that help with this process mostly pertaining to streamlining, providing transparency, and increasing predictability. Researchers also can best prepare for and navigate this cumbersome process.

Agencies and organizations have been attempting to limit redundancies, such as double review systems, where possible and remove some of the burden and mystery of the process from the researchers. For example, in years past, proposals to NASA required the researcher to identify and obtain letters of agreement from analogs before submitting a proposal. This can be very difficult and often researchers for a particular topic may not necessarily have access to analog environments in their existing network. In recent years, NASA has modified their approach such that researchers must specify their requirements for an analog, then, if the grant is awarded, NASA facilitates entry into the NASA-sponsored analog, or an analog of another agency. Multiple facilities' descriptions are often in the calls for proposals. Further, cooperative agreements may also include the removal of a second review of scientific merit, which greatly facilitates the process. For example, NASA partners with the NSF for research at multiple Antarctic bases, and with the DLR for research in their :envihab facility, so that proposals do not need to be reviewed by multiple agencies' panels.

Researchers can take steps to facilitate access and streamline the process. First, to facilitate integration of multiple protocols and protect the integrity of all projects by minimizing participant burden and redundancy, researchers should consider two versions of their study design. The versions are: (1) the best-case scenario, i.e., the ideal protocol if yours is the only study being conducted and (2) a "minimum acceptable data yield" protocol, i.e., the least amount of data needed to answer the primary research question(s), without which it would not be worth participating. The former informs the scientific merit review by communicating the project team's understanding of the underlying science and mitigates unfair criticism of methodology by reviewers less familiar with applied and operational research, whereas the latter demonstrates focus on the primary scientific aims while respecting the agency's evaluation of operational feasibility and willingness to work as part of a larger integrated project team.

Researchers also can take actions to best navigate the extensive multiyear effort. When projects are funded, budgeting should be done carefully and research funds should be spent slowly to avoid a lack of funding later in the research process if the project is extended or data collection is delayed substantially. Patience and planning are mandatory to deal with the process. An understanding of what is needed, both generally for analog research as well as specifically for the analog, is essential. Reading chapters, such as this one, and reviews of analog research in specific research areas is helpful. When reviews of previous analog research on the topic of interest are not available, becoming familiar with other primary research that has been conducted in the analog environment of interest and personal correspondence with other investigators can be helpful for a better understanding of the context and procedures used.

The extreme length of analog research (sometimes 5 years or greater) necessitates maintaining detailed records and ensuring knowledge transfer across multiple parties that are involved over the course of a project. First, technology and software can be used to manage knowledge transfer. For example, team collaboration software such as Slack or Microsoft Teams can be used to organize workflows and keep multiple teams and team members in the communication loop for a given topic. They also provide a history of the interactions and decisions made by the research team. Another strategy is to have research assistants and other parties create an informal presentation in PowerPoint or Google Slides for bi-weekly updates and reporting. These updates can be used to ensure that progress is systematic and focused on research objectives, while also documenting information and decisions for future parties to refer to when joining the research team. In addition, documents can be stored on a cloud service such as Google Drive or Dropbox. A "readme" file or folder can be kept with the proposal, along with background documents, and a Gantt (or similar) chart that delineates key responsibilities over time. Researchers can use this folder as an "onboarding" for new research team members. It should be noted that cloud services may be appropriate for some documentation (e.g., nonproprietary protocols, information about the analogs), but not others (e.g., proprietary or confidential information such as data). Further, often there will be a long gap between methodological decisions and write up for publication, final reports, or submission of protocols and cleaned data to the funding agency. To help in the documentation of these decisions, a "working manuscript" can be assigned for each aspect of the research project. As an example, if interaction data are being coded, a "Method" section should be written in real-time as decisions are made. While this may be trimmed for some publications or reports, the working manuscript provides documentation or a reminder of the decisions that were made for current and future research team members.

What Is Needed

To the extent possible, the process should continue to be demystified and streamlined to improve researcher participation. Many of the solutions mentioned increasingly are being implemented by researchers, organizations, and agencies. In addition to these, three other things are needed to address the challenge of obtaining access in analogs. First, where possible, it is important to improve the predictability of when parts of the research process are likely to occur. As timelines change or are established, this should be quickly communicated to the researchers. Researchers should solicit this information regularly. Given that analog research is a multiyear process, it is likely that externally funded researchers will be engaged in other activities, whether research, teaching, or other commitments. Improving the predictability of timing can help researchers better manage multiple obligations, conduct higher quality research that is better integrated, and wisely spend public funds with which they and the sponsoring agency have been entrusted. Second, efficiencies should be made so that the process is less cumbersome. The removal of multiple scientific reviews is an example of this that is already being done. Finally, agencies and sponsors of analogs should provide orientation training to new researchers, so that the researchers and analog managers develop a shared understanding of the parties involved and the research process in the analog. For example, a new researcher may not understand the importance of science integration meetings or science requirements documents. A lack of understanding may result in researchers not prioritizing important meetings or deadlines, and inefficiencies such as multiple drafts of documents. This chapter, as well as orientation training to a specific analog and its process, may help researchers provide better and timelier information to agencies. Researchers should regularly ask for the purpose of different meetings and requirements, so they understand their role in preparing for and facilitating data collection.

CHALLENGE 3: DATA CHALLENGES

Analog environments share one or more of the contextual features of the extreme environment. These particulars often lead to significant data challenges. For example, teams of people wintering over in the Antarctic are subject to a hostile physical environment, including extreme temperature, and may not be accessible during the winter-over period. While these features help researchers understand individual and team function in isolation while surrounded by a hostile environment, the hard-to-access locations can make data collection difficult. More details on the challenge of collecting data in hard-to-access locations, a lack of criterion data, and small sample sizes are provided next.

First, difficult-to-access locations can lead to data-collection challenges. The researcher may never actually travel to the site (e.g., Antarctica) which may result in an inadequate understanding of the context and how best to frame the research and manage data (e.g., providing appropriate instructions for administering the protocol, identifying whether measures need to be translated). Second, remote locations may have inadequate power for equipment and confined spaces may limit the amount of equipment a researcher may provide. Third, by nature of the individuals and team living in an isolated environment, they are likely to have limited access to the Internet (availability, stability, or bandwidth). This means that data will often be collected by other means such as paper-and-pencil measures or interview, or, if the Internet is intermittently available, the need for paper-and-pencil measures as backup.

Second, because of the extreme features of the environment, data collection can often be limited in amount or kind. As an example, a lack of criterion data, such as performance measures is a common problem. Participants in natural analogs and operational environments are there for purposes other than the research. This has two implications: privacy concerns may significantly limit access to some types of data (e.g., mental health, performance), and the number of hours participants dedicate to research may be significantly limited. In controlled analogs, particularly for longer missions there may be more time to collect criterion measures. Even so, adequate criterion measures can be time-consuming to develop.

Third, the sample size is often small in analog research, particularly in higher fidelity settings. For example, to examine the effects of prolonged isolation, the Mars 500 experiment placed a crew of 6 people in isolation for 520 days. These chamber simulations are logistically intense to prepare for and operate, which makes a large individual and team sample size unrealistic. A recent quantitative review of team dynamics research conducted in analog environments found that 80% of articles included a sample size of fewer than five teams (Bell, Brown, & Mitchell, 2019).

Small sample sizes in analog research also have been noted for individual-level phenomena (Shea, Slack, Keeton, Palinkas, & Leveton, 2011). Small sample size is one of the greatest data collection challenges of analog and operational research.

Generalizability of research is made through a series of inferences, including statistical inferences. In human research with adequate sample sizes, the null hypothesis significance testing (NHST) framework is often used to provide evidence that an observed effect in the sample is present in the population of interest. A hypothesis is specified and then data are analyzed with inferential statistics such as *t* tests, analysis of variance (ANOVA), and multiple regression to infer whether observed effects (i.e., the difference between two means) is more likely than would be expected by chance. Small sample sizes typically have inadequate statistical power, thereby decreasing the likelihood that the results will be reproducible (Button, Ioannidis, Mokrysz, Nosek, & Flint, 2013).

Button et al. (2013) provide a detailed discussion and demonstration of the consequences of low power for replicability. As they indicate, low-powered studies are more likely to produce more false negatives; they have a lower chance of finding a true effect when one exists in the population. Lower power also lowers the positive predictive value (PPV), or the probability that a positive research finding reflects a true effect. A lower-powered study also makes it less likely that an observed effect that passes a significance threshold (e.g., $P < .05$) reflects a true effect. Finally, even when an underpowered study discovers a true effect, the estimate of the magnitude of the effect is likely to be exaggerated, especially when the effect is newly discovered, called the 'winner's circle' (Ioannidis, 2008). *For these reasons, inferential statistics and the NHST framework are largely inappropriate in small sample size research.* As such, other approaches must be used to provide evidence that an observed effect is likely to occur in the target population.

Solutions

Current solutions and best practices exist for each of these challenges. First, because the researcher is unlikely to travel to hard-to-access locations, it is critically important that the researcher identify a point-of-contact (POC) and establish good rapport and a shared understanding of the research project with the point of contact. The POC can significantly influence the success of a project; for example, the POC in natural analogs will often help with the recruitment of participants. The researcher should discuss the purpose of the project, the target population to be recruited, and next steps for individuals who show interest in participating. The same person is often the POC for several studies; thus it is important to provide clear and concise documentation that summarizes "what is this project about?", "who can participate?", and a few bulleted points on what administrative behaviors can significantly strengthen or weaken the data (e.g., biological samples must be frozen immediately, must use a quiet room for cognitive testing). The POC can ensure any compliance with the protocol and communicate anomalies back to the researchers.

Second, it is necessary to overcome inadequate criteria in good analog research. Potential solutions to the lack of criterion data involve a two-step process. As a first step, the researcher must adequately define criteria of interest. This is important for being able to judge the relevance of possible criteria in the next step. Generally, criteria of importance are aligned with organization goals, for example, identified risks in NASA's Human Research Roadmap. An example is the risk of performance and behavioral health decrements due to inadequate cooperation, coordination, communication, and psychosocial adaptation within a team (NASA Human Research Program's "Team" risk). These risks identify important criteria to the organization (in this case NASA), and should serve as the basis for criteria in research. Second, researchers can judge currently available information, or potential criteria they could develop for their usefulness. Usefulness of criteria is judged by its practicality, relevance, and sensitivity (Cascio & Aguinis, 2019). Practicality reflects the extent to which criterion data collection will interfere with ongoing operations. Relevance captures the extent to which a measure is logically related to the criterion domain of interest. Sensitivity is the extent to which the criteria can discriminate between individuals and teams that are doing well or not doing well.

To increase the likelihood of practicality in an analog setting, where possible, unobtrusive measures or already existing data should be used. As examples of unobtrusive data, audio and video recordings can be used to determine the presence and patterns of behaviors. Gushin et al. (2016) provide an example of capturing the well-being of crew members via content analyses of communications. Existing measures also should be used where possible, and existing measures or tasks can be modified or recoded to improve criterion relevance. As an example, the multimission, space-exploration vehicle, extravehicular activity (MMSEV-EVA) is a team performance simulation in NASA's HERA. Fuel consumption is an efficiency metric automatically generated by the program; however, the simulation provides a wealth of additional information about team performance. The first author's research team was able to code and score existing MMSEV-EVA data to provide a quantitative measurement of the extent to which the team met its objectives, which was a definition of team performance more relevant for their research questions. Without collecting additional criterion data, they were able to identify predictors of team performance over time.

In some cases no relevant criterion can be identified, which will require criterion development. Best practices in criterion development should be used to ensure criteria are relevant, reliable, and able to be combined. Importantly, when such criteria are used in an analog, they should go through extensive validation before use in that setting. The special population and analog context need to be considered, and criteria must be carefully crafted to ensure distinguishability between those high and low on the criterion of interest. For example, astronauts and astronaut-like individuals may have excellent cognitive function, requiring tests of cognitive function to be refined so they may detect meaningful decline for participants who typically score at the top end of the distribution. Participants in analog research may also engage in impression management for several reasons, such as the desire to become an astronaut, or because they are working in their professional operational environment (e.g., Antarctic stations). Good criteria must be able to distinguish between desirable levels on a criterion and impression management.

Finally, there are several potential solutions to small sample research (Bell et al., 2018). Within a given study, strategies can be used to increase the statistical power. For example, when appropriate for answering the research question, the sample size can be increased by collecting data at a lower level (e.g., interaction, instead of individual, level data), or collecting repeated measures over time. Further, alternative research and data analytic approaches better suited to small sample research can be used. For example, computer simulations, such as agent-based modeling, can be used to conduct virtual experiments and alternative scenarios to those present in the analog. Bayesian analyses are particularly useful for small sample size research if the data are normally distributed (see Bell et al., 2018; McNeish, 2016; Van De Schoot, Broere, Perryck, Zondervan-Zwijnenburg, & Van Loey, 2015). Descriptive statistics and effect sizes also may effectively communicate trends. Finally, it is often prudent to conduct debrief interviews to contextualize the data for better interpretation of small sample research and to explore new issues.

Further, because any one small sample study is unlikely to provide enough evidence for firm conclusions, small sample analog research needs to be deliberately conducted with eventual data aggregation in mind. Bell et al. (2016, 2019) provide guidance for doing so as summarized next, including collecting (a) enough data so that an effect size can be generated for eventual comparison in a meta-analysis and (b) raw data in such a way that it can be analyzed and summarized across multiple studies.

First, when possible, data should be collected in such a way that an effect size estimate can be generated. Effect sizes standardize (e.g., *d, r*) the strength and direction of a relationship and allow for comparisons across studies. Meta-analysis, the preferred means for summarizing quantitative data across studies and generating cumulative knowledge, requires that an effect size is generated from each study. Second, when possible, data should be collected using the same measures and response formats that have been used in previous analog studies. For example, an International Standardization of Best Rest Studies Measures (Sundblad & Orlov, 2015) has been adopted that allows for researchers in different analogs to collect comparable data on frequently examined domains. Comparable data across analogs allow researchers to benchmark new results as well as conduct analyses with larger data sets. Further, because studies will often include too small of a sample size to generate an effect size estimate, traditional meta-analysis may not always be possible to aggregate analog research. The use of the same measures allows for systematic reviews of the research through other means. As an example, in their review of long-distance space exploration mission-analog research on teams, comparable measures across analogs allowed Bell et al. (2019) to generate descriptive figures reflecting changes over time as well as upper and lower limits in team cohesion, team efficiency, team conflict, communication with mission control, and team mood across multiple analog studies. The ability to aggregate data across studies becomes significantly limited when different measures or different response formats are used (e.g., some scales have a neutral, some do not). *Thus, a critical step in analog protocol development is to identify measures of constructs that have been previously validated in analog research, and to use these measures in the protocol when they will adequately capture the phenomena of interest.*

What Is Needed?

While many solutions exist, continued work in this area is needed. First, agencies and researchers need to make the reproducibility and replication of research a priority and adhere to best practices. Researchers and other stakeholders (e.g., funding agencies, operations, journal editors) need to ensure that appropriate inferences are being made from small sample research. Research should be judged on quality and appropriateness of the inferences, rather than the presence or absence of an effect. Second, continued understanding of the appropriate analysis and inferences of small sample size research is needed (Bell et al., 2018). For example, future research could explore the accuracy of different meta-analytic approaches for use with extremely small sample sizes (Bell et al., 2019), and alternative strategies for best representing the available data can be developed. Finally, as noted, consistent measures should be used across analogs. International agreements, such as was done with the International Standardization of Best Rest Studies Measures, can be done in additional individual and team research domains to facilitate knowledge accumulation. Analog research that operates outside of these agreements, for example, that are run by private organizations, should still adhere to these standards when possible, so that their research can be compared with, and contribute to, an emerging body of analog research. Note, the use of these standardized measures should not preclude innovation or the collection of data on constructs ought not to be captured by the standardized measures. However where possible, the use of measures and response formats that allow for comparisons across multiple studies and analogs is essential.

CHALLENGE 4: DISSEMINATING FINDINGS AND MAINTAINING PARTICIPANT CONFIDENTIALITY

The distinctive features of the analog research setting, including the high-profile nature of analog environments, public and media interest, small sample sizes, and crew members that are publicly identified, require special consideration to how findings are disseminated. Research should be disseminated through traditional outlets, such as publications and presentations; however, researchers are often provided opportunities for additional outlets likely to reach the broader public, such as education outreach, press, broadcast media, and creative works like documentaries. Researchers participating in these public information forums play an important role in engaging and fostering the public's interest in space exploration, science, and technology. Social media platforms also are increasingly used to disseminate research information or used to prime the public before more comprehensive dissemination in long form media.

There are important considerations for the content and timing of what should be presented, published in articles, or discussed in public forums. Importantly, confidentiality of a participant's information or data should be maintained throughout the process. Researchers cannot disclose human participant information that violates research ethics guidelines and regulations (e.g., HIPAA) set out by the study Institutional Review Boards even after a study ends. While most researchers are likely familiar with the ethical mandate of maintaining confidentiality, small sample sizes require extra thoughtful consideration of what is discussed or presented publicly. Small sample sizes may inherently make crews or participants identifiable when they are unique in their standing on a particular construct and this information is presented or revealed (e.g., mission length, gender composition, role). Confidentiality of participant data is put further at risk because crew member identities are often made available online through the press or other coverage, or even publicly self-disclosed by participants. This risk is further compounded by the often inherent familiarity participants may have with each other's data, and the potential for violating others' rights and the researcher's responsibility to protect confidentiality.

Research protocols can be ongoing for several missions, requiring that elements of the protocol, measures, and methods be kept confidential so that the research integrity can be maintained in a future mission, or even for operational use. For example, key manipulations for which data are still being collected should not be revealed to the press. Future participants are often interested in space and may encounter the information in the media, leading to the potential for confounded results when they participate, or potential loss of suitable volunteers in response to misinformation. As with participant confidentiality, researchers must strike a careful balance between describing their study and depicting results transparently while not breaching the confidentiality of protocols and manipulations that are still being implemented.

Solutions

Maintaining confidentiality is paramount during the dissemination process. Methods, descriptions, and results sections including tables and figures should be carefully inspected for the extent to which that information could be combined with publicly available information to match an individual to their results. For example, it may be inappropriate to present data in a figure that shows mission length and crew role information (e.g., commander) in datasets that include missions of uniquely different lengths. Codes rather than crew or position information can often be used to mask participant identity and still adequately represent the data (e.g., position A). These codes should not be the participant IDs assigned by researchers as the first level of participant identity protection. In media interactions, descriptions of analog results whether empirical or anecdotal should be careful not to reveal an individual participant's data.

Principal investigators and their research teams can be proactive in guiding how the research team will interact with the media. This can be as simple as specifying who can communicate with the media about the research, and what content and key talking points are appropriate, or a more formal coordinated public outreach plan across researchers made in conjunction with their public relations department. For example, researchers can contribute to an agency press release when there is a new finding that is released to multiple, reputable news agencies. Websites that define the project in general terms with findings added as data collection for key manipulations can be created. This information can serve as the official public statements on the project to which initial inquiries may be referred.

Researchers may be asked for interviews, either as a follow up to publications, press releases, or other information. For these interviews, the research team should identify key talking points ahead of time. These talking points should succinctly summarize key findings from the research and their implications for spaceflight. The sponsoring organization, such as a space agency, also may have talking points when new larger initiatives are implemented (e.g., return to the moon; Artemis). Preparing talking points ahead of time can help the researcher stay focused and avoid misstatements including breeches of confidential information during an interview. The press is often looking for 'soundbites' for traditional and social media platforms, which talking points can lend themselves to nicely when properly prepared. Further, the researcher should not necessarily be limited to key findings and should feel comfortable speaking from their expertise. However, it is usually prudent for a researcher to articulate when they are speaking from general expertise or reporting on others' research projects, as the press may report all comments as the results of the researcher's study. Formal development of science communication skills for public outreach is rarely offered as part of professional science training, but science communication has become a field unto itself, and organizations are increasingly investing in such training (www.aldacenter.org).

What Is Needed?

While researchers are often well versed in technical writing for publication, they are less likely to be familiar with engaging the media. Media training (or coaching) from the researcher's home institution is highly recommended for researchers who may eventually engage in many forms of public outreach that require specific interview techniques or key messaging formats. An analog itself will have an identity and operational context that it wants to present to the research community and the public, and researchers and their personnel should be provided with messaging guidelines from the analog so that the analog is accurately represented in the media. Just as important is the development of a standardized guide for ethical considerations in disseminating analog research results. Some aspects of media inclusion and public outreach will vary according to the specific analog; however, much of what constitutes appropriate disclosure of protocols, other participants' experiences, and roles of the participants is the same across analogs. For example, we cannot think of a time when revealing details to the press about a specific participant's experience, data, or the details of a manipulation or intervention from an ongoing protocol would be appropriate. Because of these concerns and their ethical implications, it is worthwhile that a standardized training related to these privacy considerations is created for distribution to all participants and key personnel including researchers.

CONCLUSION

Analog research plays an important role in preparing for future space exploration missions. It allows for the characterization of the risks of spaceflight, allows problems to be solved, and countermeasures to be developed and refined before deployment to an operational environment. Despite its importance, analog research has significant challenges. We focus on a few key issues related to identifying and selecting analogs; obtaining access to analogs; data collection challenges including the logistics of data collection in hard-to-access locations, lack of criterion data, and small sample sizes; and the challenges to protecting subject identity in research dissemination and when interacting with public forums. We highlight key challenges and current solutions, also where more work is needed so that analog research can best inform our understanding of human behavior and performance at the final frontier.

ACKNOWLEDGMENTS

Work related to this project conducted by STB was supported by National Aeronautics and Space Administration grants NNX16AQ48G, NNX15AM32G, NNJ15HK18P, and NNJ13HF07P. STB, PGR, and BJC were supported by KBR's Human Health and Performance Contract NNJ15HK11B through the National Aeronautics and Space Administration.

Author contributions: STB and PGR conceived the project. STB, PGR, and BJC wrote the chapter.

Any opinions, findings, conclusions, or recommendations expressed are those of the authors and do not necessarily reflect the views of any organization affiliated with the research. The authors declare that there are no conflict(s) of interest.

REFERENCES

Bell, S. T., Brown, S. G., & Mitchell, T. D. (2016). *Data mining review of team benchmark studies related to long duration exploration missions* (NASA/TM-2016-219280). Houston, TX: NASA/Johnson Space Center.

Bell, S. T., Brown, S. G., & Mitchell, T. (2019). What we know about team dynamics for long-distance space exploration: A systematic review of analog research. *Frontiers in Psychology*, *10*, 811. doi:10.3389/fpsyg.2019.00811.

Bell, S. T., Brown, S., Outland, N., & Abben, D. (2015). *Critical team composition issues for long-distance and long-duration space exploration: A literature review, an operational assessment, and recommendations for practice and research* (NASA/TM-2015-218568). Houston, TX: NASA Johnson Space Center.

Bell, S. T., Fisher, D. M., Brown, S. G., & Mann, K. E. (2018). An approach for conducting actionable research with extreme teams. *Journal of Management*, *44*(7), 2740–2765.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., & Flint, J. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews*, *14*(5), 365–376.

Cascio, W., & Aguinis, H. (2019). *Applied psychology in talent management* (8th ed.). Thousand Oaks, CA: Sage.

Gushin, V. I., Yusupova, A. K., Shved, D. M., Shueva, L. V., Vinokhodova, A. G., & Bubeev, Y. A. (2016). The evolution of methodological approaches to the psychological analysis of the crew communications with Mission Control Center. *REACH, 1,* 74–83.

Ioannidis, J. P. A. (2008). Why most discovered true associations are inflated. *Epidemiology*, *19,* 640–648.

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review, 31,* 386–408.

Keeton, K. E., Whitmire, A., Feiveson, A. H., Ploutz-Snyder, R., Leveton, L. B., & Shea, C. (2011). *Analog assessment tool report* (NASA/TP-2011-216146). Houston, TX: NASA Johnson Space Center.

Landon, L. B., Slack, K. J., & Barrett, J. D. (2018). Teamwork and collaboration in long-duration space missions: Going to extremes. *American Psychologist*, *73*(4), 563–575.

McNeish, D. (2016). On using Bayesian methods to address small sample problems. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*, 750–773.

Shea, C., Slack, K. J., Keeton, K. E., Palinkas, L. A., & Leveton, L. B. (2011). *Antarctica meta-analysis: Psychosocial factors related to long-duration isolation and confinement*. (NASA/TM-2011-216148). Houston, TX: National Aeronautics and Space Administration/ Johnson Space Center.

Sundblad, P., & Orlov, O. (Eds.) (2015). *Guidelines for standardization of bed rest studies in the spaceflight context*. Paris: International Academy of Astronautics. Available for download at: http://www.nasa.gov/hrp/important_documents.

Van De Schoot, R., Broere, J. J., Perryck, K. H., Zondervan-Zwijnenburg, M., & Van Loey, N. E. (2015). Analyzing small data sets using Bayesian estimation: The case of posttraumatic stress symptoms following mechanical ventilation in burn survivors. *European Journal of Psychotraumatology*, *6*(1), 25216.