Human Factors in Design, Engineering, and Computing, Vol. 238, 2026, 1373–1383
https://doi.org/10.54941/ahfe1008366

Human Testing of Spacesuit Size Preference Using 3D Printed Mockups

Cassandra Vining[1], Chris Miller[2], K. Han Kim[3], and Nate J. Newby[4]

[1]Aegis Aerospace, Houston, TX 77058, USA
[2]KBR, Inc., Houston, TX 77058, USA
[3]Leidos, Inc., Houston, TX 77058, USA
[4]NASA Johnson Space Center, Houston, TX 77058, USA

ABSTRACT

This study presents a novel spacesuit fit assessment method that parameterized suit-to-body interaction patterns as a predictor of fit quality. A statistical model was constructed by leveraging historical NASA fit check data and was validated via Human-in-the-Loop (HITL) testing. Fully customized Hard Upper Torso (HUTs) were developed for five subjects. Motion capture, laser scanning, and subjective surveys were used to collect data on each HUT's impact on fit, indexing, mobility, and reach. The developed model opens the door for customized hardware and extended population analysis.

Keywords: Spacesuit fit, Human body modeling, 3D printing, Custom-fit spacesuit design, Extravehicular Activity (EVA), Shuttle EMU

INTRODUCTION

Extravehicular Activity (EVA) is a critical element of human space exploration, required for vehicle repairs and science in space. The current-generation EVA suit, the Shuttle Extravehicular Mobility Unit (EMU; Figure 1), utilizes a modular design where each component has a discrete sizing scheme. Even with this highly reconfigurable architecture, interaction with the EVA suit is a siginficant cause of injury to crewmembers on the International Space Station (ISS) (Scheuring et al., 2009). One factor for injury risk is inadequate spacesuit fit (Dunn et al., 2022). Inadequate suit fit can also inhibit mobility, which then compounds the injury risk as crew with reduced mobility may have to work harder against the suit to complete EVA tasks.

The Shuttle EMU Hard Upper Torso (HUT) component connects the helmet bubble, life support systems, arm assemblies, and waist assembly (Figure 1). The five originally designed sizes were intended to accommodate a range of crewmembers from the 5th percentile female to 95th percentile male for critical anthropometric dimensions. However, only the three largest sizes were constructed, leading to an actual accommodated range of 40th percentile female to 95th percentile male (NASA, 1999). Additionally, the sizing scheme was established using linear dimensions (e.g., linear chest breadth) rather than the complex non-linear interactions between unique anthropometric characteristics and the suit. Given the increasing human spaceflight activities

Received October 13, 2026; Revised November 18, 2026; Accepted November 27, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

and higher number of EVAs planned for future programs, the need for more structured methods of assessing spacesuit fit becomes more critical.

This paper focuses on the human-in-the-loop (HITL) testing element of a larger study whose overarching purpose was to evaluate spacesuit accommodation rates across the NASA crew population. A statistical model was developed to conduct virtual fit checks. A HITL series was conducted to validate the model.



Figure 1: NASA extravehicular mobility unit and hard upper torso (NASA, 2019).

MODELING

Precise HUT and body geometries are required to effectively evaluate the interactions between them. The internal geometry of each Shuttle EMU HUT size was collected using a hand-held 3D laser scanner (FARO Laser Line Probe), and the 3D body geometry of each subject was collected with a whole-body laser scanner (Humanetics Vitus). The resulting geometries were overlaid to digitally estimate the suit-to-body contact pattern between the person and the HUT. Combining these estimated contact patterns with historic EMU fit check data from NASA's Neutral Buoyancy Laboratory (NBL; n = 105 crew and non-crew test subjects) allowed the study team to correlate suit-to-body interaction patterns (i.e., "contact patterns") with the fit condition (Figure 2). In the figure, fit condition is represented by an integer denoting how many sizes the HUT is larger/smaller than the subject's preferred. For instance, the "Optimal+1" condition represents the fit condition produced when a subject wears a HUT that is one size larger than their preferred. A model was then developed that inverts the relationship: By defining a targeted level of fit for an individual subject's body geometry, a HUT design can be generated that will produce suit-to-body interaction patterns associated with the desired level of fit.

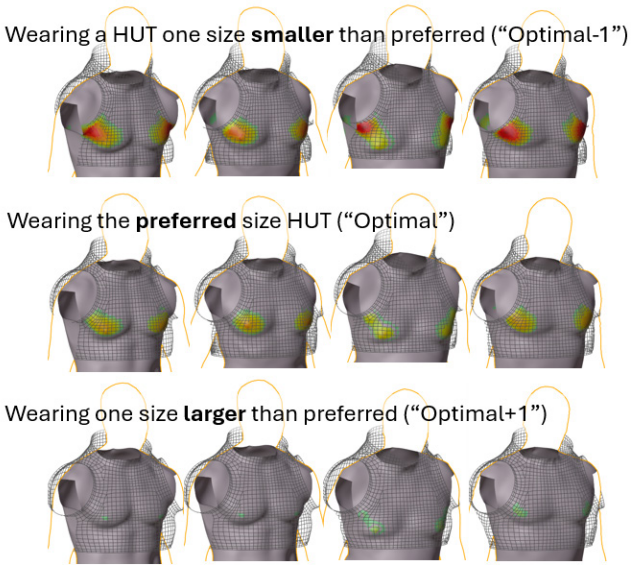


Figure 2: Contact patterns in different HUT fit conditions.

HUMAN SUBJECT TESTING

The model was validated through a HITL test series with five subjects (3 male, 2 female) with a range of anthropometric characteristics, as outlined in Table 1. The protocol was approved by the NASA JSC Test Readiness Review Board, and all subjects were cleared for participation through the NASA test subject clinic and signed an informed consent form.

Table 1: Test subject anthropometry represented in gender-specific percentile.

	Minimum	Maximum
Stature	38 F	89 M
Weight	17 F	95 M
Chest Breadth	66 F	79 M
Chest Depth	24 F	93 M

For each subject, the statistical model was used to generate three custom HUT internal surface geometries—an "Optimal" design, a design one size greater ("Optimal+1"), and a design one size smaller ("Optimal-1"). Each HUT surface design was extruded to a 3D geometry that could be 3D printed. Due to time and cost constraints, it was not feasible to 3D print all 15 unique HUTs (5 subjects × 3 targeted fit conditions). Instead, a faster method was developed where only 7 full HUTs and 11 sets of "sizing inserts" were printed.

Each set of sizing inserts represented the intersection between 3D mesh volume of one of the fabricated HUT geometries (e.g., a subject's "Optimal+1") with that of one of the non-printed HUT geometries (e.g., their "Optimal" or "Optimal-1"; Figure 3). Using the resulting intersecting

geometry, sizing inserts were designed and 3D printed that allowed the study team to reconfigure the full HUTs into any of the remaining 11 targeted geometries.

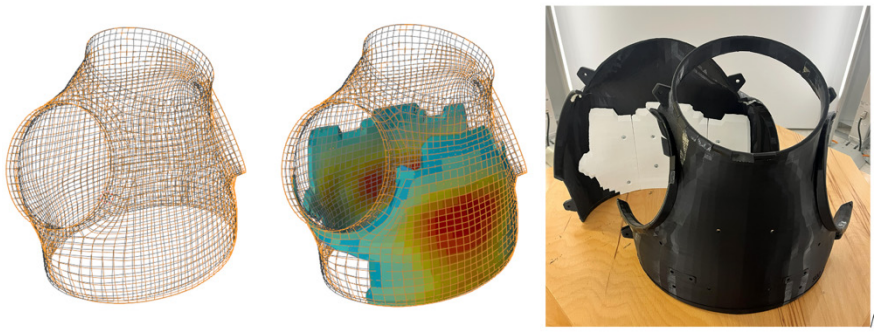


Figure 3: Sizing insert workflow (left: reference HUT geometry; middle: intersection with a target HUT; right: 3D printed mockup with reference HUT and inserts).

In an actual spacesuit, the HUT is held in place by the surrounding suit components, such as the arm softgoods and the waist-brief assembly, in addition to shoulder straps. Because the HUT in this study is isolated from the suit, it was necessary to add hardware to minimize unwanted HUT motion. Subjects wore a climbing-style harness under the HUT with connecting straps that held the HUT in place relative to the subject's torso. Several elements were also added to prevent possible subject injury during testing, such as shoulder pads and foam around the arm and head openings.

For each subject, data were collected across two sessions. In the first session, body scan geometry was collected using a whole-body laser scanner (Humanetics Vitus). The subject's custom HUTs and sizing inserts were fabricated before the second session. Finally, during the second test session, the subject donned and evaluated each of their three HUT geometries ("Optimal," "Optimal-1," and "Optimal+1").



Figure 4: Subject evaluating custom HUTs. Backgrounds removed with Microsoft Copilot.

After donning, the following activities were performed for each HUT (Figure 4):

- **Mobility assessments:** The subject completed mobility tasks aimed at producing a variety of waist, torso, and shoulder postures. The subject's kinematics were recorded with a Vicon motion capture system. These tasks also provided subjects with the opportunity to get a perception of how the HUT fit in a more dynamic condition before providing subjective feedback.
- **Reach envelope generation:** The subject moved their arms in up-and-down and side-to-side sweeping motions. The motion capture system recorded the paths of optical markers placed on the subjects' elbows and middle fingertips. Comparing reach envelopes within each subject across their three HUT sizes provided insight into how the reach capability was affected by HUT condition.
- **3D body scanning:** To evaluate the HUT fit, it was necessary to confirm the HUT was sitting on the subject's body as expected—or, in other words, to verify the indexing of the HUT. Each subject was laser scanned in the Vitus whole-body laser scanner while wearing each of their HUTs. Scans of the subjects posing in various postures were collected to ensure the HUT did not move relative to the body more than expected. The resulting scans were overlaid with the initial model results for comparison.
- **Subjective feedback surveys:** For each HUT, the subject completed a suit fit survey twice. The first was collected immediately after the subject donned the HUT to represent a static condition and record initial perceptions. The surveys were conducted again after the subject completed mobility tasks to collect their perceptions from the dynamic condition (i.e., when the subject had moved around while wearing the HUT).
- **Subjective contact pattern mapping:** Immediately after completing a subjective feedback survey (i.e., both before and after the dynamic mobility tasks), the subject was presented with a digital manikin displayed on a touchscreen tablet (Figure 5). The subject used a stylus to draw on the manikin, color-coding regions based on the level of contact felt between their body and the current HUT. Each subject created a total of six contact maps– two for each of the three HUT sizes.

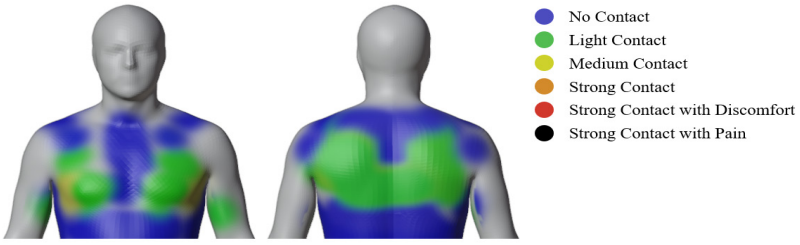


Figure 5: Subjective contact pattern mapping example.

RESULTS

This paper focuses on the reach envelopes and the subjective feedback assessments conducted through Thematic Analysis (TA).

Reach Envelopes

For each subject-HUT combination, the elbow and fingertip reach envelopes were generated by radial basis function approximation of the raw marker traces. The elbow location was defined as the midpoint between the lateral and medial elbow markers. Each marker path was down sampled and template-fitted to a parametrically deformed hemisphere (Figure 6).

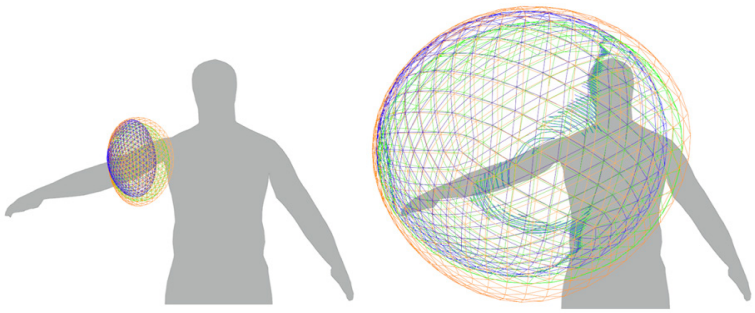


Figure 6: Reach envelopes for one subject. "Optimal" = Green, "Optimal-1" = Orange, "Optimal+1" = Blue. (Left) Elbow reach envelopes; (Right) Fingertip reach envelopes.

As a measure of reach capability, the surface area of each subject's reach envelope was calculated and compared between the conditions. Because reach envelope size relies heavily on subject anthropometry, it is necessary to evaluate each reach envelope metrics against an established baseline within each subject rather than between subjects. Defining each subject's "Optimal" HUT condition as the baseline, the change in envelope size was computed for each HUT (Table 2). Because the "Optimal" condition is used as the baseline, its change in reach computes to 1 for every row and thus is not included in the table. For elbow and fingertip reach envelopes, the "Optimal+1" HUT envelope had a smaller surface area than the "Optimal" HUT. The "Optimal-1" HUT's reach envelope was slightly smaller for the elbow and larger for the fingertip.

Table 2: Percent change in reach envelope surface areas (Mean ± SD).

	Optimal-1 (Compared to Optimal)	Optimal+1 (Compared to Optimal)
Elbow Envelope	–0.57%±34.58%	–12.97%±24.48%
Fingertip Envelope	+9.32%±13.26%	–5.46%±13.54%

Cross-body reach distance was also measured as the maximum horizontal displacement of the elbow (or the fingertip) beyond the straight-forward line from the shoulder, then the ratios were compared against the baseline

"Optimal" HUT (Table 3). This metric quantifies the extent to which shoulder mobility was constrained by the geometry of the HUT. Compared to the "Optimal" HUT condition, the elbow reach was slightly smaller and the fingertip reach was larger in the "Optimal-1" condition. The "Optimal+1" condition saw a significant reduction in shoulder and upper arm mobility, resulting in a negative ratio for the elbow envelope. This indicates that subjects generally could not bring the elbow further medial than their acromion while wearing the "Optimal+1" HUT.

Table 3: Percent change in cross-body reach (Mean ± SD).

	Optimal-1	Optimal+1
Elbow Maximal Cross-Body Reach	–6.53±179.97%	–214.83±267.30%
Fingertip Maximal Cross-Body Reach	12.91±18.10%	–11.03±19.39%

Subjective Feedback

In the suit fit survey, subjects were asked to evaluate how tight or loose each HUT felt on a 5-point scale from –2 ("Very loose") to 2 ("Very tight"), with 0 being ideal (Figure 7a). Additionally, subjects rated each HUT for level of comfort on a 1-5 scale (Figure 7b), with 1 indicating least comfortable and 5 indicating most comfortable. The minimum, maximum, and average rating were compiled and compared for both the static condition (feedback before mobility tasks) and the dynamic condition (feedback after mobility tasks).

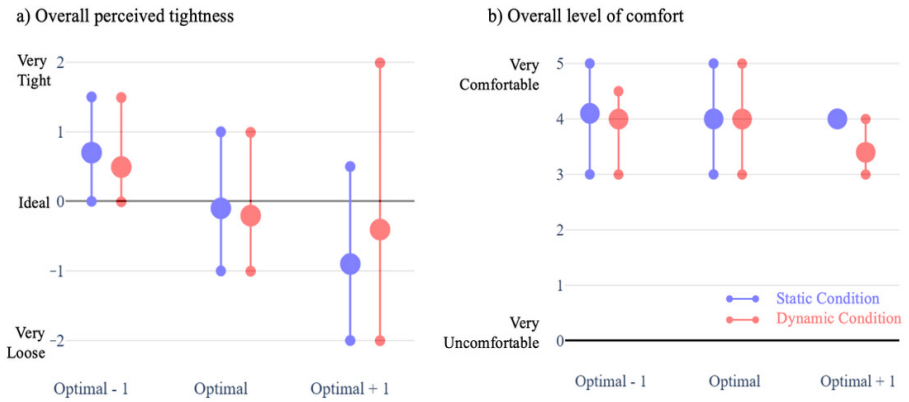


Figure 7: Subjective score distribution; circles represent min, mean, and max.

A repeated-measure analysis of variance (ANOVA) was conducted to characterize the impact of the HUT size and the task condition (i.e., dynamic vs. static) on the perceived tightness shown in Figure 7a. The perceived tightness relied almost entirely on HUT size ($p < 0.05$), whereas the task condition did not have a significant impact ($p = 0.69$). However, subjects perceived isolated "hot spots" and areas with strong suit-to-body contact increased while moving.

Assessing the overall comfort (Figure 7b) with repeated-measure ANOVA, the HUT size has a less pronounced impact ($p = 0.65$). The task condition, on the other hand, has a marginally significant impact on the reported levels of comfort ($p = 0.14$). This indicates that the comfort effects of an off-nominal fit may not be observed until the suit is evaluated in a dynamic condition.

In addition to the survey evaluation, subject feedback was compiled and Thematic Analysis (TA) was conducted (McLeod, 2024) to identify correlations between themes and fit conditions. To begin, the subjective feedback was consolidated and reviewed for recurring themes across HUT sizes. Five predominant themes were identified, as outlined in Table 4. The number of subjects who reported each theme was counted for each HUT size.

In general, subjects were more likely to identify themes while wearing a larger HUT (6 total themes identified in "Optimal-1", 11 in "Optimal", and 16 in "Optimal+1"). Notably, one of the two most prevalent themes in any suit configuration was "Excessive underarm contact," particularly in larger HUTs (5 out of 5 subjects reported the issue with "Optimal+1," and 4 out of 5 subjects reported with "Optimal" HUT). Additionally, although the "Optimal-1" size was preferred for comfort and mobility, two of the five subjects reported an inability to take a full breath in this configuration "Inhibited Respiration" theme). The "Optimal-1" size HUT is the only one that was reported to impact breathing.

Table 4: Recurrence of Identified themes in subjective feedback.

	Optimal-1	Optimal	Optimal+1
Excessive Underarm Contact	2	4	5
Noticeable Voids	2	4	5
Inhibited Shoulder Mobility	0	2	3
HUT Not Secure	0	1	3
Inhibited Respiration	2	0	0
Total	6	11	16

The themes were also related back to the suit-body relationship. Several subjects noted that their mobility seemed reliant on their shoulder indexing (i.e., whether their shoulders sat outside the HUT's scye bearing (i.e., arm opening). To evaluate this, the ratio between the subject's inter-acromion distance (IAD) and the HUT's "Q-distance" was computed for every subject-HUT combination (3 HUTs × 5 subjects = 15 combinations). The Q-distance is defined as the distance between the center of each scye bearing (Figure 8).

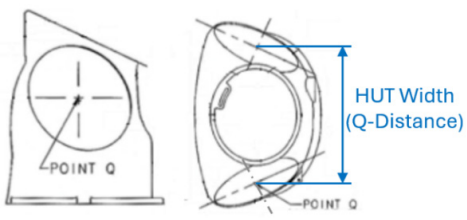


Figure 8: Point Q and Q-Distance on a HUT.

Theme prevalence was plotted in relation to the ratio of IAD to Q-distance (IAD/Q; Table 5), and combinations where "Inhibited shoulder mobility" was observed are represented with an X-marker (Figure 9). Notably, there appears to be an IAD/Q threshold. All tested subjects with IAD/Q of 1.11 or below experienced the "Inhibited shoulder mobility" theme, while all subjects with IAD/Q of 1.15 or higher do not. This implies that subjects were afforded greater shoulder mobility the further their shoulders sat outside of the HUT's scye bearing.

Table 5: Summary of theme count compared to IAD/Q.

Group	*N*	Inhibited Mobility?	Mean # of Themes	Std Dev
IAD/Q	5	100% (5/5)	3.20	± 0.45
IAD/Q	10	0% (0/10)	1.70	± 1.06

DISCUSSION

This study provided insight into evaluation methods for human testing of spacesuit mockups. Five HITL subjects each evaluated three custom HUTs. Laser scanning, motion capture, and surveys were used to capture a comprehensive picture of the resulting fit level for each subject-HUT combination. Novel elements of the evaluation methodology (namely, the contact pattern mapping activity) are already being incorporated into new studies.

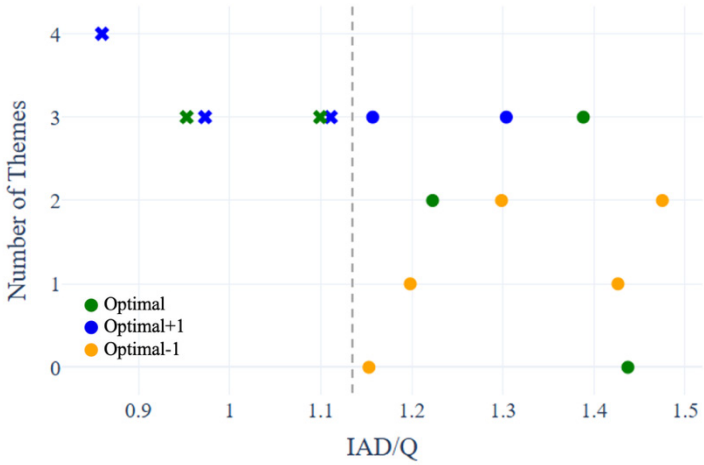


Figure 9: Number of themes plotted against the IAD/Q. Values greater than 1.0 indicate that inter-acromion distance is larger than the Q-distance (i.e., the acromion landmarks are outside the HUT's scye bearings). X-markers represent the subject reported inhibited shoulder mobility.

One unique element in this study is the evaluation of subject experience in both a "static" and "dynamic" condition. Subjects were significantly more likely to find a HUT comfortable in a static condition, indicating that

elements of spacesuit fit that relate to comfort may not be identifiable until the subject can experience it in a dynamic condition. This highlights an important consideration for full suited testing and mockup testing alike—fit evaluations should not end at the static condition.

New procedures were developed for this study to map levels of contact for test subjects in suited tests (Figure 5). While previous methods relied on segment-wise rating systems, this new method allows subjects to mark specific areas of note with granularity, improving the fidelity of subjective data. This new procedure is currently being implemented into other NASA studies.

Subjective evaluations were also structurally analyzed, primarily through the subjective contact pattern mapping and Thematic Analysis. The "Optimal+1" HUT elicited more themes that can increase injury risk. In particular, the "Excessive underarm contact" theme could reasonably be anticipated to increase injury risk as hard contact with suit components is a significant risk factor for suited injury (Dunn et al., 2022). Furthermore, because the existing Shuttle EMU HUT is not expected to accommodate crew with critical anthropometric measurements below 40th percentile female, it is likely that a significant amount of the crew population is wearing an oversized HUT and are at increased risk of injury due to excessive hard contact from being in a HUT too large. Furthermore, this segment of the crew population could be more prone to experiencing other themes associated with wearing too large a HUT, such as the hardware not feeling secure, inhibited shoulder mobility, and gaps where there is not contact between the suit and the subject's body.

This study has several important limitations. Each HUT was statistically generated using a model that correlated body geometry, suit geometry, and level of fit in historic NBL fit checks. Because only three HUT sizes were developed, the entire anthropometric range tested in this study was based on geometric models based on Medium, Large, and Extra-Large HUT sizes; thus, the tested range of the HUT size variation is fairly limited. Additionally, there are suit elements that were not represented by the 3D printed components, including adjacent suit components (e.g., softgoods, arms, waist-brief, etc.) and ancillary hardware (e.g., drink bag, liquid cooling and ventilation garment), which are known to substantially influence fit and comfort. Future work can incorporate these elements into the study. The study participants were not crewmembers, so they lacked EVA or realistic suit operation experience, which may have limited their evaluation.

Additionally, donning and doffing HUT hardware was not evaluated in this study. The excessive width of larger HUTs was identified as a risk factor for suited activity, but because the EMU utilizes a waist-entry HUT architecture, significant width is required to safely don and doff the HUT. A potential avenue for future work would be to consider the balance between don/doff capability and suit fit after donning.

CONCLUSION

This study developed new methods for structured human testing and fit modeling that can be leveraged in future suit fit assessments. A complete population analysis can be conducted to identify how much of the crew population is accommodated by existing hardware and how many crewmembers may be forced to use non-optimally fitting equipment. This study's novel HUT generation methodology acts as proof of concept for a custom spacesuit design framework, which can help enable future advancements in improved sizing schema design or a new suit architecture with components fully individualized for each crewmember.

ACKLNOWLEDGMENT

This study was supported by NASA Human Research Program (HRP) Human Factors and Behavioral Performance Element (HFBP) Directed Study.

REFERENCES

Dunn, J., Benson, E., Norcross, J., and Newby, N. (2022). "Risk of Injury and Compromised Performance Due to EVA Operations." *NASA Evidence Report*.

McLeod, S. (2024). Thematic Analysis: A Step by Step Guide. https://www.simplypsychology.org/thematic-analysis.html

NASA. (1999). Validation of Extravehicular Mobility Unit (EMU) Planar Hard Upper Torso (HUT) Capability. NASA Lessons Learned Information System. https://llis.nasa.gov/lesson/1114

NASA. (2019). The Extravehicular Mobility Unit (EMU). https://www.nasa.gov/image-article/extravehicular-mobility-unit-emu/

Scheuring, R. A., Mathers, C. H., Jones, J. A., & Wear, M. L. (2009). Musculoskeletal Injuries and Minor Trauma in Space: Incidence and Injury Mechanisms in U.S. Astronauts. *Aviation, Space, and Environmental Medicine*, 80(2), 117–124.