Gesture-based Interaction in Extreme Environments: Findings from a Mars-Analog Mission with a Finger-Worn Device

Jean Vanderdonckt
LouRIM & ICTeam Institutes
Université catholique de Louvain
Louvain-la-Neuve, Belgium
jean.vanderdonckt@uclouvain.be

Radu-Daniel Vatavu
MintViz Lab, MANSiD Center
Ştefan cel Mare University of Suceava
Suceava, Romania
radu.vatavu@usm.ro

Nuwan Attygalle
LouRIM Institute
Université catholique de Louvain
Louvain-la-Neuve, Belgium
nuwan.attygalle@uclouvain.be

Loriane Baes
MARS–UClouvain Space Mission
Louvain-la-Neuve, Belgium
baesloriane@gmail.com

Louis Baltus
MARS–UClouvain Space Mission
Louvain-la-Neuve, Belgium
louis.baltus@student.uclouvain.be

Arnaud de Wergifosse
MARS–UClouvain Space Mission
Louvain-la-Neuve, Belgium
dewergifosse@marsuclouvain.be

Hippolyte Hilgers
Université catholique de Louvain
Louvain-la-Neuve, Belgium
hippolyte.hilgers@uclouvain.be

Romain Maddox
MARS–UClouvain Space Mission
Louvain-la-Neuve, Belgium
rmaddox@marsuclouvain.be

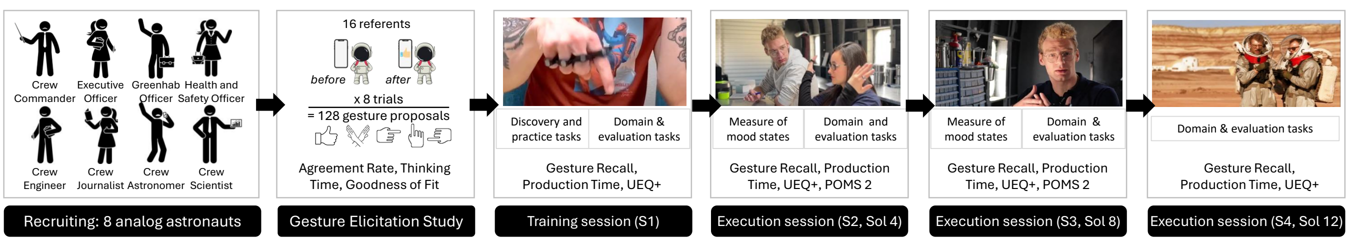

Figure 1: Overview of our experiment protocol for understanding user preferences, performance, and experience of gesture interaction from Earth (first three blocks) to a Mars-analog environment (last three blocks) at the Mars Desert Research Station.

Abstract

Extreme environments, such as analog space stations, can impact the dexterity with which astronauts perform gestures for communication within the crew and for controlling interactive systems. To investigate this impact, we conducted a gesture elicitation study with eight analog astronauts, followed by several user performance and user experience evaluation sessions at the Mars Desert Research Station. Our findings reveal high gesture recall accuracy, short production times, and an overall favorable user experience across thirteen UEQ+ dimensions, from perceived efficiency to trust, which we interpret alongside the crew's psychological mood during the two-week mission. Based on our findings, we derive design implications for gesture interaction in extreme environments.

CCS Concepts

• **Human-centered computing** → *Field studies*; **Gestural input**;
• **Applied computing** → **Aerospace**.

ACM ISBN 979-8-4007-2812-9/2026/10
https://doi.org/10.1145/3822518.3830049

Keywords

Gesture interaction, Mars, Gesture elicitation, Space interaction, Wearable devices, Mars Desert Research Station

ACM Reference Format:
Jean Vanderdonckt, Radu-Daniel Vatavu, Nuwan Attygalle, Loriane Baes, Louis Baltus, Arnaud de Wergifosse, Hippolyte Hilgers, and Romain Maddox. 2026. Gesture-based Interaction in Extreme Environments: Findings from a Mars-Analog Mission with a Finger-Worn Device. In *ACM Symposium on Spatial User Interaction (SUI '26), October 10–11, 2026, Bari, Italy*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3822518.3830049

1 Introduction

Finger gesture interaction refers to translating finger movements, captured by various sensors into commands for controlling or operating an interactive system via natural and intuitive mappings [29]. However, an extreme environment [15], such as that of an analog space station [9,87], can negatively affect the dexterity with which these gestures are performed [77]. In a broader context, extreme environments enable extreme interactions [71], which occur in physical or digital contexts that are significantly different from typical settings pose severe challenges to users, and require rethinking traditional interaction design [67]. Examples include astronaut-oriented design [5,6] and design for space [22,33,54], which support users' cognitive and physical performance under unusual conditions [30], such as microgravity [43] and prolonged isolation [15].

Extreme constraints can be psychological [35,44], cognitive [24, 51], physical [43,77], medical [21,53], sensorimotor [77], social [44], and technological [5,14]. The success of communication in extreme environments depends on modalities robust to such constraints. For example, speech-based communication can be degraded by radio interference, background noise from ventilation systems and equipment from the habitat, and other environmental factors that must be considered during system design [74], while microgravity and spaceflight conditions may alter speech production and reduce recognition accuracy [83]. Gesture-based communication, however, can prove vital when speech or other channels fail [62].

However, research on gesture interaction in extreme settings has been limited, primarily because of the challenges of conducting human-computer interaction experiments [72] in enclosed spaces and habitats [4], small participant samples [23], and requirements for wearing protective gear that restricts body movement [70]. Moreover, current gesture sensing technologies, such as wearables for the fingers and hand [59], may not be optimized for such extreme settings, which could negatively affect user performance, user experience [69], and, ultimately, mission success [5]. In this paper, we investigate the impact of an extreme environment, during an analog space mission at the Mars Desert Research Station (MDRS),[1] on the design of finger gesture-based interaction using a wearable device, and make the following contributions:

- A gesture elicitation study with eight analog astronauts, who proposed finger gestures for sixteen referents using the Tap Strap 2,[2] a wearable finger-worn device (Section 3).
- An experiment conducted in the extreme environment of the MDRS, where the analog astronauts performed the elicited gestures across sessions with increasing physical and psychological constraints. We report combined results of user performance, user experience, and mood states (Section 4).
- Lastly, we derive design implications for gesture interaction in extreme environments via wearable devices (Section 5).

Although limited by the small sample size imposed by accommodation constraints at the MDRS, we report the first gesture elicitation study with analog astronauts in an extreme environment to advance gesture interaction research for future space missions.

2 Related Work

Our work focuses on gesture interaction in extreme environments, at the intersection of finger-worn devices and human-computer interaction (HCI) research on extreme environments.

2.1 Wearables and Gesture-based Interaction

Prior research on gesture interaction using wearables has focused primarily on balancing low-power sensing [20] with the computational cost of running machine learning algorithms for gesture recognition [10,78,85], while also introducing novel device form factors [28,80]. For example, Wen *et al.* [78] showed that fine-motor finger gestures can be recognized using sensors integrated into a smartwatch with an F1-score of 87%. PPG sensors have also been used to recognize finger gestures, as in Zhao *et al.* [85], who reported over 88% accuracy from training on a set of 3,600 gestures collected from ten users. AO-Finger [82] is a fine-grained finger gesture recognizer based on acoustic-optic sensor fusing, achieving 95% accuracy. More recently, Cornella-Barba *et al.* [10] reported 98% accuracy for twenty-six American Sign Language symbols. Despite these technological advances, major challenges remain regarding limited battery life, sub-optimal recognition accuracy and user comfort, and limited context-aware adaptability [28]. A further limitation is that most systems employ designer-defined gesture sets, which prioritize recognition accuracy over user preferences.

2.2 Human-Computer Interaction Research for Extreme and Space Environments

An extreme environment impacts organizational, interpersonal, and individual working and living systems [68]. Thus, many adaptability attributes for operating effectively in such environments have been proposed [4], such as intelligence, emotional stability, self-control, openness, optimism, mastery, introversion, task-oriented coping, and adequate sleep. These influence crew performance [3] as the human cognitive system tries to adapt to extreme constraints [45].

SpaceCHI [6,46,47,86] explores HCI research for space as a representative example of an extreme environment, by "designing novel interactive systems and mobile interfaces to support human living and working in space and other regions of the solar system" [46, p.1]. In this context, it encompasses participatory design approaches [40], experience design [41], and human factors [12,13]. Thus, HCI research for extreme environments [12] serves a variety of purposes, including designing for software efficiency on Mars [14], medical support [53], decision support systems [58] and planning [52]. In extreme environments, gesture-based interaction may be limited or even impractical [39,40]; for example, graviception interferes with manual dexterity, making gesture movement difficult [79]. Furthermore, microgravity creates a haptic illusion [43]: whereas Earth gravity induces the perception that upward forces are larger than downward ones, this effect disappears in microgravity and is increased in hypergravity environments. Since astronauts wear gloves that prevent direct physical contact [65], they exhibit a "need for touch." We refer to Vatavu *et al.* [72], who highlighted methodological challenges in conducting HCI scientific experiments in extreme environments, such as the small samples, a constraint also encountered in our experiment described in Sections 3 and 4.

2.3 Gesture-based Interaction in Space

Gesture-based interaction in extreme environments, such as in analog or space missions, presents both benefits and challenges. For example, in microgravity [43] and high-risk failure contexts [42], gesture input reduces reliance on physical controls and enables hands-free or contactless interaction. However, bulky spacesuits [70] and pressurized gloves [57], resulting in limited finger dexterity, constrain gesture articulation, therefore affecting recognition accuracy. Furthermore, environmental factors, such as radiation, vibration, and strict power constraints affect negatively the reliability of wearable sensors and onboard processing. User fatigue and cognitive load are also amplified in such environments, where astronauts

[1] https://mdrs.marssociety.org
[2] https://shop.tapwithus.com/products/tapwithus-tap-strap-2-wearable-keyboard-mouse-air-gesture-controller

typically operate under a variety of physical and psychological stressors. As a result, gesture interaction for space missions requires adaptive, context-aware design balancing expressiveness with safety, reliability, and physical effort.

Few papers have directly addressed finger gesture-based interaction for gloved hands in isolated, confined, and extreme settings. For example, Ceruti *et al.* [7] developed wireless data gloves with magnetic sensors for gesture recognition in extreme military, space, and hazardous environments; Zhang [84] proposed 3D-printed self-powered sensing gloves for finger gestures in underwater and smoke environments; and Tseng *et al.* [66] examined VR for confined spaces via small-scale finger motions. Regarding gesture recognition technology, prior work has addressed system design for astronauts, but did not explicitly report on what type of gestures are recommendable for optimal user performance and user experience. For example, Pavithran *et al.* [48] developed electromagnetic wave-based recognition with 80% accuracy on four hand gestures for spacewalk communication; Liu *et al.* [27] created a data glove system for recognizing sixteen American Sign Language gestures to control a robot during extravehicular activities (EVA); Gu *et al.* [26] introduced an astronaut gesture dataset and a hierarchical attention network for human-robot interaction at space stations; and Hu *et al.* [19] developed vision-based recognition for astronaut training systems.

Although several Gesture Elicitation Studies (GESs) [81] have explored finger gestures, none has targeted astronauts or users in extreme environments, according to recent literature reviews [8,76]. In Section 3, we present the first study of this kind, to the best of our knowledge, followed by evaluations of the consensus gesture set at the MDRS, described in Section 4.

3 A Gesture Elicitation Study on Earth

We conducted a gesture elicitation study in a controlled, typical Earth-like environment, during which analog astronauts were asked to propose finger gestures with a finger-worn device for use during missions to a Mars-analog space station.

3.1 Participants

We involved a crew of eight analog astronauts,[3] representing the maximal accommodation allowed at the MDRS facility. Participants were recruited through an open call advertising a two-week mission to a Mars-analog environment. The selected crew comprised five men and three women, aged between 22 and 31 years (*M*=25.6, *SD*=2.6), with no physical impairments, and different educational backgrounds in astronomy, biology, engineering, geography, neuroscience, and psychology. All participants were right-handed, frequent users of computers and smartphones, and had no prior experience with the gesture sensing wearable used in the study. Given the small sample size, which is nevertheless representative of HCI studies conducted in extreme environments [72], we present participants' individual demographic details in Table 1.

[3]An analog astronaut is a trained individual who assumes the role of an astronaut during a simulated crewed mission.

Table 1: Demographic details of crew members.

ID	Gender, age (years)	Academic Background	Crew role during the mission
CM1	Male, 27	Physics	Commander
CM2	Female, 25	Psychology and Occupational Sciences	Executive officer
CM3	Male, 24	Bioengineering	Engineer
CM4	Female, 25	Pharmacy	Health/Safety officer
CM5	Male, 22	Electrical Engineering	Greenhab officer
CM6	Male, 24	Geography	Astronomer
CM7	Male, 31	Neuroscience and Physiotherapy	Scientist
CM8	Female, 27	Biotechnology	Journalist

3.2 Apparatus

We selected the Tap Strap 2[4] for our study, a wearable device consisting of five interconnected electronic rings, one for each finger, with integrated accelerometers for detecting mid-air movements and on-surface taps. A key reason for selecting this device is that it can be worn underneath astronaut gloves, making it suitable for space operations. Unlike other sensing technology for finger-augmentation devices [59], such as imaging-based gesture sensing that can be affected by camera placement, lighting conditions, and obstructed views, or biometric sensing, such as electromyography that may be affected by extreme bodily responses to psychological [35,44] and physiological stressors [43,77] in extreme settings, the Tap Strap 2 relies on body-worn inertial sensing, making it well suited to confined and extreme operational environments. Moreover, accelerometer-based gesture recognition has been extensively studied, and robust performance reported [10,37].

3.3 Task and Design

We selected 16 specific referents relevant for the analog Mars mission (Figure 2, bottom left), for implementing gesture-based interaction and communication [62], with the following categories: (i) communication [25] (*e.g.*, communicating with mission control and crew members), (ii) monitoring [38] (*e.g.*, constantly keeping an eye on the life support systems), and (iii) troubleshooting [52]. Referents were presented one at a time, in a random order per participant, via before/after visual representations using PowerPoint slides (see Figure 2, top for examples). The study followed a within-subjects design with three dependent variables:

- Agreement-Rate (AR), expressing the level of agreement among participants' gesture proposals for each referent, calculated using the AGATe [73] tool.
- Thinking-Time, defined as the time elapsed between the presentation of a referent and the participant's demonstration of their preferred gesture, measured in seconds.
- Goodness-of-Fit, representing participants' ratings, from 1 (low) to 10 (high), of how well they considered their proposed gestures to fit the corresponding referents.

[4]https://www.tapwithus.com/product/tap-strap-2

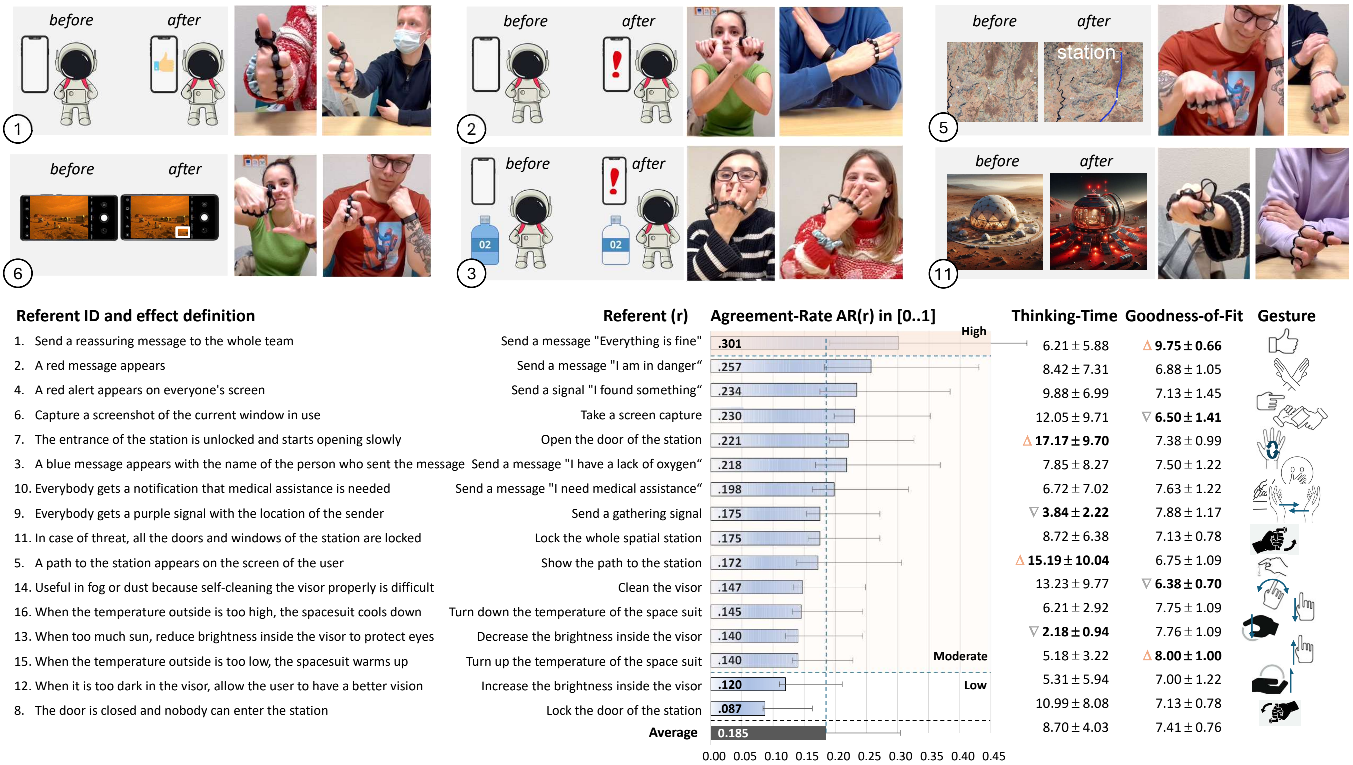

Figure 2: Examples of gesture proposals (top), the set of referents and the consensus gestures (bottom). Error bars show 95% CIs.

3.4 Results

Overall, we collected 8 (analog astronauts) × 16 (referents) = 128 gesture proposals for our analysis.

Agreement-Rate values were moderate (M=.185) with little variation (SD=.054), ranging between .087 (low agreement) for "Lock the door of the station" and .301 (high agreement) for "Send the message 'Everything is fine'." Overall, 13 of the 16 agreement rates (81%) were classified as moderate, two as low, and one as high, according to the interpretation guidelines from [73]. A few exceptions aside, the agreement rates exceeded those typically reported in gesture elicitation studies [73,76], indicating strong consensus among the crew. Figure 2, bottom right shows the consensus gesture set in descending order of Agreement-Rate. According to the taxonomy of Aigner *et al.* [2], five gestures in this set are iconic-dynamic (*e.g.*, "Lock the whole spatial station"), four are iconic-static (*e.g.*, "Take a screen capture"), three are semaphoric-dynamic (*e.g.*, "Send a gathering signal"), two are semaphoric stroke-like, one is semaphoric-static, and one is pointing.

Goodness-of-Fit ratings were high (M=7.41, SD=0.76 on the 1-10 scale), indicating that participants were generally satisfied with their proposed gestures. In particular, participants gave the highest ratings to the "Everything is fine" (9.75 ±0.66) and "Turn up the temperature of the space suit" (8.00 ±1.00) referents, and the lowest ones for "Clean the visor" (6.38 ±0.70) and "Take a screen capture" (6.50 ±1.41), respectively.

The Thinking-Time analysis revealed that participants proposed gestures quickly (M=8.70s, SD=4.03s), ranging from 3.84s (±2.22s) for "Send a gathering signal" to 17.17s (±9.70s) for "Open the door of the station." Furthermore, we found that Thinking-Time was negatively correlated with Goodness-of-Fit (Pearson's $r_{(8)}=-.53$): referents with longer thinking time tended to receive lower fit ratings, suggesting that participants were less satisfied with the gestures when they took longer to propose them.

3.5 Outcomes for the Mars-Analog Environment

To reduce the cognitive effort associated with memorizing and recalling gestures, which would add to other existing challenges in extreme environments [24,51], we retained only eight referent-gesture pairs for the analog mission at the MDRS, as follows:

- Thumbs up for "Send the message 'Everything is fine' " (6/8 crew members, representing a max-consensus rate[5] of 75%).
- Make a cross with both arms across the chest for "Send the message 'I am in danger' " (2/8, 25%).
- Raise the thumb and index finger, followed by the index finger bending twice to simulate taking a picture for the referent "Take a screen capture" (6/8, 75%).
- Draw a circle with the index finger and then point to its center for "Show the path to the station" (3/8, 37.5%).
- Bring arms in front of the body twice for "Send the message 'I have a lack of oxygen' " (3/8, 37.5%).
- Point with the index finger at the other arm for "Send the message 'I need medical assistance"' (2/8, 25%).

[5]Following Morris [36], the max-consensus rate is the percent of participants suggesting the most popular proposed interaction for a given referent.

- Tap with the open hand on the thigh for "Lock down the entire station" (2/8).
- Repeatedly move the open hand in front of the head for "Clean the visor" (6/8, 75%).

4 Experiment at the Mars Desert Research Station

We carried out an evaluation at the MDRS to understand the user performance and user experience of the identified gesture set.

4.1 Location

The experiment was conducted during a two-week mission at the MDRS, Hanksville, UT, USA. Designed based on NASA's Design Reference Mission 3, the MDRS is composed of several modules, including a three-floor cylindrical habitat, an Extra-Vehicular Activity (EVA) preparation room, and a science dome with a laboratory [61] (Figure 3a), all providing a Mars-analog environment to enable scientists to experience extreme conditions [64]. A two-week interval represents a reference duration for analog Mars missions [50].

4.2 Tasks and Protocol

Three tasks were carried out throughout the experiment (Figure 1):

- A *discovery and practice task*, in which participants received a tutorial on how to use the Tap Strap 2 device (5 minutes) and freely practiced the domain task (10 minutes).
- A *domain task*, in which participants performed the gestures.
- An *evaluation task*, in which the participants were administered the UEQ+ questionnaire.

The discovery and practice task was performed once, three weeks before the mission, representing the (Earth) baseline condition. The domain and evaluation tasks were performed repeatedly across four sessions (Figures 1 and 3): a *training session* (S1) took place one week before the mission in the tutorial room; an *execution session* (S2) after four Mars days (*i.e.*, Sol 4) in the main habitat at the MDRS (Figures 3b,c); a second *execution session* (S3) after eight Mars days (Sol 8) in the science laboratory (Figure 3d); and a final *execution session* (S4) after twelve Mars days (Sol 12), conducted outside the station (Figures 3e, 3f), with participants wearing analog spacesuits (suit, helmet, backpack with life support system, and 3mm-thick gloves; total weight approximately 12kg). During S4, crew members walked outside, drove rovers, and performed the domain and evaluation tasks over approximately two hours.

4.3 Measures

4.3.1 User Performance.

We evaluated the user performance of gesture production using the following two measures:

- Recall-Accuracy, a categorical variable adapted from the Nielsen-Norman success rate, as follows: 1=*first trial* (participants succeed in producing the correct gesture for the presented referent, getting it right from the first time), 2=*second trial* (participants succeed during the second attempt, achieving success with only a minor issue), 3=*third trial* (participants succeed during the third attempt, achieving success with one major issue), 4=*incorrect gesture* (participants produce the wrong gesture in response to the referent), and 5=*forgotten gesture* (participants could not recall the corresponding gesture for the presented referent).
- Production-Time, defined as the time required by the participant to perform gestures corresponding to the referents, measured in seconds with a stopwatch.

Since these measures focus exclusively on gesture execution, automatic gesture recognition was not required in our evaluations, although possible with existing implementations [10,37].

4.3.2 User Experience.

We used UEQ+ [55], a modular extension of the User Experience Questionnaire (UEQ) [56], which consists of several scales and is supported by analysis instruments and published norms [34] for interpreting results [17]. We administered the following scales: *Attractiveness*, *Efficiency*, *Perspicuity*, *Dependability*, *Stimulation*, *Novelty*, *Adaptability*, *Trust*, *Usefulness*, *Value*, *Visual Aesthetics*, *Intuitive Use*, and *Trustworthiness of Content*, from which we computed the following aggregate measures:

- Scale-Mean-Score (SMS), representing the average score for a given user experience scale across all crew members, ranging from −3 (negative) to +3 (positive).
- Scale-Mean-Importance (SMI), measuring the average perceived importance of the scale, from −3 to +3.
- Scale-Mean-Ratio (SMR), defined as the ratio of each S2-S4 sessions' SMS to the baseline SMS in the control session S1.
- Importance-Mean-Ratio (IMR), defined as the ratio of each S2-S4 sessions' SMI to the baseline SMI in the S1 control.

4.3.3 Psychological Evaluation.

To evaluate the psychological state of the crew during the two-week mission, we administered the Profile of Mood States (POMS 2) [31,32] questionnaire, a 65-item adjective checklist eliciting responses to "How you have been feeling during the past days, including today." POMS 2 allows for the rapid assessment of transient and fluctuating feelings [16, p. 1], and takes approximately 3 to 7 minutes to complete in typical conditions. We administered this assessment on sessions S2 and S3 only to capture changes in the crew's mood under comparable indoor conditions, whereas S4 was outdoor with different operational demands and physical exertion; in this way, S2 served as a baseline (pre) assessment for S3 (post), conveniently positioned at the midpoint of the two-week mission. Based on POMS, we measured the following:

- POMS-Score, a 5-point Likert rating, from 0 (not at all) to 1 (a little), 2 (moderately), 3 (quite a bit), and 4 (extremely).
- POMS-Scale-Mean, representing the average rating across all crew members, for the seven scales of POMS 2.
- Total-Mood-Disturbance as an arithmetic combination of the POMS 2 scales: (*Tension* + *Depression* + *Anger* + *Fatigue* + *Confusion*) - *Vigor*. TMD gauges the overall mental state of crew members, with lower scores indicating more positive mood states, while higher scores signify greater disturbance. With this measure, we are looking for patterns like the "Iceberg Profile" with high vigor and low negatives.

4.4 Results

4.4.1 User Performance.

Figure 4 shows the distribution of Recall-Accuracy levels across the eight referents and four sessions. Performance improved from S1 to S2 and S3, and ended with perfect recall in S4. It took four sessions for the crew members in our study

Figure 3: The Mars Desert Research Station (a), where the last three execution sessions (S2-S4) of our experiment were conducted: the main habitat for S2 (b,c), the science laboratory dome for S3 (d), and the extra-vehicular activity for S4, where analog astronauts were driving outside the station in a rover (e) while wearing full analog astronaut equipment with gloves (f).

to achieve this level of performance. Differences among referents were also observed due to some of the corresponding gestures being easier to recall, *e.g.*, for "Everything is fine," while others were more complex, *e.g.*, for "Show the path to the station." We also observed that most participants, regardless of their prior experience with wearables, were able to use the Tap Strap 2 effectively despite the restrictive physical conditions of the MDRS environment.

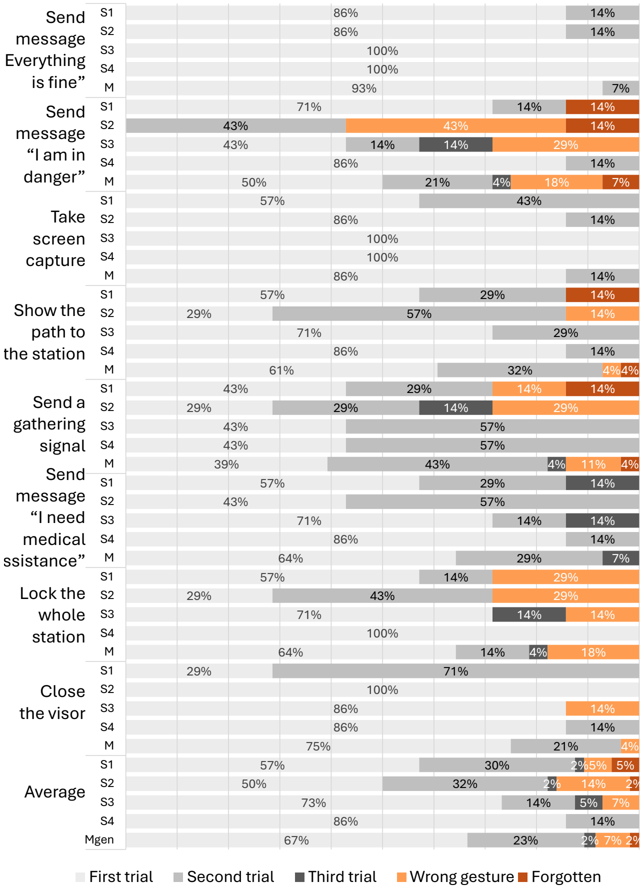

Figure 4: Distribution of RECALL-ACCURACY-LEVEL across all referents and sessions.

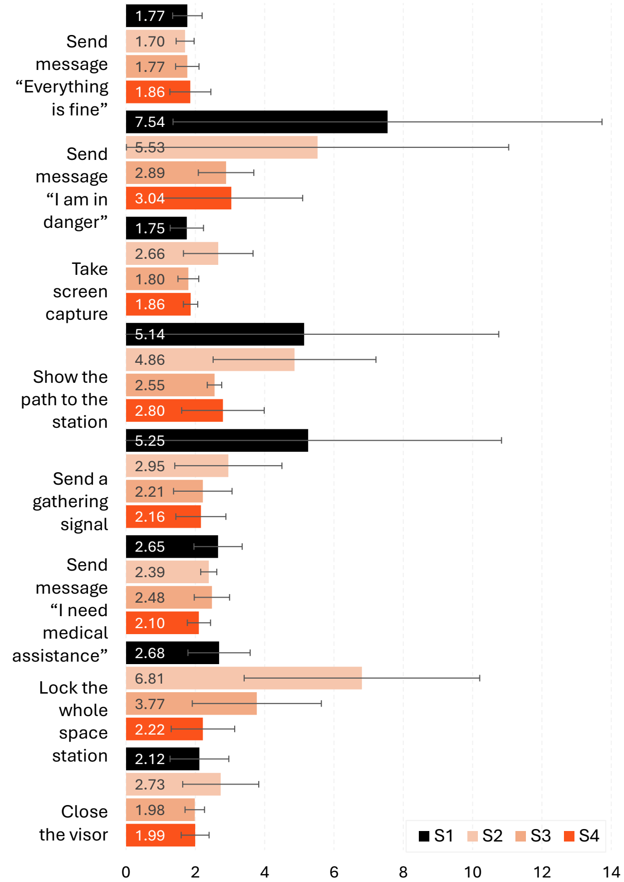

Figure 5: Average gesture PRODUCTION-TIME computed across all referents and sessions. Error bars show 95% CIs.

Figure 5 shows the average PRODUCTION-TIME for gestures performed by the crew members across the four sessions. Overall, production times were largely consistent across gestures, except for minor variations in the first sessions, *e.g.*, for the "I am in danger" and "Send a gathering signal" referents. Three gestures, corresponding to "I am in danger," "Show the path to the station," and "Send a gathering signal," were performed with short production times after S2, while only one gesture, corresponding to the "Lock the whole spatial station" referent, initially exhibited a sharp increase before returning to the same level as the others. All crew members achieved low production times by S4, between 1.86s for "Everything is fine" and 3.04s for "I am in danger." These results suggest that gestures via finger-worn devices represent a plausible interaction modality in terms of production time: execution was fast across sessions, despite increasingly challenging conditions.

4.4.2 User Experience. Figure 6a shows the SMS and SMI values for the UX dimensions evaluated in our experiment. Overall, the scores are high, above the 0.8 threshold in the "good" interval, with the only exception of *Adaptability* in S3 (*M*=0.63, *SD*=1.75). This dimension was affected by the crew's first exposure to more severe conditions, but recovered in S4 at a level comparable with the other dimensions. Ratings in S1 ranged from 1.09 to 1.81, with *Visual Aesthetics* (*M*=1.09, *SD*=1.21) receiving the lowest score and *Perspicuity* (*M*=1.81, *SD*=1.01) the highest. In S4, ratings ranged from 1.25 to 2.00, with *Novelty* (*M*=2.00, *SD*=0.94) and *Trustworthiness* (*M*=2.00, *SD*=1.06) receiving the highest ratings. Overall, the following trends can be distinguished based on SMS (Figure 6a):

- *V-shaped curves* (4/13=31%) undergo a noticeable drop, but then rise: *Efficiency*, *Stimulation*, and *Adaptability*. For example, *Efficiency* started from a mean score of *M*=1.75 (*SD*=0.90) and reached *M*=1.47 (*SD*=1.75) by S4; *Stimulation* was rated at *M*=1.56 (*SD*=1.25) and reached *M*=1.38 (*SD*=1.60) in S4.
- *Inverted V-shaped curves* (3/13=23%) grow locally, but end with a value below that of the previous sessions: *Attractiveness*, *Trust*, and *Intuitive Use*. For example, *Attractiveness* started from *M*=1.56 (*SD*=0.79) in S2, reached *M*=1.78 (*SD*=1.22) in S3, and fell to *M*=1.59 (*SD*=1.73) in S4.
- *Overall ascending trends* (5/13=38%) progressively increase across sessions: *Dependability*, *Novelty*, *Usefulness*, *Visual Aesthetics*, and *Trustworthiness of Content*. For example, the *Trustworthiness of Content* dimension gradually improved from *M*=1.63 (*SD*=1.24) in S2 to *M*=2.00 (*SD*=1.06) in S4.
- *Descending trend* (1/13=8%) for *Perspicuity*, with an average rating of 1.81 in S2, which decreased to 1.56 in S4.

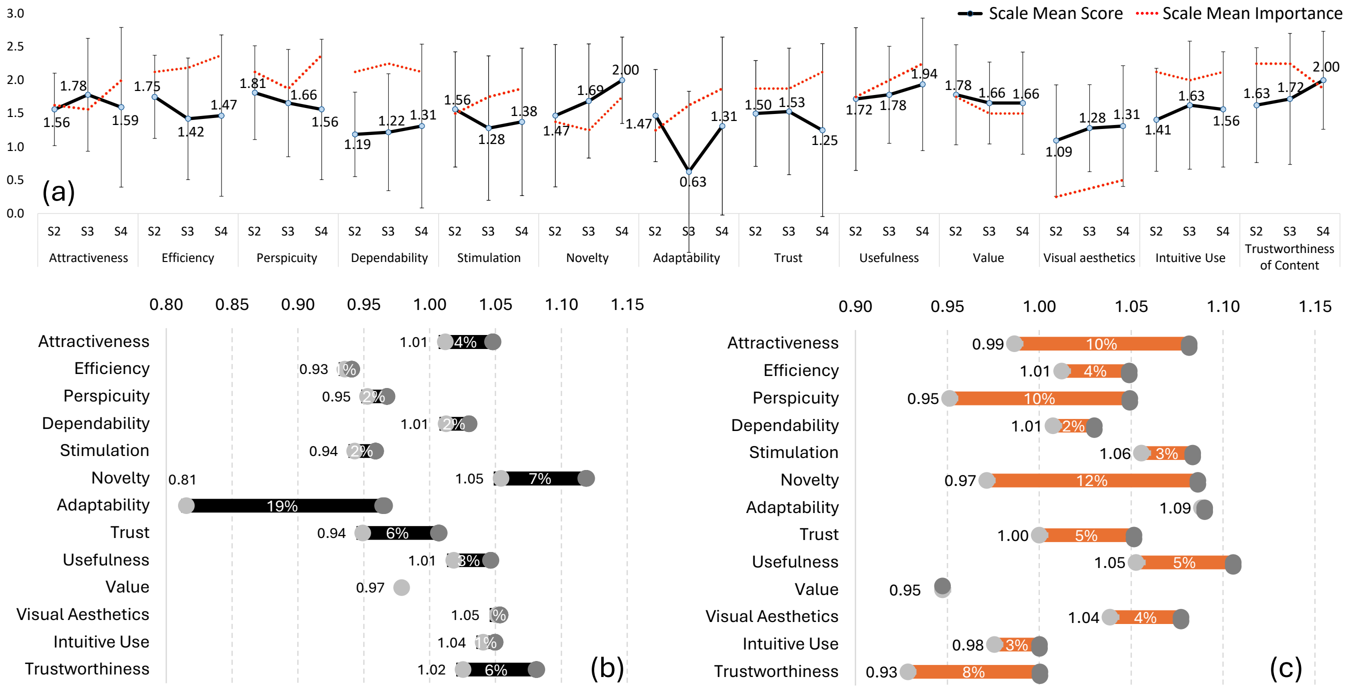

Figure 6: Overview of UX evaluation results, in terms of the scale mean scores and their importance (a), evolution of the scale mean ratios (b) and scale importance ratios (c), respectively. Error bars show 95% CIs.

The SMIs (Figure 6a) are positive throughout the evaluation sessions. Although impacted by the challenging conditions of the sessions, crew members progressively improved their assessments.

Figures 6b and 6c show the SMR and IMR measures as dumbbell plots to present their evolution over the mission. In these plots, any point below or above the reference value of 1.0 indicates a UX dimension perception as less or more positive, respectively, than in the baseline, condition S1. Most ratios started close to 1.0, except for *Adaptability*, which increased by Δ=20% from 0.81 to 0.97. This dimension was followed by *Novelty* (Δ=7%), *Trust* (Δ=6%), and *Attractiveness* (Δ=4%). Overall, the majority of the ratios were close to the baseline and consistent across the evaluation sessions.

4.4.3 Psychological Evaluation. Figure 7 shows the results of the POMS questionnaire for sessions S2 (pre) and S3 (post). We observed for both S2 and S3 the "Iceberg Profile," considered the ideal mood profile for physical performance: low scores on the negative scales and a significant peak, above average, on Vigor-Activity, suggesting that crew members were in positive moods while engaging in gesture interaction. Overall, mood improved from S2 to S3, as reflected by the values for the negative scales decreasing and those for the positive scales increasing, while Total-Mood-Disturbance in S2 (M=46.69) was higher than in S3 (M=39.11). These results indicate that, as a group, the crew experienced an improvement in mental well-being and a reduction in negative mood states, despite the extreme conditions in our experiment. In particular, the commander (CM1) exhibited a decrease in the Total-Mood-Disturbance score from S2 (45.10) to S3 (35.99), indicating significant mood improvement, especially visible in terms of fatigue (64.55 to 53.38); see Figure 7, bottom. CM8 consistently reported low scores and high vigor, suggesting a generally positive emotional state compared to the group average. CM4 exhibited the largest improvements in terms of Anger-Hostility (40.69 to 34.62) and Confusion-Perplexity (58.57 to 42.17) across sessions S2 and S3.

5 Implications for Gesture-based Interaction

Based on our empirical findings, we propose a set of implications for gesture-based interaction in extreme environments, in particular during analog and space missions:

- *Prioritize finger gestures that are robust to extreme constraints: fatigue, limited dexterity, gloves, and psychological stress.* In extreme environments, finger and hand gestures should rely on coarse movements instead of the fine motor precision expected in typical settings. The crew members in our experiment, were able to reliably perform the gestures in Figure 2, despite psychological stress and physical fatigue, with low gesture production times and high recall accuracy. Consequently, gestures should be simple to enable easy execution in such settings, and the consensus set in Figure 2 constitutes a good starting point, to be confirmed in future replications.
- *Favor gesture memorability through semantic mapping and a small vocabulary.* Gestures with clearer semantic relationships to their referents, *e.g.*, for "Everything is fine," exhibited higher recall accuracy and faster user performance and experience stabilization across the evaluation sessions, while abstract mappings, such as for the "I am in danger" referent, required more practice. To support easy recall under extreme conditions, gestures

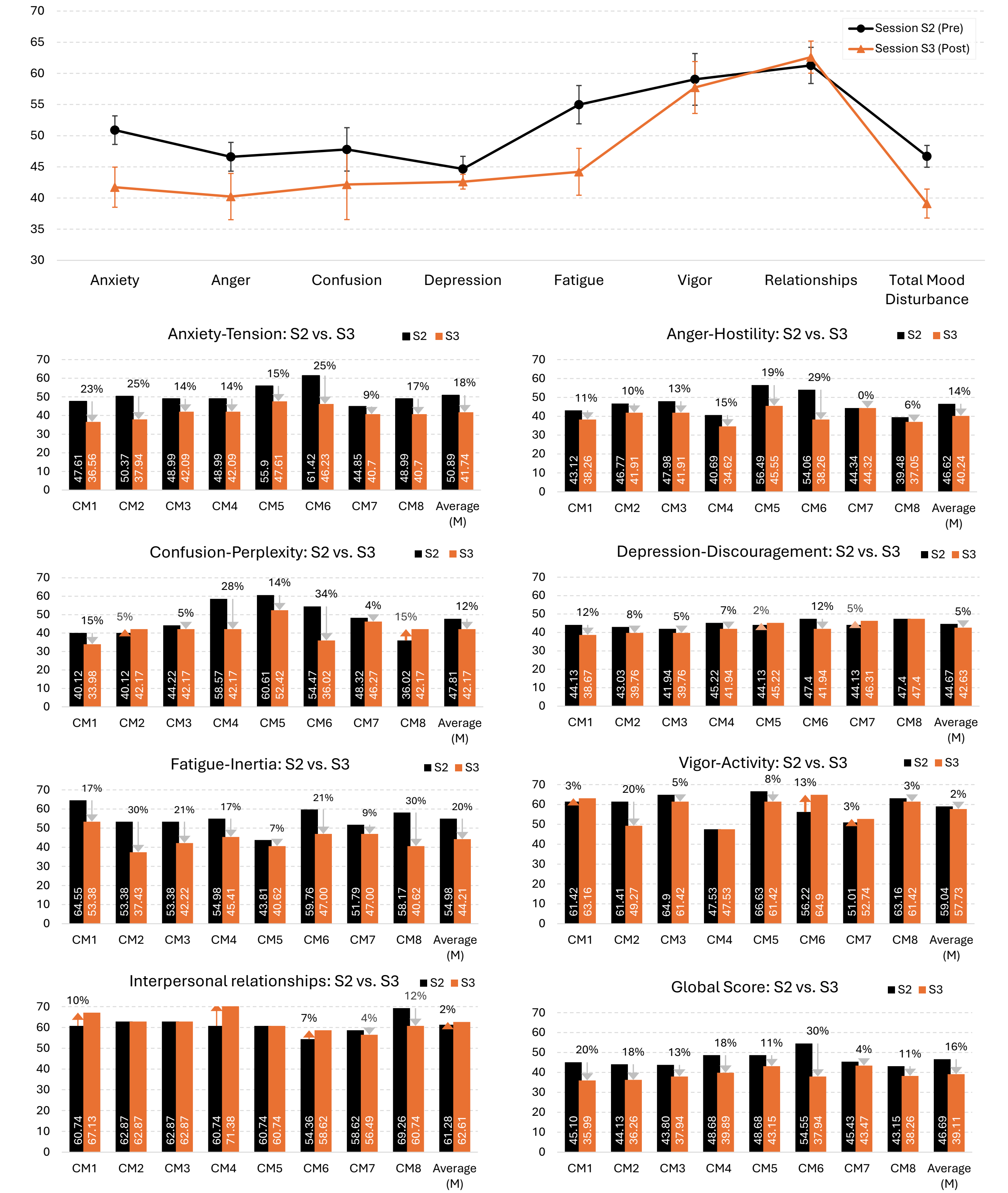

Figure 7: Profile of Mood States (POMS) evaluations in S2 and S3: overview (top) and per crew member. Error bars show 95% CIs.

should integrate intuitive metaphors even more than in typical settings [63]. Furthermore, our experiment successfully employed a user-defined vocabulary of finger gestures, which crew members mastered after repeated exposure. The vocabulary was small, fixed [60], and consistent [11], all characteristics that help reduce cognitive load and support long-term retention during prolonged missions in extreme environments [24].

- *Support repeated reinforcement for consistent user performance and sustained user experience.* Perfect (100%) gesture recall accuracy was achieved after multiple sessions, indicating that gesture-based interfaces should be designed with progressive onboarding, repetition, and rehearsal. Moreover, gesture production times converged to approximately two seconds and remained stable over the more complex physical challenges of the evaluation sessions and increasing mission fatigue. The user experience results revealed that perceived *Trustworthiness of Content*, *Dependability*, and *Usefulness* improved across the evaluation sessions, even when *Visual Aesthetics* was relatively low. In extreme environments, gesture design should prioritize user performance and experience with proper feedback over visual aspects.

Translating these implications further into design principles according to NASA or ESA terminology and quality factors (*e.g.*, reliability, dependability, and trustworthiness) is beneficial [1,18]. For instance, gesture interaction for analog astronauts should be designed as *operationally robust under EVA constraints,* tolerating reduced dexterity, gloves, fatigue, and environmental stress. To support *cognitive economy and low mental workload,* gestures should rely on semantically meaningful mappings and a small, stable command vocabulary that promotes long-term memorability during prolonged missions. Furthermore, gesture execution should exhibit *temporal predictability and performance consistency* across mission environments (*e.g.*, habitat, laboratory, EVA, or driving a rover) for reliable task sequencing and crew trust. Gesture systems should also support *progressive learning and training compatibility* by allowing skill consolidation over repeated mission exposure. Lastly, gesture interaction should remain *resilient to psychological stressors* and fluctuating human performance, which are hallmarks of isolated, confined, and extreme environments.

6 Limitations

Several threats to validity [75] might have influenced our experiment, as follows. Concerning *external validity*, the experiment was conducted at the MDRS, a state-of-the-art facility designed to replicate Mars conditions as closely as possible on Earth; accordingly, our findings should be interpreted in the context of a high-fidelity simulation. Regarding *ecological validity*, the MDRS offered a realistic setting where participants experienced many physical and psychological constraints relevant to space missions, except for very specific ones, such as altered respiratory and gravity conditions, which may further affect user performance and experience. The MDRS represents a simplified approximation of the real conditions imposed by a space station. Pletser [49] explains how researchers design, prepare, develop, test, and fly their experiments on microgravity platforms before sending them to space, which go beyond the MDRS conditions: drop towers, gliders, parabolic flights, sounding rockets, short-duration orbital platforms, and long-duration orbital platforms. Our analog astronauts are considered simplified proxies of real astronauts, who are involved in more than three years of intensive training. Regarding *internal validity*, the crew size was limited to eight astronauts because of logistical constraints at the MDRS. Concerning *construct validity*, our within-subject approach could be complemented in future work by a between-subject design to reduce potential accommodation effects across sessions. Although carry-over effects are unavoidable in a within-subject design, they mirror the training process of real astronauts, who practice the same tasks and procedures before a mission; despite this potential confound, we observed some variations in performance, including degradation across some sessions. Lastly, regarding *conclusion validity*, although all crew members received the same training, this preparation was not comparable to that of actual astronauts, which spans months. Future work could address these limitations by conducting studies with larger samples, for example during multiple missions at the MDRS, and involving crew members with different demographic characteristics in terms of age, handedness, and cultural background, which may reflect in different gesture preferences and user experience findings.

7 Research Ethics and Participant Privacy

The experiment design was reviewed and approved by the Institutional Review Board (IRB), under reference no. Project 2022-60 and by the "Comité d'éthique Hospitalo-Facultaire", under reference no. 2025/27JAN/037 (Advancing Crisis Resource Management Strategies for Space Exploration: Lessons from Mars-Analog Simulation). Before data collection, the participants were informed of the experiment's objectives, procedure, and potential risks, and signed a GDPR-compliant consent form. During data processing, all personally identifiable information was anonymized or removed. Psychological and behavioral data were stored on a local computer with restricted access and used solely for the purposes of this research. The data are not publicly released without additional anonymization and renewed participant consent.

8 Conclusion

We reported empirical findings from the first gesture elicitation study involving a crew of analog astronauts, with a consensus gesture set evaluated in terms of user performance and user experience measures, alongside the monitoring of the crew's psychological profile. Based on our results, we proposed design implications for gesture-based interaction in extreme environments, such as those encountered during Mars-analog missions, performed with finger-worn devices due to their favorable characteristics aligning well with gloved interaction and wearing protective equipment. Further research is needed to validate our design implications, integrate them with established quality factors from NASA/ESA, and extend them to other types of gestures and gesture-sensing devices.

Acknowledgments

Nuwan Attygalle and Jean Vanderdonckt are supported by the EU EIC Pathfinder-Awareness Inside challenge Symbiotik project (1 October 2022-30 September 2026) under Grant no. 101071147. We also acknowledge support by the Fonds de la Recherche Scientifique (FNRS) under Grant n°40035986 COSMIC project.

References

[1] Vicki Ahlstrom. 2016. Human Factors Design Standard. HF-STD-001B. https://hf.tc.faa.gov/hfds

[2] Roland Aigner, Daniel Wigdor, Hrvoje Benko, Michael Haller, David Lindbauer, Alexandra Ion, Shengdong Zhao, and Jeffrey Tzu Kwan Valino Koh. 2012. *Understanding Mid-Air Hand Gestures: A Study of Human Preferences in Usage of Gesture Types for HCI.* Technical Report MSR-TR-2012-111. Microsoft Research. https://www.microsoft.com/en-us/research/publication/understanding-mid-air-hand-gestures-a-study-of-human-preferences-in-usage-of-gesture-types-for-hci

[3] Matthew Allner, Sheryl Bishop, Vadim Gushin, Caoimhe McKay, and Vadim Rygalov. 2008. Crew performance Analysis of a simulated Mars mission at the Mars Desert Research Station (MDRS) in Utah, USA. In *Proceedings of the 59th International Astronautical Congress (IAC '08).* Int. Astronautical Federation, Paris, France, 11–23. https://iafastro.directory/iac/archive/browse/IAC-08/A1/1/173

[4] Paul T. Bartone, Gerald P. Krueger, and Jocelyn V. Bartone. 2018. Individual Differences in Adaptability to Isolated, Confined, and Extreme Environments. *Aerospace Medicine and Human Performance* 89, 6 (June 2018), 536–546. https://doi.org/10.3357/AMHP.4951.2018

[5] Leonie Bensch, Tommy Nilsson, Paul de Medeiros, Florian Dufresne, Andreas Gerndt, Flavie Rometsch, Georgia Albuquerque, Frank Flemisch, Oliver Bensch, Michael Preutenborbeck, and Aidan Cowley. 2023. Towards Balanced Astronaut-Oriented Design for Future EVA Space Technologies. In *Proceedings of the 3rd Int. Conference on Human-Computer Interaction for Space Exploration (SpaceCHI 3.0).* MIT Media Lab, Boston, MA, USA, 10 pages. https://elib.dlr.de/201717

[6] Leonie Bensch, Tommy Nilsson, Martin Nisser, Pat Pataranutaporn, Albrecht Schmidt, and Valentina Sumini (Eds.). 2025. *Advancing Human-Computer Interaction for Space Exploration.* OASIcs, Vol. 130. Leibniz-Zentrum für Informatik, Schloss Dagstuhl. https://www.dagstuhl.de/dagpub/978-3-95977-384-3

[7] Marion G. Ceruti, Vincent Vinh Dinh, Nghia Xuan Tran, Hoa Van Phan, LorRaine T. Duffy, Tu-Anh Ton, Guy Leonard, Emily Medina, Omar Amezcua, Sunny Fugate, Gary J. Rogers, Robert Luna, and Jeffrey Ellen. 2009. Wireless communication glove apparatus for motion tracking, gesture recognition, data transmission, and reception in extreme environments. In *Proceedings of the ACM International Symposium on Applied Computing (SAC '09).* ACM, New York, NY, USA, 172–176. https://doi.org/10.1145/1529282.1529320

[8] Yuting Cheng, Zhanwei Wu, and Ruowei Xiao. 2024. Exploring Methods to Optimize Gesture Elicitation Studies: A Systematic Literature Review. *IEEE Access* 12 (2024), 64958–64979. https://doi.org/10.1109/ACCESS.2024.3387269

[9] William J Clancey. 2006. Participant Observation of a Mars Surface Habitat Mission Simulation. *Habitation* 11, 1 (2006), 27–47. https://doi.org/10.3727/154296606779507132

[10] Guillem Cornella-Barba, Bruno Dumas, Mehdi Ousmer, Santiago Villarreal-Narvaez, Jean Vanderdonckt, Eudald Sangenis, and Adrien Chaffangeon Caillet. 2025. TapStrapGest: Elicitation and Recognition of Ring-based Multi-Finger Gestures. *Proceedings of the ACM on Human-Computer Interaction* 9, 4, Article EICS001 (June 2025), 27 pages. https://doi.org/10.1145/3733047

[11] Tilman Dingler, Rufat Rzayev, Alireza Sahami Shirazi, and Niels Henze. 2018. Designing Consistent Gestures Across Device Types: Eliciting RSVP Controls for Phone, Watch, and Glasses. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18).* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3173574.3173993

[12] Sands Fish. 2018. How to Design Interplanetary Apps. Medium. https://sandsfish.medium.com/how-to-design-interplanetary-apps-22ebefec097d

[13] Sands Fish. 2022. Orientation-Responsive Displays for Microgravity. In *Proceedings of the SpaceCHI 2.0 Workshop, Advancing Human-Computer Interaction for Space Exploration at CHI 2022 (SpaceCHI 2.0).* MIT Media Lab, MA, USA, 5 pages. https://drive.google.com/open?id=1BUQDkbt6tSCJ759Z00FuaEmm6Af5h2z9

[14] Gregory Goth. 2012. Software on Mars. *Commun. ACM* 55, 11 (2012), 13–15. https://doi.org/10.1145/2366316.2366321

[15] Sandra Häuplik-Meusburger and Sheryl Bishop. 2021. Introduction. In *Space Habitats and Habitability: Designing for Isolated and Confined Environments on Earth and in Space.* Springer, Cham, 1–8. https://doi.org/10.1007/978-3-030-69740-2_1

[16] Juvia P. Heuchert and D.M. McNair. 2012. Profile of Mood States (POMS), 2nd Edition. *APA PsycTests* 2, 1 (2012). https://doi.org/10.1037/t05057-000

[17] Andreas Hinderks, Dominique Winter, Martin Schrepp, and Jörg Thomaschewski. 2019. Applicability of User Experience and Usability Questionnaires. *Journal of Universal Computer Science* 25, 13 (2019), 1717–1735. http://www.jucs.org/jucs_25_13/applicability_of_user_experience

[18] Kritina L. Holden, Jennifer L. Boyer, Neta Ezer, Keith Holubec, Anikó Sándor, and John-Paul Stephens. 2013. Human Factors in space vehicle design. *Acta Astronautica* 92, 1 (2013), 110–118. https://doi.org/10.1016/j.actaastro.2012.10.020

[19] Hong Hu, Jian Gang Chao, and Zai Qian Zhao. 2014. Study of Vision-Based Hand Gesture Recognition System for Astronaut Virtual Training. In *Advances in Applied Sciences, Engineering and Technology II (Advanced Materials Research, Vol. 998).* Trans Tech Publications Ltd, Switzerland, 1062–1065. https://doi.org/10.4028/www.scientific.net/AMR.998-999.1062

[20] Woosoon Jung and Hyung Gyu Lee. 2022. Energy–Accuracy Aware Finger Gesture Recognition for Wearable IoT Devices. *Sensors* 22, 13, Article 4801 (2022), 17 pages. https://doi.org/10.3390/s22134801

[21] Wiesław Kopeć, Grzegorz Pochwatko, Monika Kornacka, Wiktor Stawski, Maciej Grzeszczuk, Kinga Skorupska, Barbara Karpowicz, Rafał Masłyk, Pavlo Zinevych, Stanisław Knapiński, Steven Barnes, and Cezary Biele. 2024. Human Factors in Space Exploration: Opportunities for International and Interdisciplinary Collaboration. arXiv:2403.12344 [cs.HC] https://arxiv.org/abs/2403.12344

[22] Vassilis Kostakos. 2011. The Challenges and Opportunities of Designing Pervasive Systems for Deep-Space Colonies. *Personal and Ubiquitous Computing* 15, 5 (2011), 479–486. https://doi.org/10.1007/S00779-010-0317-5

[23] Lauren Blackwell Landon, Grace L. Douglas, Meghan E. Downs, Maya R. Greene, Alexandra M. Whitmire, Sara R. Zwart, and Peter G. Roma. 2019. The Behavioral Biology of Teams: Multidisciplinary Contributions to Social Dynamics in Isolated, Confined, and Extreme Environments. *Frontiers in Psychology* 10 (2019), 442358. https://doi.org/10.3389/fpsyg.2019.02571

[24] Lauren B. Leveton, Camille Shea, Kelley J Slack, Kathryn E. Keeton, and Lawrence A. Palinkas. 2009. Antarctica Meta-analysis: Psychosocial Factors Related to Long-duration Isolation and Confinement. In *Proceedings of Human Research Program Investigators Workshop.* Universities Space Research Association, USA. https://ntrs.nasa.gov/citations/20090007551

[25] Rhema Linder, Chase Hunter, Jacob McLemore, Senjuti Dutta, Fatema Akbar, Ted Grover, Thomas Breideband, Judith W. Borghouts, Yuwen Lu, Gloria Mark, Austin Z. Henley, and Alex C. Williams. 2022. Characterizing Work-Life for Information Work on Mars: A Design Fiction for the New Future of Work on Earth. *Proceedings of ACM Human-Computer Interaction* 6, GROUP, Article 40 (jan 2022), 27 pages. https://doi.org/10.1145/3492859

[26] Gu Lingyun, Zhang Lin, and Wang Zhaokui. 2020. Hierarchical Attention-Based Astronaut Gesture Recognition: A Dataset and CNN Model. *IEEE Access* 8 (2020), 68787–68798. https://doi.org/10.1109/ACCESS.2020.2986473

[27] Jinguo Liu, Yifan Luo, and Zhaojie Ju. 2016. An Interactive Astronaut-Robot System with Gesture Control. *Computational Intelligence and Neuroscience* 2016, 1 (2016), 7845102. https://doi.org/10.1155/2016/7845102

[28] Yilin Liu, Fengyang Jiang, and Mahanth Gowda. 2020. Finger Gesture Tracking for Interactive Applications: A Pilot Study with Sign Languages. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 3, Article 112 (Sept. 2020), 21 pages. https://doi.org/10.1145/3414117

[29] Mary Lou Maher, Tim Clausner, Barbara Tversky, David Kirsh, Judy Kay, Andreea Danielescu, and Kazjon Grace. 2014. Gesture-based Interaction Design: Communication and Cognition. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14).* ACM, New York, NY, USA, 61–64. https://doi.org/10.1145/2559206.2559220

[30] Jessica J. Márquez, Lauren Blackwell Landon, and Eduardo Salas. 2023. The Next Giant Leap for Space Human Factors: The Opportunities. *Human Factors* 65, 6 (2023), 1279–1288. https://doi.org/10.1177/00187208231174955

[31] Douglas M. McNair, Juvia W.P. Heuchert, and E. Shilony. 2003. *Research with the Profile of Mood States (POMS) 1964-2002: A comprehensive bibliography.* Multi-Health Systems (MHS), Toronto, Canada.

[32] Douglas M. McNair, M. Lorr, and L.F. Droppleman. 1981. *Educational and industrial testing service manual for the Profile of Mood States.* Educational and Industrial Testing Service, San Diego, CA, USA.

[33] Kaitlin R. McTigue, Megan E. Parisi, Tina L. Panontin, Shu-Chieh Wu, and Alonso H. Vera. 2023. How to Keep Your Space Vehicle Alive: Maintainability Design Principles for Deep-Space Missions. In *Proceedings of SpaceCHI 3.0, A Conference on Human-Computer Interaction for Space Exploration (SpaceCHI 3.0).* MIT Media Lab, MA, USA, 8 pages. https://human-factors.arc.nasa.gov/publications/SpaceCHI2023_Maintainability.pdf

[34] Anna-Lena Meiners, Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2024. A Benchmark for the UEQ+ Framework: Construction of a Simple Tool to Quickly Interpret UEQ+ KPIs. *Int. J. Interact. Multim. Artif. Intell.* 9, 1 (2024), 104. https://doi.org/10.9781/IJIMAI.2023.05.003

[35] Susmita Mohanty, Jesper Jørgensen, and Maria Nyström. 2006. Psychological Factors Associated with Habitat Design for Planetary Mission Simulators. In *Proceedings of Space Architecture Symposium: Space Human Factors (Space 2006).* American Institute of Aeronautics and Astronautics, Reston, VA, USA. https://doi.org/10.2514/6.2006-7345

[36] Meredith Ringel Morris. 2012. Web on the wall: Insights from a multimodal interaction elicitation study. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces (ITS '12).* ACM, New York, NY, USA, 95–104. https://doi.org/10.1145/2396636.2396651

[37] Kristian Mrazek and Tauheed Khan Mohd. 2021. Using LSTM Models on Accelerometer Data to Improve Accuracy of Tap Strap 2 Wearable Keyboard. In *Proceedings of 13th International Conference on Intelligent Human Computer Interaction.* Springer, Berlin, 27–38. https://doi.org/10.1007/978-3-030-98404-5_3

[38] Joost Nelis, Christopher Elliott, and Katrina Campbell. 2018. "The Smartphone's Guide to the Galaxy": In Situ Analysis in Space. *Biosensors* 8, 4, Article 96 (2018), 31 pages. https://doi.org/10.3390/bios8040096

[39] Tommy Nilsson, Leonie Bensch, Florian Dufresne, Flavie Rometsch, Paul de Medeiros, Enrico Guerra, Florian Saling, Andrea Casini, and Aidan Cowley. 2023. Out of this World Design: Bridging the Gap between Space Systems Engineering and Participatory Design Practices. In *Proceedings of SpaceCHI 3.0, A Conference on Human-Computer Interaction for Space Exploration*. MIT Media Lab, MA, USA, 9 pages. https://spacechi.media.mit.edu/spacechi-2023-program

[40] Tommy Nilsson, Flavie Rometsch, Leonie Becker, Florian Dufresne, Paul Demedeiros, Enrico Guerra, Andrea Emanuele Maria Casini, Anna Vock, Florian Gaeremynck, and Aidan Cowley. 2023. Using Virtual Reality to Shape Humanity's Return to the Moon: Key Takeaways from a Design Study. In *Proceedings of the ACM Int. Conference on Human Factors in Computing Systems (CHI '23)*. ACM, New York, NY, USA, Article 305, 16 pages. https://doi.org/10.1145/3544548.3580718

[41] Marianna Obrist, Yunwen Tu, Lining Yao, and Carlos Velasco. 2019. Space Food Experiences: Designing Passenger's Eating Experiences for Future Space Travel Scenarios. *Frontiers in Computer Science* 1 (2019), 17 pages. https://doi.org/10.3389/fcomp.2019.00003

[42] Elif Oguz, Martin Kubicek, and David Clelland. 2018. Failure modes and criticality analysis of the preliminary design phase of the Mars Desert Research Station considering human factors. *Reliability Engineering & System Safety* 178 (2018), 247–254. https://doi.org/10.1016/j.ress.2018.06.023

[43] Laurent Opsomer, Benoit Delhaye, Vincent Théate, Jean-Louis Thonnard, and Philippe Lefèvre. 2023. A haptic illusion created by gravity. *iScience* 26, 7 (2023), 16 pages. https://doi.org/10.1016/j.isci.2023.107246

[44] Lawrence A. Palinkas and Peter Suedfeld. 2021. Psychosocial issues in isolated and confined extreme environments. *Neuroscience & Biobehavioral Reviews* 126 (2021), 413–429. https://doi.org/10.1016/j.neubiorev.2021.03.032

[45] Paul Parsons, Zixu Zhang, and Jackson Murray. 2022. Adaptive Performance: A Generative Theory for HCI Design in Extraterrestrial Habitats. In *Proceedings of the SpaceCHI 2.0 Workshop, Advancing Human-Computer Interaction for Space Exploration at CHI 2022 (SpaceCHI 2.0)*. MIT Media Lab, MA, USA, 4 pages. https://drive.google.com/open?id=1vYNe4HXcDVmMeNn0X_lXWn7hCUXIDOKU

[46] Pat Pataranutaporn, Valentina Sumini, Ariel Ekblaw, Melodie Yashar, Sandra Häuplik-Meusburger, Susanna Testa, Marianna Obrist, Dorit Donoviel, Joseph Paradiso, and Pattie Maes. 2021. SpaceCHI: Designing Human-Computer Interaction Systems for Space Exploration. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21)*. ACM, New York, NY, USA, Article 96, 6 pages. https://doi.org/10.1145/3411763.3441358

[47] Pat Pataranutaporn, Valentina Sumini, Melodie Yashar, Susanna Testa, Marianna Obrist, Scott Davidoff, Amber M. Paul, Dorit Donoviel, Jimmy Wu, Sands A Fish, Ariel Ekblaw, Albrecht Schmidt, Joe Paradiso, and Pattie Maes. 2022. SpaceCHI 2.0: Advancing Human-Computer Interaction Systems for Space Exploration. In *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, Article 77, 7 pages. https://doi.org/10.1145/3491101.3503708

[48] Shilpa Pavithran, Sruthi Pallathuvalappil, Elizabeth George, Javed G S, and Alex James. 2025. EM Wave-Based Hand Gesture Recognition for Astronauts Using 3D Memristive Neural Network. *IEEE Journal of Microwaves* 5, 1 (2025), 48–58. https://doi.org/10.1109/JMW.2024.3506736

[49] Vladimir Pletser. 2020. *Preparation of Space Experiments*. IntechOpen, London. https://doi.org/10.5772/intechopen.92406

[50] Tereza Pultarova. 2014. Two weeks on Mars. *Engineering & Technology* 9, 3 (2014), 34–37. https://doi.org/10.1049/et.2014.0328

[51] Erin E. Richardson. 2024. Predicting and Protecting the Cognitive Health of Operators in Isolated, Confined, and Extreme Environments. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS '24)*. Int. Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 2764–2766. https://dl.acm.org/doi/abs/10.5555/3635637.3663280

[52] Michael Saint-Guillain, Jean Vanderdonckt, Nicolas Burny, Vladimir Pletser, Tiago Vaquero, Steve Chien, Alexander Karl, Jessica Marquez, Cyril Wain, Audrey Comein, Ignacio S. Casla, Jean Jacobs, Julien Meert, Cheyenne Chamart, Sirga Drouet, and Julie Manon. 2023. Enabling astronaut self-scheduling using a robust advanced modelling and scheduling system: An assessment during a Mars analogue mission. *Advances in Space Research* 72, 4 (2023), 1378–1398. https://doi.org/10.1016/j.asr.2023.03.045

[53] Inderpal S. Saluja, David R. Williams, Daniel Woodard, Janusz Kaczorowski, Ben Douglas, Philip J. Scarpa, and Jean-Marc Comtois. 2008. Survey of astronaut opinions on medical crewmembers for a mission to Mars. *Acta Astronautica* 63, 5 (2008), 586–593. https://doi.org/10.1016/j.actaastro.2008.05.002

[54] Irene Lia Schlacht and Henrik Birke. 2011. Space design - Visual interface of space habitats. *Personal and Ubiquitous Computing* 15, 5 (2011), 497–509. https://doi.org/10.1007/S00779-010-0326-4

[55] Martin Schrepp. 2021. Measuring User Experience with Modular Questionnaires. In *Proceedings of International Conference on Advanced Computer Science and Information Systems*. IEEE, USA, 1–6. https://doi.org/10.1109/ICACSIS53237.2021.9631321

[56] Martin Schrepp and Jörg Thomaschewski. 2019. Design and Validation of a Framework for the Creation of User Experience Questionnaires. *International Journal of Interactive Multimedia and Artificial Intelligence* 5, 7 (2019), 88–95. https://doi.org/10.9781/IJIMAI.2019.06.006

[57] Sue Ann Seah, Marianna Obrist, Anne Roudaut, and Sriram Subramanian. 2015. Need for Touch in Human Space Exploration: Towards the Design of a Morphing Haptic Glove – ExoSkin. In *Proceedings of 15th IFIP TC 13 Int. Conf. on Human-Computer Interaction, Lecture Notes in Computer Science (INTERACT '15, Vol. 9299)*. Springer-Verlag, Berlin, Heidelberg, 18–36. https://doi.org/10.1007/978-3-319-22723-8_3

[58] Shivang Shelat, John A. Karasinski, Erin E. Flynn-Evans, and Jessica J. Marquez. 2022. Evaluation of User Experience of Self-scheduling Software for Astronauts: Defining a Satisfaction Baseline. In *Engineering Psychology and Cognitive Ergonomics*. Springer, Cham, 433–445. https://doi.org/10.1007/978-3-031-06086-1_34

[59] Roy Shilkrot, Jochen Huber, Jürgen Steimle, Suranga Nanayakkara, and Pattie Maes. 2015. Digital Digits: A Comprehensive Survey of Finger Augmentation Devices. *ACM Comput. Surv.* 48, 2, Article 30 (Nov. 2015), 29 pages. https://doi.org/10.1145/2828993

[60] Arthur Sluyters, Quentin Sellier, Jean Vanderdonckt, Vik Parthiban, and Pattie Maes. 2023. Consistent, Continuous, and Customizable Mid-Air Gesture Interaction for Browsing Multimedia Objects on Large Displays. *Int. J. of Human-Comp. Interaction* 39, 12 (2023), 2492–2523. https://doi.org/10.1080/10447318.2022.2078464

[61] David Smitherman and Andrew Schnell. 2020. Gateway Lunar Habitat Modules as the Basis for a Modular Mars Transit Habitat. In *Proceedings of the IEEE Aerospace Conference*. IEEE, USA, 1–12. https://doi.org/10.1109/AERO47225.2020.9172540

[62] Jeff Spry. 2022. Watch how NASA astronauts communicate in space without speaking. https://www.space.com/how-astronauts-communicate-without-speaking-video

[63] Zhida Sun, Sitong Wang, Chengzhong Liu, and Xiaojuan Ma. 2022. Metaphoraction: Support Gesture-based Interaction Design with Metaphorical Meanings. *ACM Trans. Comput.-Hum. Interact.* 29, 5, Article 45 (Oct. 2022), 33 pages. https://doi.org/10.1145/3511892

[64] Andrew Terhorst and Jason A. Dowling. 2022. Terrestrial Analogue Research to Support Human Performance on Mars: A Review and Bibliographic Analysis. *Space: Science & Technology* 2022 (2022), 18 pages. https://doi.org/10.34133/2022/9841785

[65] Shelby Thompson, Miranda Mesloh, Scott England, Elizabeth Benson, and Sudhakar Rajulu. 2011. The Effects of Extravehicular Activity (EVA) Glove Pressure on Tactility. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 55, 1 (2011), 1385–1388. https://doi.org/10.1177/1071181311551288

[66] Wen-Jie Tseng, Samuel Huron, Eric Lecolinet, and Jan Gugenheimer. 2022. Enabling Virtual Reality Interactions in Confined Spaces by Re-Associating Finger Motions. In *Proceedings of the IEEE Interntational Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW '22)*. IEEE, USA, 574–575. https://doi.org/10.1109/VRW55335.2022.00138

[67] Selen Türkay, Abigail Allwood, Ross Brown, Christoph Schrank, Luke Nothdurft, Thierry Peynot, and Cael Gallagher. 2024. CHI in the Sky: Role of Human Computer Interaction for Space and Planetary Surface Exploration. In *Proceedings of the 35th Australian Computer-Human Interaction Conference (OzCHI '23)*. ACM, New York, NY, USA, 712–715. https://doi.org/10.1145/3638380.3638439

[68] Martine Van Puyvelde, Daisy Gijbels, Thomas Van Caelenberg, Nathan Smith, Loredana Bessone, Susan Buckle-Charlesworth, and Nathalie Pattyn. 2022. Living on the edge: How to prepare for it? *Frontiers in Neuroergonomics* 3 (2022), 20 pages. https://doi.org/10.3389/fnrgo.2022.1007774

[69] Jean Vanderdonckt, Radu-Daniel Vatavu, and Romain Maddox. 2025. Between Bulky Suits and Isolated, Deserted Landscape: Measuring the User Experience of Astronaut-Drone Interaction MHCI005. *Proc. ACM Hum. Comput. Interact.* 9, 5 (2025), 1–20. https://doi.org/10.1145/3743712

[70] Jean Vanderdonckt, Radu-Daniel Vatavu, Julie Manon, Romain Maddox, Michael Saint-Guillain, Philippe Lefevre, and Jessica J. Marquez. 2025. UX, but on Mars: Exploring User Experience in Extreme Environments with Insights from a Mars Analog Mission. In *Proceedings of the ACM International Conference on Designing Interactive Systems Conference (DIS '25)*. ACM, New York, NY, USA, 3235–3250. https://doi.org/10.1145/3715336.3735706

[71] Radu-Daniel Vatavu and Jean Vanderdonckt. 2025. What is Extreme in Human-Computer Interaction Research?. In *Adjunct Proceedings of the 27th International Conference on Mobile Human-Computer Interaction, MobileHCI 2025, Sharm El-SheikhEgypt, September 22-25, 2025*, Yomna Abdelrahman, Passant Elagroudy, and Florian Alt (Eds.). ACM, 17:1–17:7. https://doi.org/10.1145/3737821.3749561

[72] Radu-Daniel Vatavu, Jean Vanderdonckt, Julie Manon, Michael Saint-Guillain, Philippe Lefevre, Romain Maddox, and Jessica J. Marquez. 2025. Conducting Human-Computer Interaction Scientific Experiments in Extreme Environments: Insights from Analog Mars Missions. *Romanian Journal of Information Science and Technology* 28, 3 (2025), 247–259. https://doi.org/10.59277/ROMJIST.2025.3.01

[73] Radu-Daniel Vatavu and Jacob O. Wobbrock. 2015. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1325–1334. https://doi.org/10.1145/2702123.2702223

[74] Tara Vega, Ariana Bueno, Chad Cerutti, Haoran Chang, Matthew Garvin, Catalina Garza, Parker Kurlander, Nilton Renno, Fernando Figueroa, and Lauren Underwood. 2022. Enabling a Voice Management System for Space Applications. In *Proceedings of the IEEE Aerospace Conference*. Stennis Space Center, USA, 17 pages. https://ntrs.nasa.gov/citations/20210022373

[75] Roberto Verdecchia, Emelie Engström, Patricia Lago, Per Runeson, and Qunying Song. 2023. Threats to validity in software engineering research: A critical reflection. *Information and Software Technology* 164 (2023), 107329. https://doi.org/10.1016/j.infsof.2023.107329

[76] Santiago Villarreal-Narvaez, Arthur Sluÿters, Jean Vanderdonckt, and Radu-Daniel Vatavu. 2024. Brave New GES World: A Systematic Literature Review of Gestures and Referents in Gesture Elicitation Studies. *Comput. Surveys* 56, 5 (2024), 128:1–128:55. https://doi.org/10.1145/3636458

[77] Bernhard Weber and Martin Stelzer. 2022. Sensorimotor impairments during spaceflight: Trigger mechanisms and haptic assistance. *Frontiers in Neuroergonomics* 3 (2022), 16 pages. https://doi.org/10.3389/fnrgo.2022.959894

[78] Hongyi Wen, Julian Ramos Rojas, and Anind K. Dey. 2016. Serendipity: Finger Gesture Recognition using an Off-the-Shelf Smartwatch. In *Proceedings of the ACM International Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 3847–3851. https://doi.org/10.1145/2858036.2858466

[79] Olivier White, Yannick Bleyenheuft, Renaud Ronsse, Allan M. Smith, Jean-Louis Thonnard, and Philippe Lefèvre. 2008. Altered Gravity Highlights Central Pattern Generator Mechanisms. *Journal of Neurophysiology* 100, 5 (2008), 2819–2824. https://doi.org/10.1152/jn.90436.2008

[80] Mathias Wilhelm, Daniel Krakowczyk, and Sahin Albayrak. 2020. PeriSense: Ring-Based Multi-Finger Gesture Interaction Utilizing Capacitive Proximity Sensing. *Sensors* 20, 14, Article 3990 (2020), 23 pages. https://doi.org/10.3390/s20143990

[81] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 1083–1092. https://doi.org/10.1145/1518701.1518866

[82] Chenhan Xu, Bing Zhou, Gurunandan Krishnan, and Shree Nayar. 2023. AO-Finger: Hands-free Fine-grained Finger Gesture Recognition via Acoustic-Optic Sensor Fusing. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. ACM, New York, NY, USA, Article 306, 14 pages. https://doi.org/10.1145/3544548.3581264

[83] Chengzhu Yu and John H.L. Hansen. 2017. A study of voice production characteristics of astronuat speech during Apollo 11 for speaker modeling in space. *The Journal of the Accoustical Society of America* 141 (2017), 1605–1614. https://doi.org/10.1121/1.4976048

[84] Shanfei Zhang, Zhigang Xia, Zhufeng Liu, Qi Wang, Yamei Yue, Jianyu Huang, and Bin Su. 2023. Magnetic/conductive/elastic multi-material 3D-printed self-powered sensing gloves for underwater/smoke environmental Human-Computer Interaction. *Chemical Engineering Journal* 463 (2023), 142388. https://doi.org/10.1016/j.cej.2023.142388

[85] Tianming Zhao, Jian Liu, Yan Wang, Hongbo Liu, and Yingying Chen. 2018. PPG-based Finger-level Gesture Recognition Leveraging Wearables. In *Proceedings of the IEEE Conference on Computer Communications (INFOCOM 18)*. IEEE, USA, 1457–1465. https://doi.org/10.1109/INFOCOM.2018.8486006

[86] Jimin Zheng, Shivang M. Shelat, and Jessica J. Marquez. 2023. Facilitating Crew-Computer Collaboration During Mixed-Initiative Space Mission Planning. In *Proceedings of SpaceCHI 3.0, A Conference on Human-Computer Interaction for Space Exploration (SpaceCHI 3.0)*. MIT Media Lab, MA, USA, 7 pages. https://ntrs.nasa.gov/citations/20230008619

[87] Yubin Zhong, Tao Wu, Yan Han, Feiyang Wang, Dan Zhao, Zhen Fang, Linxin Pan, and Chen Tang. 2025. Advancements in Mars Habitation Technologies and Terrestrial Simulation Projects: A Comprehensive Review. *Aerospace* 12, 6, Article 510 (2025), 18 pages. https://doi.org/10.3390/aerospace12060510