Evaluating the Effect of Spacesuit Glove Fit on Functional Task Performance

by

Seamus Lombardo

B.S., University at Buffalo, State University of New York (2018)

Submitted to the Department of Aeronautical and Astronautical Engineering
in partial fulfillment of the requirements for the degree of
Masters of Science in Aerospace Engineering
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
May 2020

Author . . . .
Department of Aeronautical and Astronautical Engineering
May, 2020

Certified by . . . .
Leia Stirling
Visiting Associate Professor
Thesis Supervisor

Certified by . . . .
Kevin R. Duda
Principal Engineer, Space and Mission Critical Systems, Draper
Thesis Supervisor

Accepted by . . . .
Sertac Karaman
Chair, Graduate Program Committee

Evaluating the Effect of Spacesuit Glove Fit on Functional Task Performance

by

Seamus Lombardo

Submitted to the Department of Aeronautical and Astronautical Engineering
on May, 2020, in partial fulfillment of the
requirements for the degree of
Masters of Science in Aerospace Engineering

Abstract

As the number of suited operations per mission increases with exploration beyond low Earth orbit (LEO), it is essential that crewmembers conduct suited activities in a manner that enables acceptable performance and minimizes the risk of injury. Currently, knowledge gaps exist in how to define optimal suit fit, how to more effectively incorporate fit into the suit design process, and how fit is related to performance. While it is understood that fit influences suited performance, the relationship between fit and performance has not been quantified. This research effort investigates the effects of spacesuit glove fit on tactile, dexterous, cognitive, and technical flight performance. This study adapted functional performance tasks from the literature and developed novel tasks to assess performance. Through these tasks, the hypothesis that static fit (as derived from glove and human anthropometry dimensions) is related to performance in spacesuit glove was evaluated.

Subjects wore prototype gloves, developed by David Clark Company, Incorporated (DCCI). These gloves are similar to the DCCI Orion Crew Survival System intravehicular activity (IVA) gloves that will be utilized on NASA's Orion spacecraft. Participants completed a battery of functional assessment tasks in a glovebox vacuum chamber (4.3 psid). The subject's prescribed fit within the DCCI glove sizing scheme specific to this design was determined using their anthropometry. The subjects then conducted the tasks in gloves one size below their prescribed fit, their prescribed fit size, and gloves one size larger than their prescribed fit in both a pressurized and unpressurized state.

To evaluate general tactility, blindfolded subjects attempted to detect bumps of different widths (0.59 in, 0.39 in, 0.20 in) and heights (0.05 in, 0.20 in,0.39 in) while the correct detection was recorded. An operationally-relevant tactility task was also designed. A mock spacecraft control panel was created in consultation with subject matter experts and designed to NASA specification. Blindfolded subjects then actuated a pre-defined sequence of these controls on the switchboard. The accuracy and completion time of the sequence was recorded. To evaluate general dexterity, subjects completed a pegboard task, which required moving and rotating pegs between locations on the board. Dexterity was also measured using a functional tool task where subjects attached and detached an extravehicular activity (EVA) tether hook to fixtures designed to NASA specification. For both dexterity tasks completion time was recorded. The Draper real-time performance metrics workstation lunar landing simulator was used to assess technical flight performance and mental workload (through a secondary task response time measure).

It was found that direct measures of static fit derived from hand length and glove length had a significant relationship to performance on the switchboard tactility task. Additionally, it was found that in the unpressurized case, subjects performed significantly better on the switchboard task when wearing the glove size larger than the prescribed fit as compared to small and prescribed fit. No consistent significant relationships with respect to glove sizing were found for the dexterity tasks or the the lunar landing simulator task. This study also reaffirms tactile and dexterous performance decreases with a spacesuit glove pressurization, with tactile performance also decreasing with the addition of unpressurized gloves over barehanded conditions.

Thesis Supervisor: Leia Stirling
Title: Visiting Associate Professor

Thesis Supervisor: Kevin R. Duda
Title: Principal Engineer, Space and Mission Critical Systems, Draper

Acknowledgments

This research effort would not have been possible without the consistent contributions of many individuals at MIT and in the broader human spaceflight community. I first have to thank my amazing advisor Leia Stirling for her support and mentorship. Not only were you always willing to provide insightful and detailed guidance on the technical aspects of my project, but you were also always available to help me through stressful situations, provide career advice and always did what was best for me as a student. I can't understate the value of your mentorship and it has been an honor to be your student. I must also extend my sincere thanks to Kevin Duda, my Draper mentor. Whether it was helping me navigate the process to make the best use of Draper's resources, providing me with expert advice on my project from your experience, or just chatting with me in the Atrium about how my project was going, your mentorship was important to my success.

There are also many people I must thank in the broader aerospace community. To Shane Jacobs at David Clark, thank you for the essential resources you provided, the advice you gave during the planning process, and helping me through more than one stressful situation during testing and data collection. This project would not have been possible without your generous time and effort. To the folks at NASA JSC (Shane McFarland, Dick Watson, Andrew Abercromby, Omar Bekdash, Sarah Jarvis and others) thank you for always being available to answer my questions whenever I was confronting a new challenge. Additionally, to the staff at Draper (Steve York, Alison Siegman, and Sherrie Holder), thank you for helping me through all the bumps in the road I encountered when using the Lunar Landing Simulator. To Sheila and Martha, the support, funding, and resources provided through the Draper Fellowship program were essential to my success throughout the entirety of my master's program.

Here at the Institute there have also been many others who have helped bring my project to fruition. To Maddie Garcia, Katya Bezugla, Thomas Rick, Mary Dahl, and Thomas Adebiyi, your work as UROPs was essential to completing this project and I could never have finished it without your committed help. To Beata, Liz and

the other support staff, thank you for helping me order supplies, figure out travel, and for always working tirelessly behind the scenes to make sure all of our research runs smoothly. Additionally, to everyone in the aeroastro machine shop, The Deep, and the COUHES office, your support is helpful in bringing many students' projects to fruition.

In addition the aid I received on my project itself, I have also had a tremendous support network to help me get through the trials of graduate school. To my friends in the HSL, thank you for being such an awesome community. Our talks in the office, group Halloween costumes, parties, and movie nights played a huge role in keeping me sane. To my friends from Nanuet, UB and elsewhere, staying in touch with all of you was a way for me to de-stress and stay grounded during stressful graduate school experiences.

To my family, you have always been there for me no matter what and these past two years have been no exception. Whether its a care-package filled with sweets, a big hug on a visit home, or sharing a meal in Boston or Nanuet, I couldn't have gotten through this without your help. Lastly, to Niamh, for the past 8 years you have been my biggest supporter and best friend. I know these past two years have been stressful for a multitude of reasons, but you always stood by my side, helped me cope, and brought me back down to earth no matter what was happening. Your love and support means the world to me and is essential to everything I do.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The following section first describes the motivation for this research effort by outlining the coming increase in suited activities with exploration missions and the knowledge gaps related to the relationship between spacesuit fit and task performance. A literature review is then conducted that surveys the state of the field on categories of fit, spacesuit testing, gloved performance assessment and previous work on the relationship between fit and performance. Lastly the specific aims and hypotheses of the study are laid out in the final section.

1.1 Motivation

As the world's space agencies and private companies begin to move beyond low earth orbit (LEO) to explore the moon and Mars, the use of spacesuits will increase beyond the current levels associated with International Space Station (ISS) operation [2]. NASA's Human Research Program reports that during future missions to planetary surfaces, crewmembers could perform up to 24 hours a week of suited activity, a substantial increase over the maximum of 4 extravehicular activities (EVAs) for a shuttle mission and the moonwalks performed during the Apollo era (of which there were only 15 across all of the Apollo missions). The tasks performed on future missions will be diverse and complex, and could include operating spacecraft, conducting scientific investigations, constructing equipment and maintaining hardware and mission infrastructure. EVAs will require astronauts to wear spacesuits to provide life support and protect from these new environments, necessitating further spacesuit development to facilitate this increase in suited activity. The spacesuits of the Apollo era were noted for their limited mobility and reduced dexterity [2]. The current Extravehicular Mobility Unit (EMU) is designed for ISS operations, limiting lower extremity mobility for the microgravity environment, but providing upper extremity mobility for ISS EVAs. Although many successful missions have been completed in the EMU, injuries have resulted from use of the suit during training as well as during mission operations [3]. These injuries are of importance because similar incidents could adversely affect mission goals on a planetary mission.

To meet the challenge of increased suited operations, it is important that crewmembers be able to perform their suited tasks in a manner that enables acceptable performance for frequent and complex tasks while also minimizing injury risk. One important factor in this performance is spacesuit fit. NASA acknowledges that proper suit fit is a critical element in preventing astronaut injury [2]. Additionally, the NASA Human Research Roadmap [4], acknowledges the role of spacesuit fit in gap EVA 7B, stating "How does suit sizing and fit affect crew health and performance in exploration environments?" Knowledge gaps currently exist with respect to how to design for optimal spacesuit fit and how spacesuit fit is related to task performance. While spacesuit fit has long been incorporated into the development process in a subjective manner, there is no consensus on methods for determining optimal suit fit [1, 2, 5]. Additionally, while the relationship between fit and performance has been examined in other domains (including body armor and protective overalls) [6, 7] and it is acknowledged that spacesuit fit plays a role in crewmember performance [2, 8], this relationship has not been quantified.

Spacesuit fit is further complicated by the phenomenon that what is a "good" fit may change based on the gravitational environment. In March 2019, NASA had to cancel the highly anticipated first all-female spacewalk on the ISS as astronaut Anne McClain decided that she would perform better utilizing a medium-sized EMU torso, even though two of the medium-size torso were not prepared for use [9]. Although

McClain had trained in the prepared large-size torso on Earth, a previous spacewalk in microgravity changed her perception of which suit would facilitate her best performance. This situation highlights that the need to further understand the relationship between fit and performance (including how this relationship changes with different environments and tasks) is necessary for mission success. This thesis investigates this relationship between spacesuit glove fit and performance towards the goal of enabling improved functional task performance on future space missions.

1.2 Literature Review

The following section reviews the literature relevant to an examination of the relationship between spacesuit glove fit and task performance. First the different categories of fit metrics are described including static, dynamic, and cognitive fit. An overview of both spacesuit testing and gloved performance assessment across multiple fields is then provided. Lastly, past studies into the relationship between fit and performance in other domains are detailed.

1.2.1 Categories of Fit

Fit can be defined at a high level as "an optimized relationship between the human and the environment" [10], where the environment often refers to the garment or exosystem (such as an exoskeleton or a spacesuit [11]) that the human is wearing. However, it is important to note there are several categories of fit that describe the way an external system relates to a human. These categories (static fit, dynamic fit, and cognitive fit) describe different aspects of fit with relation to human movement as well as different areas of performance.

Static fit

Static fit is defined as the relationship between the human and the relevant garment in a given static posture. Direct measures of this category of fit generally makes use of the anthropometric measurements of the human and the dimensions of the garment or exosystem. One method with which static fit is measured is “ease”, which is defined as the difference between a given anthropometric measurement and the relevant portion of the surface of the garment or exosystem [12]. Scalar anthropometric measurements relevant to static fit (e.g. hand length) can be measured using traditional tools such as calipers or tape measures. Direct measures of static fit can also be assessed utilizing 3D scanning and volumetric measures, such as work by Choi et al. [10] who aligned 3D scans of a nude figure with 3D scans of a clothed figure to assess the relationship between the human and a garment for a given cross section.

Static fit can also be assessed subjectively, where the wearer of the garment provides input on their perceived fit. Subjective static fit is often assessed through a questionnaire given while wearing the garment. These questionnaires range in fidelty from broad subjective ratings (ex:“too tight” and “too loose”) [13] to more in-depth numerical scales that assess both indexing and feature alignment [5]. These direct measures of static fit are often combined with subjective measures to provide an overall assessment of static fit [14].

Static fit is an essential way of assessing that a given garment or exosystem matches the dimensions of the wearer. However, it is important that the given static postures used to assess static fit be functionally or operationally relevant [11]. Additionally, the fit of the wearer often changes between different postures or during movement, necessitating other categories of fit assessment described in the following section.

Dynamic Fit

Dynamic fit is defined as the interaction between the human and the garment or exosystem during movement [15]. This category of fit is important because the relationship between the human and the garment can change for different postures and movements. Understanding this relationship during dynamic motion can help increase mobility and decrease the exerted forces necessary to perform a given task [11]. Additionally, misalignment between the human and a garment during movement can cause unnecessary fatigue and potentially lead to injury. Specifically, many spacesuit injuries (such as abrasions or muscle soreness) are a result of interactions between the human and the spacesuit [16] and it is therefore essential to assess this category of fit.

Dynamic fit can be assessed in a wide variety of ways due to the diversity of motions and garments that humans utilize. One metric of dynamic fit is range of motion (ROM) which can be measured in many ways. Choi et al. used standard anthrompmetry tools to assess ROM in body-armor [6], Reid et al. [17] utilized a 3D motion capture system to assess ROM in the EMU, and Fineman et al. [15] used wearable inertial measurement units (IMU) to assess ROM in the lower extremities of the Mark III spacesuit. Schmidt et al. also used human subjects and a specialized robot to measure joint torques and joint angles in the EMU [18]. The use of functional assessment tasks to indirectly evaluate dynamic fit has also been used, such as the Load Effects Assessment Program which uses mobility assessment tasks to evaluate dynamic fit of soldiers' equipment [19] and work by Cullinane et al. [20] that utilized 3D motion capture to assess gait parameters in the Mark III spacesuit during walking tasks. Dynamic fit can also be assessed by measuring the physical interaction between the human and garment or exosystem. Anderson et al. [21] utilized pressure sensors to assess human-spacesuit interaction in the Mark III spacesuit during a series of movements and functional tasks. Compton et al. [22] evaluated conductive materials that can be integrated into garments that could be used to assess dynamic fit by measuring contact between the garment and body or spacesuit. Lastly, relative motion between a garment or exosystem and the wearer has also been utilized to assess dynamic fit, with Fineman et al. [15] utilizing wearable IMUs to assess coordination between the human and the Mark III spacesuit in the lower extremities and Lombardo et al. [23] utilizing similar metrics to evaluate dynamic fit in the EMU torso.

Cognitive Fit

Cognitive fit is defined as the ability of a exosystem or garment to support the perception-cognition-action decision process of the wearer [11]. Understanding cognitive fit is important so that the given garment or exosystem does not inhibit the wearer's decision making process or ability to perform operational tasks.

There are several different ways to examine the concept of cognitive fit. One is somatosensation which pertains to the tactile perception and proprioception [11]. This form of cognitive fit can be assessed using methods such as the Semmes-Weinstein monofilament test [24] which facilitiates the evaluation of peripheral nerve sensory function, or robotic assessment of movements to evaluate proprioception [25]. Another aspect of cognitive fit is motor action selection, or the ability of the wearer to map motor commands to goals based on internal cognitive models of the situation which can be measured by assessing motor trajectories, response times and accuracy [11].

Of particular importance to this research effort is the concept of executive function with respect to cognitive fit. Executive function is defined as the cognitive processes that "enable goal-directed behavior, including inhibition of behavior, working memory, and cognitive flexibility" [11]. One important component of executive function is mental workload which is a measure of the level of attentional resources required for an operational task [26]. The concept of mental workload describes the relationship between attention and performance. The nature of this relationship can be described by the Yerkes-Dodson law, where an inverted u-shaped curve shows low performance at both low and high attention states [27]. In high workload situations, the ability to shift focus to extraneous tasks or the environment is reduced [26]. Also important to understanding mental workload are the concepts of capacity and resource models. Kahneman [27] described a model where there is a general capacity for the amount of mental work a human can do, but that this capacity can be divided up among different tasks in different amounts (known as the Single Resource Model). This capacity-based model leads to the concept of using a secondary task to assess mental workload. A person can distribute their total attentional capacity, with part of the capacity required to perform the primary task. The spare attentional capacity can be measured via performance a secondary task. Wickens [28] further elaborated on this concept by introducing the Multiple Resource Model, wherein different attentional resources exist for different sensory modalities (such as visual or auditory inputs) and domains (such as spatial and verbal). This concept therefore influences the choice of primary and secondary task used to assess mental workload in order to ensure that they are assessing the appropriate category of attentional capacity.

A garment or exosystem that requires increased attention to utilize, such as a pressurized spacesuit with poor fit or limited biomechanical degrees of freedom, can cause increased mental workload that inhibits performance on both the primary functional task under consideration as well as secondary tasks. One type of secondary task used to assess mental workload are reaction time tasks [29, 30]. In these tasks, the response time to an indicator is measured to assess the mental workload required to perform a primary task for a given scenario. In addition to these aforementioned secondary tasks, subjective measures [28] such as the NASA Task Load Index (TLX) [31] can also be used to assess mental workload.

Another component of executive function is situation awareness. Endsley [32] breaks down the concept of situation awareness into three levels: 1) perception of elements in the current situation; 2) comprehension of the current situation; and 3) projection of future status. Maintaining all 3 levels of situation awareness is important in spacesuit operations, such as the monitoring, understanding and future projection of spacesuit life support consumables. There are multiple ways with which to measure situational awareness, including explicit, implicit, and subjective techniques. Implicit techniques are those derived from task performance and subjective techniques are those where participants rate their perception of their own situation awareness [33]. Explicit techniques directly assess the operator's perception of a given situation. One example of explicit situation awareness measures was employed by Ma et al. [34] by freezing a driving simulation and posing queries about the driving situation. Challenges such as the intrusiveness of probes administered during a simulation freeze or the additional workload imposed by a concurrent probe are associated with explicit techniques. However, explicit techniques more directly assess situation awareness when compared to implicit techniques (where poor task performance may not necessarily be a function of situation awareness) and subjective techniques (where survey responses may be affected by subject confidence). The lunar landing simulator employed in this study used the explicit technique of concurrent verbal call-outs of the simulation state assessed for accuracy and timing via automatic speech recognition [35].

Interaction Between Fit Categories

It is important to note that all three of the aforementioned fit categories interact with each other in various ways. Examples of these interactions include those between static fit and dynamic fit. Changes in ease (a measure of static fit) have been shown to affect ROM (a measure of dynamic fit) for a variety of garments such as body armor [6] and protective overalls [7]. Additionally, there are interactions of cognitive fit with static and dynamic fit. These interactions are evidenced by results by Bequette et al. [36] who found that despite adjusting an exoskeleton to physically fit the wearer, decrements of cognitive fit as measured by visual response times were still observed for some subjects. For current planetary spacesuit prototypes, limited ROM results indicate reduced dynamic fit (such as in the Mark III hip assembly as modeled by Cowley et al. [37]), but also reduced cognitive fit as the altered gait that is caused by the suit can increase mental workload [11]. Crewmembers have specifically mentioned that EVA operations are both physically and cognitively demanding, demonstrating the necessity of assessing all the aspects of spacesuit fit [38]. The interactions between these various categories of fit as well as their effects of performance, can be evaluated through spacesuit testing.

1.2.2 Spacesuit Testing and Evaluation

Spacesuits have baseline functional demands that are necessary to allow humans to survive the harshness of the space environment, performance demands to allow astronauts to conduct their mission effectively, and must also mitigate injury risk during operations. Spacesuit testing evaluates many baseline life support functions such as pressurization, waste removal, and thermal regulation that are necessary to sustain human life in space. Suits must undergo structural testing to ensure that they can endure the mechanical loads associated with suit pressurization (such as recent testing of the Orion Crew Survival System (OCSS) Intravehicular Activity (IVA) suit [39]) and that the composition of gases within the suit maintains approriate levels of CO2 and oxygen (such as the recent CO2 washout study performed by Bekdash et al. [40]). To ensure the removal of human waste, the integrated testing of waste removal systems is part of the suit development process (including the recent testing OCSS urine removal system while a pressurized suit was worn by subjects [41]). The thermal regulation and cooling systems of the suit are also important to sustaining human life and their evaluation was incorporated into testing of a planetary suit during simulated EVA operations by Watts et al. [42].

Beyond these baseline life support functionalities, spacesuits also undergo testing before missions to ensure acceptable human performance. These testing efforts have examined specific areas of human performance within the spacesuit such as strength and energy expenditure, both of which were evaluated by Reid et al. [17] in an evaluation of the EMU torso. Carr et al. also examined the energy expenditure associated with locomotion in spacesuits using a lower body exoskeleton [43]. Of particular interest to this research effort are evaluations of spacesuit performance measures that relate to spacesuit fit. Efforts to evaluate mobility and range of motion are related to the investigation of spacesuit fit. Cullinane et al. [20] examined spacesuit mobility using 3D motion capture and Fineman et al. [15] utilized IMUs to evaluate ROM, with both studies examining the Mark III spacesuit during a walking task. Mobility and ROM have also been examined for the upper extremities in both the Mark III and David Clark Company, Inc. (DCCI) Mobility mockup suit utilizing IMUs [44] and in the EMU hard upper torso (HUT) during a sizing investigation using 3D motion capture [17]. Mobility evaluations tasks were also part of the pre-delivery testing done by ILC Dover for the Z-2 spacesuit [45]. Metrics of relative motion have been employed to evaluate spacesuit fit using IMUs both by Fineman et al. [15] in the lower extremities of the Mark III and by Lombardo et al. [23] in the torso of the EMU. Spacesuit fit has been subjectively evaluated through the use of surveys that evaluate indexing, feature alignment, and mobility [5].

In addition, to pre-flight baseline and performance testing of spacesuits, it is also important to examine post-mission evaluations of spacesuit related injuries. A range of injury mechanisms have been examined in spacesuits including fingernail delamination [46] and shoulder injuries which occur primarily during suited training on Earth [16]. Injuries specifically related to spacesuit fit include blisters and abrasions resulting from shear pressures due to unwanted interactions between the human and the suit, as well as joint injuries that result from feature misalignment [16]. In-suit sensing systems to gain a better understanding of these injury mechanisms are actively under development and make use of pressure sensors to measure contact points [47] and IMUs to assess indexing and joint alignment [48].

While this subsection provides an overview of the functional and performance testing of spacesuits, as well as evaluations of injury risk, it is important to further evaluate how to assess performance in gloves, which are the garment under examination in this study.

1.2.3 Glove Performance Assessment

Spacesuit gloves are a critical suit component as they allow the astronaut to manipulate their environment, permitting completion of relevant operational tasks. If these tasks are not able to be completed, the mission could be compromised [2]. Astronauts on the Apollo, ISS, and Space Shuttle missions have reported that spacesuit glove use led to hand fatigue, pointing to an area for improvement in the spacesuit development process. Another factor related to spacesuit gloves is injury risk, as hand injuries such as fingernail delamination have been reported, as well abrasions, contusions and perhipheral nerve impingements which are related to glove fit [2, 16, 46]. As spacesuit gloves are critical to performance on human spaceflight missions and also an area of the suit that can cause fatigue or injury, it is essential to understand the relationship between spacesuit glove fit and task performance. This section surveys the literature on how best to assess gloved task performance and on previous work evaluating spacesuit gloves.

General Considerations for Gloved Performance Assessment

For any evaluation of spacesuit glove fit on performance, relevant measures of performance must be selected. Gloved performance has been assessed in many ways, including muscle load and fatigue, functional tasks, and subjective comfort measures [49]. Another common way to assess gloved performance is the use of functional tasks [49]. No agreed upon set of tasks with which to assess gloved performance exists, although methods across different industries agree on the importance of assessing key areas such as dexterity, tactility, mobility, and grip strength [49]. The relevance of the performance assessment task to a given setting or job is also an important consideration, and both generalized tasks (such as a pegboard dexterity task) and operationally relevant tasks (such as a data entry or assembly tasks) have been employed in the literature [49].

Past research efforts into spacesuit gloves specifically have utilized a wide variety of performance assessment techniques ranging from simple pegboard based dexterity tasks to the use of external motion capture systems and specially designed sensor gloves [1, 50, 51, 52]. While some studies have opted for tasks that mimic relevant operations [51], others seek to use more general evaluation methods [1].

Dexterity Performance Assessment

One key metric used to assess gloved performance is manual dexterity. Manual dexterity is defined as ”a motor skill that is determined by the range of motion of arm, hand and fingers and the possibility of manipulation with hand and fingers” [53]. It is thought that factors such as restricted finger movement and bunching of glove material inhibit manual dexterity when gloved. A variety of functional tasks have been employed in the literature to assess dexterity including pegboard tasks, nut and bolt assembly tasks and rope tying tasks [10]. Generally, results across the literature show that the use of gloves in various domains results in dexterity performance decrements as measured by task completion time, although the level of performance decrement varies based on the glove used [49].

Dexterity in spacesuit gloves has also been assessed using functional tasks. O'Hara et al. [54] utilized a pegboard task, nut and bolt assembly task and a knot tying task to assess the dexterous performance of EVA gloves. This work found that that unpressurized glove conditions led to decrements in dexterity performance over ungloved trials in EVA gloves as measured by task completion time when task rates were not specified. The study also supported that the pressurization of the gloves further reduced performance over the corresponding unpressurized trial [54]. Newton et al. [52] assessed dexterous performance in an unpressurized IVA glove from Final Frontier Designs and found that the spacesuit glove led to a decrease in performance on a Lafayette Purdue Pegboard Task, Lafayette Hand Tool Dexterity Test and rope-tying task over ungloved conditions. As part of NASA's High Performance EVA Glove (HPEG) evaluation study, Korona et al. [1] utilized a pegboard task, a knot tying task, and a bow tying task (where task rates were not specified) to assess the performance of the current EMU glove as well as new prototype EVA gloves from DCCI and ILC Dover. The results showed that the pressurized EVA gloves resulted in significant decreases in dexterous performance over ungloved conditions, however these tasks were unable to discern differences in dexterous performance between the 3 protoype gloves [1].

Tactility Performance Assessment

Tactility is a common attribute that is evaluated when assessing gloved performance and is defined as ”sensitivity to texture, size, shape and other attributes that can be sensed through touch and the ability to detect changes in any of these attributes” [55]. General evaluation tasks for tactility include two point discrimination tests (where the smallest spatial threshold that can be detected is assessed), monofilament tests (which evaluates the ability to perceive filaments of various size with the sense of touch), and shape or object identification tests [49]. When surveying the results of these studies, some have found that gloves reduce tactile performance while others have found that the gloves have no effect [49].

With regards to tactility in spacesuit gloves, O'Hara et al. [54] found a substantial performance decrement when wearing EVA gloves on a two point discrimination task as measured by detectable gap width when comparing gloved cases to ungloved cases, and also that the addition of glove pressurization led to a decrease in performance over the unpressurized case. Additionally, O'Hara et al. [54] conducted a shape detection task and found that shapes were identified correctly 80 percent of the time in the unpressurized gloved case and 85 percent of the time in the pressurized glove case, versus a higher success rate of 97.5 percent in the ungloved case (however the authors note that there were larger differences associated with the size of the shapes than the gloved condition). Thompson et al. [56] found that for a resin bump detection tactility task, the average force needed to correctly detect resin bumps using sense of touch while blindfolded increased in the pressurized case over the unpressurized case.

Flight Technical Error Performance Assessment

NASA has long recognized the importance of utilizing flight simulators to prepare astronauts for flight, making use of high-fidelity simulators to train shuttle astronauts [57] and lunar landing simulators during the Apollo program [58]. It is therefore important to examine how flight performance is affected by spacesuit glove fit. To assess this performance attribute, this research effort examines technical flight performance on an operationally relevant lunar landing simulator task. While the lunar landing simulator used in this study has not previously been used to assess performance with respect to glove fit, Hainely et al. [35] utilized the simulator to assess the effect of flight mode on performance. The primary flight task involved nulling attitude error in the roll and pitch axes. It was found that root mean square error (RMSE) in the pitch axis was higher for simulator trials that contained a landing point redesignation and that RMSE was significantly higher in trials without a landing point redesignation where the pilot controlled all 3 axes and rate of descent, compared to trials where the pilot only controlled 1 axis.

Southern et al. [59] examined a range of performance measures (including gloved dexterity) while wearing a pressurized IVA suit and flying a Saab 2000 flight simulator. In the simulation, gloved performance was subjectively evaluated using a the

Modified Cooper-Harper rating scale where 1 is the most desirable rating and 10 is the least desirable rating. Overall the gloves received a rating of 2, corresponding to good performance with minor deficiencies. While this rating was provided, no additional qualitative insights into gloved performance were described in that study. Additionally, in further testing of IVA suits by Southern et al. [60], a flight simulator "busy board" that consisted of a throttle, joystick, switch bank, and flight simulator screen was employed while on a parabolic microgravity flight. While the objectives of this microgravity test were primarily targeted at assessing the suit's life support systems, an average subjective rating of 2.25 (where 1 is the best rating and 10 is the worst rating) was given by subjects with respect to their dexterous performance using the controls and it was reported that "busy board tasks were generally completed without issue." However, further research is required to specifically assess the effect of spacesuit gloves on flight technical error.

Cognitive Performance Assessment

The NASA Human Research Program acknowledges that inadequate spacesuit fit could lead to mental workload above acceptable levels [2] and it is therefore important to assess mental workload for representative tasks. As mentioned in section 1.2.1, secondary reaction time tasks have an extensive history of being utilized to measure mental workload [29, 30] , and these tasks are used in conjunction with flight simulators [61]. Hainely et al. [35] used a two-choice response time in conjunction with a lunar landing simulator and found a significant increase in mental workload associated with a mode transition from automatic control to a manual control mode, as well as significant increases in mental workload with increasing the number of control loops the pilot was responsible for closing during the flying task. [35].

With regards to gloved performance, Taylor and Berman [62] found that the addition of gloves impaired performance on a secondary manual tracking task using a joystick during a primary data entry task. The NASA TLX subjective assessment has been previously used to assess mental workload [31]. Using NASA TLX to assess workload, Llanos et al. [63] found that when wearing a pressurized IVA suit while flying a spaceflight simulator, 28 percent of subjects reported mental demands higher than moderate levels. It is important to note however, that no unsuited baseline trial was conducted in this study. Further research is required to specifically examine the the effects of spacesuit gloves on mental workload as assessed by a secondary response time task.

Spacesuit Glove Pressurization

Unique to assessing the performance of spacesuit gloves is the importance of pressurization. Previous spacesuit glove studies have varied in their approach to evaluating pressurized suit state, with some opting to have subjects perform tasks in a full suit [51], while others made use of a glovebox [1]. The benefits of using a full suit include a realistic interaction between the suit torso and arm components, as well as mitigating the discomfort and task difficulty that can sometimes accompany subjects whose anthropometry does not match well to glovebox dimensions. However, the use of the full spacesuit presents logistical challenges, such as hardware availability, the requirement for subject physicals, and the additional time required to don and pressurize the full suit. In contrast, gloveboxes allow the subject to easily don and doff the gloves, transition between glove fits, and pressurize and depressurize the gloves.

Although a lack of uniformity among assessment tasks makes direct comparisons across studies difficult, similar trends in performance can be seen between studies that utilize gloveboxes and those that utilize full spacesuits. Thompson et al. [56] who utilized a full EMU during their study, found performance decrements on a bump detection tactility task due to the addition of the spacesuit, with additional decrements resulting from pressurization. Similar trends in tactile performance were found by O'Hara et al. [54], where tactile performance decrements were observed due to the addition of spacesuit gloves and spacesuit glove pressurization in a glovebox-based study.

1.2.4 The Relationship Between Fit and Performance

Studies into the effects of changing static fit on task performance have been performed in a variety of fields. Choi et al. [6], examined the effect of different body armor sizing on warfighter range of motion (ROM) and found that while the decreased sizing did not significantly decrease mobility, increased sizing led to a significant decrease in ROM across all measured movements. In the area of firefighter garments, McQuerry [64] found that the static fit of structural turnout suits varied based on gender and had gender specific restrictions in mobility. Park et al. [65] found that there was a significant correlation between firefighter boot height and lower body mobility, with shorter firefighters likely to have limited mobility due to fixed boot height standards. In the domain of work garments, Huck et al. [7] examined the effect of crotch ease on ROM in custom protective work overalls (such as those that might be used in work like asbestos abatement) and found that adding crotch ease to the back of the overalls maximized wearer mobility in terms of trunk flexion.

Examinations of the effects of fit and sizing on performance have also be conducted in the domain of spacesuits. Reid et al. [17] utilized subjective, mobility, strength, and metabolic metrics to assess the effect of Hard Upper Torso (HUT) sizing in the EMU, however, minimal performance differences between the nominal size and the larger size were found. Fineman et al. [15] assessed the effect of changes in static fit through the addition of padding in the hip brief assembly of the Mark III on range of motion and relative coordination metrics but found mixed effects of padding on gait performance and these measures of dynamic fit. A preliminary investigation of dynamic fit measured via IMUs in the EMU HUT by Lombardo et. al [23] found that subjects with smaller arms moved their torsos more when suited than subjects with longer arms on a cycle ergometer task, suggesting that the relationship between the suit dimensions and subject anthropometry affected their task strategy.

The effect of static fit on performance has been considered in the context of protectives gloves. Tremblay-Lutter and Weihrer [12] defined their measure of ease as “the difference in space (i.e., length, girth, volume) between the outer surface of the hand and the inner surface of the glove". This study found an differences in completion time on the Minnesota Rate of Manipulation Turning Test, O'Connor Fine Finger Dexterity Test, Cord Manipulation and Cylinder Stringing Test, and Magazine Loading Test dexterity tasks. Results supported that decreased ease did not lead to a significant decrease in performance from the self-selected best fit condition, whereas increased ease consistently resulted in the slowest completion times. However, further research is needed to specifically assess the relationship between static fit and performance in spacesuit gloves. This research effort seeks to address this knowledge gap.

1.3 Aim

The aim of this research was to address knowledge gaps in the literature by assessing the hypothesis that there is an effect of spacesuit glove fit on task performance. This aim led to the design of a human study using a battery of both generalizable and operationally relevant performance assessment tasks and direct measures of static fit, which were used to quantify the effect of spacesuit fit on performance. A glovebox approach was used for the evaluations.

This work hypothesized that static fit metrics derived from gloved dimensions and human anthropometry are related to spacesuit glove performance on:

(1) A generalized tactility task where subjects attempt to perceive bumps of different dimensions.
(2) An operationally relevant tactility task where subjects attempt to perceive and actuate controls commonly found in spacecraft cockpits.
(3) A generalized dexterity task where subjects remove, reorient and reinsert u-bolts into a pegboard.
(4) A tool-based dexterity task where subjects hook and unhook an EVA tether hook to loops and a handle designed to NASA specification.
(5) An operationally relevant task assessing flight technical error and mental work-

load utilizing a lunar landing simulator.

Chapter 2

Methods

The following chapter describes the methods utilized to investigate the study's aims regarding the relationship between spacesuit glove fit and peformance. The various components of the experimental design are first detailed, descriptions of each of the performance assessment tasks are then given, and the statistical analyses utilized to evaluate the resulting data are outlined.

2.1 Experimental Design

This section provides a description of the study participants, the experimental protocol, the metrics employed to assess static fit, and the glovebox vacuum chamber used to conduct the study.

2.1.1 Participants

Nine subjects (n=9) participated in the study in total. Inclusion in the study required nominal static fit in the "prescribed" (P) size glove based on criteria from David Clark Company Incorporated (DCCI). Subjects were included if their hand length and hand circumference were within the range of the prescribed fit for the available 3 glove sizes. Subjects were excluded if they had any recent upper extremity impairments or if they had any auditory impairments that would prevent them from following verbal instructions. The study protocol was approved by MIT Committee on the Use of Humans as Experimental Subjects.

Out of the 9 total subjects, 7 subjects completed the tactility tasks. Two subjects did not perform the tactility tasks due to a leak in the glove preventing task completion. Of those 7 subjects, ages ranged between 23 and 27 (mean= 24.43, SD = 1.62), with 6 of the subjects being male and 1 being female.

With regards to the lunar landing simulator task, 8 subjects successfully completed the primary flight subtask. One subject was removed from the analysis of all subtasks for the lunar landing simulator due to an inability to meet completion criteria for the simulator task (several trials after the training run resulted in failure to successfully fly the spacecraft to the landing point). Of the 8 subjects who completed the primary flight subtask, ages ranged from 22 to 27 (mean = 24, SD = 1.78) with 6 subjects being male and 2 being female. An additional subject was unable to be included in the analysis of the response time subtask for the lunar landing simulator due to a failure to respond to any of the communication indicators for one of the gloved conditions. Of the 7 subjects who completed the response time subtask, ages ranged from 22 to 27 (mean = 23.71, SD = 1.70) with 5 male and 2 female subjects. Subjects were asked about their previous flight or flight simulator experience and their recorded answers ranged from no previous exposure to licensed pilots.

For the u-bolt pegboard dexterity assessment task, 8 subjects successfully completed the task. One subject did not perform the task due to glove leaks preventing task completion. For the 8 subjects who completed the u-bolt pegbord task, ages ranged from 23 to 27 (mean = 24.25, SD = 1.58) with 6 subjects being male and 2 being female. For the EVA tether dexterity assessment task, 7 subjects successfully completed the task. Two subjects did not perform the task due to glove leaks preventing task completion. For the EVA tether task, subject ages ranged from 23 to 27 (mean = 24.43, SD = 1.62) with 6 subjects being male and 1 being female.

2.1.2 Experimental Protocol

Subjects completed performance evaluation tasks in three IVA glove fit configurations to assess the effect of fit on performance. Prototype gloves, similar in design to the DCCI Orion Crew Survival System (OCSS) IVA gloves were utilized in the study. The subject's prescribed fit within the DCCI glove sizing scheme specific to this design was determined using their anthropometry. The subjects conducted the battery of tasks in gloves one size below their prescribed fit (S), their prescribed fit size (P), and gloves one size larger than their prescribed fit (L). The full set of tasks included a cognitive performance task (the Draper real-time performance metrics workstation lunar landing simulator [35] to assess mental workload and situational awareness as well as flight technical error), dexterity tasks (a general U-bolt pegboard task and operationally relevant EVA tether hook task), and tactility tasks (a general bump detection task and an operationally relevant switchboard sequence task) (Table 2.1). These tasks were performed across two days.

Table 2.1: Task Implementation and Metrics

Day	Task	Implementation	Metric
1	**Turntable Tactility Task**	While blindfolded, detected bumps of various width and height	Number of correctly detected bumps and false positives
	Switchboard Tactility Task	While blindfolded, use of IVA analogue control panel	Accuracy score for control sequence
	U-bolt Pegboard Dexterity Task	U-bolts are removed and inserted into the pegboard	Completion time per bolt
2	**Lunar Landing Simulator Cognitive Task**	Simulator with secondary communication indicator response time task and tertiary verbal callout task	Flight technical error, response time and percentage of correct callouts
	EVA Tether Dexterity Task	Hook and un-hook an EVA tether to handles and loops	Task completion time

The tasks performed on Day 1 were the U-bolt pegboard dexterity task, turntable tactility task, and switchboard tactility task. On Day 2, the lunar landing simulator and EVA tether dexterity tasks were performed. Detailed descriptions of each task are found in the following section.

First, the subjects conducted several ungloved training trials to familiarize them with each task and then conducted one ungloved trial to obtain their ungloved performance baseline. The subjects then conducted the tasks in the three glove fit configurations in both a pressurized and unpressurized state, resulting in six configurations per subject (Table 2.2). The order of the tasks within glove configuration was fixed and the order with which the subjects wore each glove was randomized between subjects. Subjects were blinded to the glove size that they were wearing.

Table 2.2: Experimental Configurations

		Glove Fit Configuration		
		Small	Prescribed	Large
Pressurization State	4.3 psid	Configuration 1	Configuration 2	Configuration 3
	Unpressurized	Configuration 4	Configuration 5	Configuration 6

2.1.3 Static Fit Metrics

Two different categories of fit metrics were gathered for this study: direct measures and perceived measures of static fit. The direct measures of static fit metrics were calculated by subtracting the anthropometric measurement of the subjects' hands from the corresponding outer dimensions of the pressurized gloves, providing a measure of ease. The anthropometric measurements gathered were hand length, hand circumference as well as the length and circumference of both the index finger and thumb. Thus, 6 direct measures of static fit were calculated. For the present analysis the direct measure with respect to hand length is presented. The direct measure with respect to hand length was selected as glove length was the dimension in which the gloves under examination varied, making it the most relevant for evaluation.

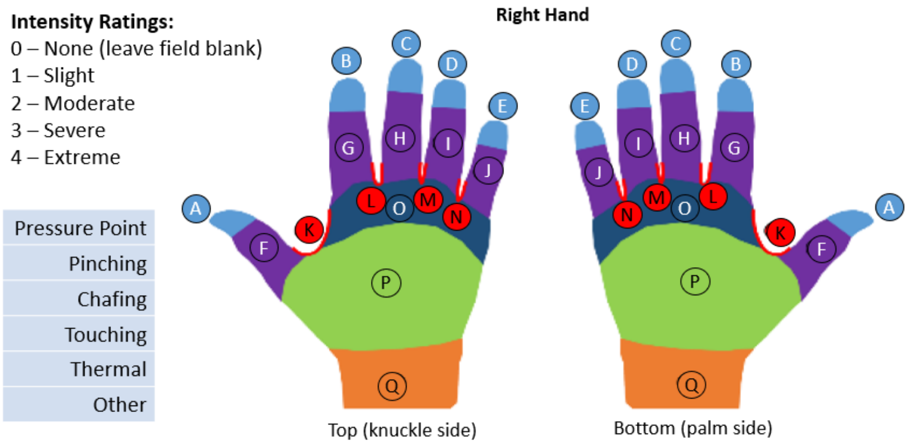


Figure 2-1: Comfort survey with areas of evaluation, range of intensity ratings, and possible discomfort categories. Figure adapted from Korona et al. [1]

Two different surveys were used to assess subject's perception of their static fit. The first was a comfort survey from the literature [1] wherein the subject would first identify any areas of discomfort based on the diagram, then categorize the sensation, and finally assign the discomfort a level of intensity (Figure 2-1). The survey was modified from the literature from a range of 0-10 to 0-4 to simplify the administration of the survey. If the subject did not feel discomfort in a given area, that rating was left blank.

The second survey was a custom fit survey informed by interactions with NASA Johnson Space Center. The subject was asked to rate their level of indexing at various points in the glove and provide a fit rating (Figure 2-2). The complete results of this survey across gloved conditions are presented in Section 3.1.4. Additionally, the ratings from the fit perception survey for fields A and B taken on Day 1 (when the tactility tasks were performed) were examined for possible correlation with the direct static fit metric to investigate whether subjects' perceptions of their fit were consistent with the direct measurements. The index fingertip and thumb tip were chosen as they were most relevant to the tactility tasks.

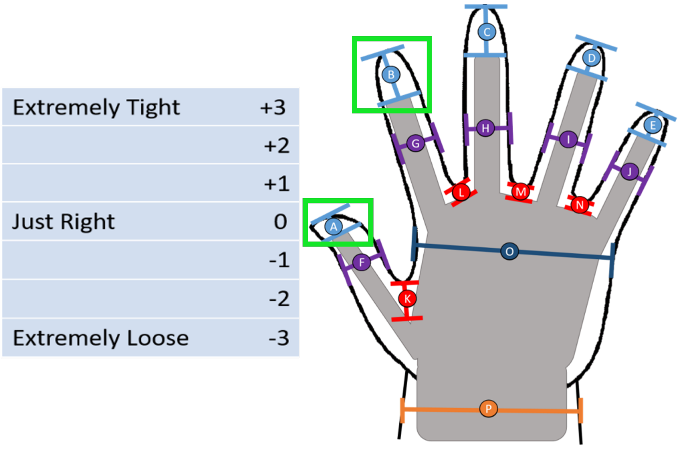


Figure 2-2: Subjective fit survey with areas of evaluation and range of possible fit ratings. Areas highlighted in green were specifically examined for possible correlations with the direct static fit metric.

2.1.4 Glovebox Vacuum Chamber

All tasks were conducted in a glovebox vacuum chamber (24 in diameter, 48 in length) (Figure 2-3). The chamber created a pressure differential to mimic the pressurization of the full suit. The vacuum chamber was kept between 4.0 psid and 4.5 psid for all trials to roughly match the nominal 4.3 psi operating pressure of the OCSS suit during a long duration contingency scenario [39].

Figure 2-3: Subject performing lunar landing simulator task in glovebox vacuum chamber

The workspace surface height was selected to limit unnatural postures when using the glovebox. Through in-person testing and information gathered on other glovebox systems, the workspace was designed such that it was 4 in below the cylinder midline. Another dimension of anthropometric concern was the distance between the two cylinder armholes. The center-to-center distance of the armholes was 18 in. Subjective feedback on any discomfort or difficulties imposed by the glovebox was gathered from all subjects and no complaints related to armhole placement or discomfort in the shoulders or upper arm due to the glovebox were recorded. To ensure postural consistency, the seated height of the subject was adjusted such that their arms in the forward reach position were parallel to the glovebox workspace.

2.2 Performance Assessment Tasks

A battery of tasks assessing the key areas of dexterous, tactile and cognitive performance were employed in this study. Both general and operationally relevant tasks were utilized. The hardware, procedures and metrics used in each task are described in the following section.

2.2.1 Tactility Tasks

Tactile performance can be essential to mission scenarios where the accurate perception of and use of controls via sense of touch may be required. This study utilized both a general turntable bump-detection task and an operationally relevant switchboard task to assess the key area of tactile performance.

General Turntable Tactility Task

The generalized turntable tactility assessment task was a modified version of the tactility task from a previous NASA EVA gloves study [56]. In the original NASA task, subjects attempted to perceive resin bumps of various size while blindfolded as force plate measurements under the bumps were collected. The modified version of the task employed in this research effort consisted of subjects attempting to detect bumps of different widths (0.59 in, 0.39 in, 0.20 in) and heights (0.05 in, 0.20 in,

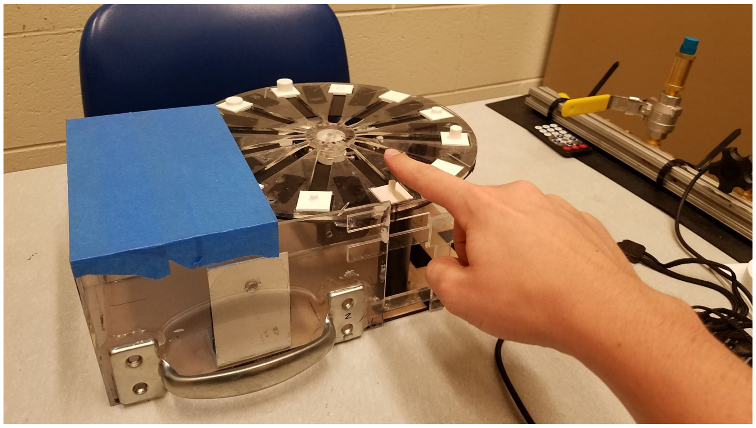


Figure 2-4: General turntable tactility task with hand in typical position for task

0.39 in), a narrower range than the original NASA study to decrease study time. An additional modification was that the bumps were placed on a motorized turntable such that the bumps could be interchanged while the glovebox chamber was under vacuum (Figure 2-4).

The turntable had 12 pads where the subject was asked whether or not they perceived a bump (9 bumps utilizing a combination of the aforementioned dimensions and 3 blank pads to assess the occurrence of false positives, Table 2.3 and Figure 2-5). The subjects were blindfolded during the task so that their ability to perceive the bumps with their sense of touch could be evaluated. The bump sizes were randomized on the turntable. The metrics recorded were the correctness of the subject's answer for each pad as well as the dimensions of any bumps where the subjects answered incorrectly.

Table 2.3: Dimensions of Bumps on Turntable Pads - Pads 10, 11, and 12 are blank pads

		Widths		
		0.20 in	0.39 in	0.59 in
Heights	0.39 in	7	8	9
	0.20 in	4	5	6
	0.05 in	1	2	3

Figure 2-5: General turntable tactility task numbered bumps

Switchboard Tactility Task

The operationally relevant switchboard tactility task involved following a sequence of control inputs on a mock spacecraft control panel (Figure 2-6). The panel was created in consultation with subject matter experts and laid out according to NASA standards [66]. The panel contained two rotary encoders of different size (1.1 in and 0.50 in in diameter, 0.56 in and 0.63 in height), five toggle switches (0.24 in diameter, 0.69 in height) with varied use of switch guards, two larger rotary switches (1.63 in diameter, 0.95 in height), and nine push-buttons (0.77 in width, 0.77 in length, 0.05 in height) laid out in number pad configuration.

Blindfolded subjects actuated a pre-defined sequence of these controls that was first presented to the subject via a training presentation and then practiced with the aid of the researcher during the practice trials. After these practice trials, the sequence was memorized by the subject. The sequence was always the same and consisted of the following:

1. Turn encoder 1 90 degrees clockwise
2. Flip toggle switch 1 down and then back up

3. Flip toggle switch 4 down and then back up
4. Turn dial 1 180 degrees clockwise and then 180 degrees counter clockwise
5. Press push button 9
6. Press push button 2
7. Press push button 4

Figure 2-6: Switchboard tactility task with control types annotated

The accuracy and completion time of the sequence was recorded. The accuracy score for each trial of the switchboard tactility task utilized a rubric (Appendix B-7) where one point was awarded or subtracted for each action in the sequence, subjects were penalized for the accidental activation of another control while utilizing the intended control, and points were awarded or subtracted based on the order in which the controls were performed. The maximum achievable score was 14 points. The switchboard task was completed 3 times for each configuration, with the accuracy and completion time metric the average of these 3 trials.

2.2.2 Dexterity Tasks

Dexterity is another area of performance that is key to accomplishing tasks on space exploration missions, but that is often inhibited by spacesuit gloves. To assess dexterity, a general u-bolt pegboard task and tool-based EVA tether task were utilized.

U-bolt Pegboard Task

The u-bolt pegboard task was a modified version of a u-bolt pegboard dexterity task from the NASA HPEG study [1]. In this task, subjects attempted to remove u-bolts from a pegboard, reorient the u-bolts and reinsert them into the pegboard. Modifications from the original NASA implementation included utilizing only one row of u-bolts as opposed to two to decrease study time. An additional modification recommended by NASA subject matter experts was the use of a metronome set at 50 beats per minute to ensure uniform timing and mitigate learning effects often associated with tasks utilizing completion time metrics. The task equipment consisted of a pegboard with one row of 5 u-bolts (Figure 2-7). Subjects were instructed to remove each u-bolt from the pegboard, turn it 90 degrees, then place it back into the pegboard for all u-bolts in the first row. If subjects dropped a bolt, they were instructed to leave it and move on to the next bolt. In addition to several ungloved practice trials given at the beginning of the task, subjects were also given a practice trial the first time they experienced the pressurized condition. After the practice trial, the u-bolt pegboard task was conducted 3 times for each gloved condition.

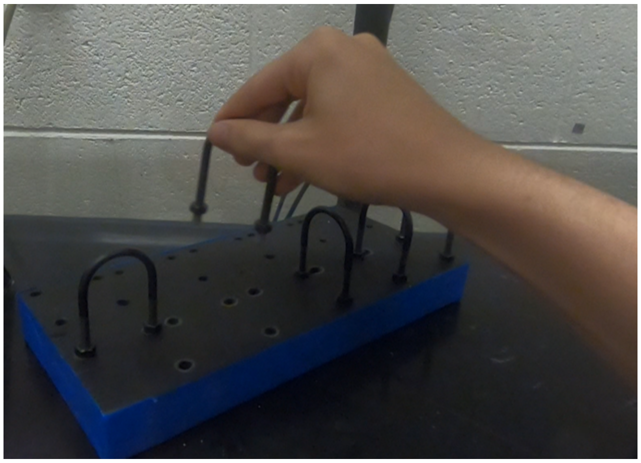


Figure 2-7: U-bolt Pegboard Task

For the u-bolt pegboard task, the first metric gathered was the average completion time per bolt for each trial. This value was obtained by dividing the overall completion time for the trial by the number of undropped bolts for that trial. Additional metrics were the number and percentage of dropped bolts each trial.

EVA Tether Task

The tool-based EVA tether task involved hooking and unhooking an EVA tether hook from a series of loops and a handle. The panel of loops and handles (Figure 2-8) was constructed and laid out according to NASA standards [66]. The EVA tether hook utilized was a non-flight test article obtained on-loan from NASA JSC. Subjects hooked and unhooked the EVA tether hook from each of the loops and handle in a predetermined sequence. The sequence entailed first hooking and unhooking from the rightmost and smallest loop, then the center larger loop, then the handle. To obtain timestamps and ensure that subjects fully removed their hand from the hook in between in each motion, subjects were instructed to hit a push button in between each loop or hook. Subjects were instructed to complete the sequence of actions as quickly and as accurately as possible. In addition to several ungloved practice trials given at the beginning of the task, subjects were also given a practice trial the first time they experienced the pressurized condition. After the practice trial, the EVA tether task was completed 3 times for each gloved condition. The metric for the EVA tether task was the average task completion time for each condition.

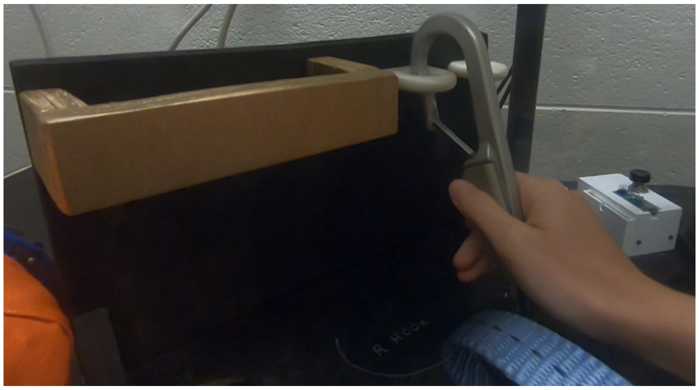


Figure 2-8: EVA Tether Task

2.2.3 Lunar Landing Simulator

The Draper real-time performance metrics workstation lunar landing simulator [35] consists of flying a simulated lunar lander during the terminal descent phase. The physical simulator setup consisted of a translational hand controller to control spacecraft ascent/descent rate, a rotational hand controller to control spacecraft attitude, a headset to record verbal callouts of altitude and fuel, and a monitor to display the simulation (Figure 2-3). The joysticks were placed 16 in. apart and positioned such that the subjects arms were in the forward reach position during flight. The monitor was placed such that the bottom of the screen rested on the top of the glovebox and was tilted downwards to allow for easier viewing by the subjects (Figure 2-9).

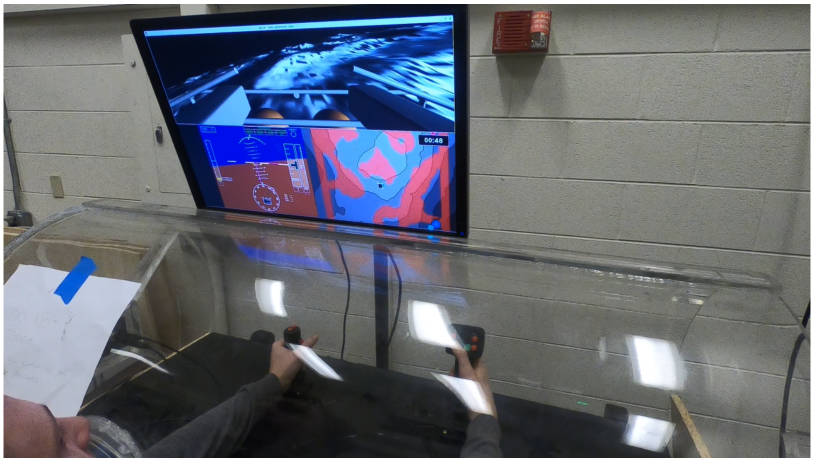


Figure 2-9: Lunar landing simulator monitor placement on glovebox setup

The simulation scenario entailed the spacecraft beginning in fully automatic mode, where the computer performed all necessary control of the spacecraft, then at a point during the simulation (approximately 20 seconds into the simulation) a transition occurred to 3-axis manual control mode. At the time of transition, the landing point for the spacecraft was re-designated, which introduced an attitude error which the pilot was required to null. In Figure 2-10 the yellow and pink crosshairs in the center of the screen indicate the current and target attitude, respectively. The rightmost meter indicates the current and target ascent/descent rate.

The lunar landing simulator consisted of 3 subtasks: the primary flight task (to assess flight technical error), a response time task (to assess mental workload) and verbal callouts of simulation state (to assess situational awareness). The subject's primary goal was to complete the flight task, followed by a secondary goal of the

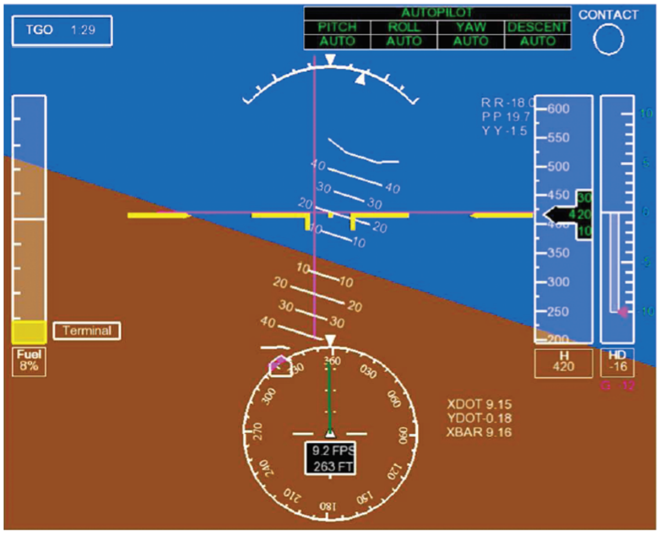


Figure 2-10: Primary Flight Display

response time task, and a tertiary goal of the verbal callouts. Subjects were explicitly told not to sacrifice performance on the primary flight task for either of the additional subtasks. Subjects completed all 3 subtasks, but only the primary flight task and the response time task are discussed in this thesis. The goal of the response time task was to minimize response time to a communication indicator. Starting from the beginning simulation and finishing before the final spacecraft descent, a communication indicator in the corner of the horizontal situation display (Figure 2-11) would change from an inactive state to having either a blue or green outer circle at random intervals between 5-7 seconds.

To respond to the communication indicator, subjects would press the corresponding blue or green button on the top of the rotational hand controller. 10 total communication indicators were displayed each trial, however the time and color of each communication indicator were randomized. Subjects were instructed to respond to the communication indicator as quickly as possible. Subjects were instructed not to compromise their performance on the primary flight task in order to respond to the communication indicator. In Figure 2-11 the red areas on the display indicate hazardous areas on the terrain map, however this aspect of the simulation was not incorporated into these tasks. The black square indicates the position of the space-

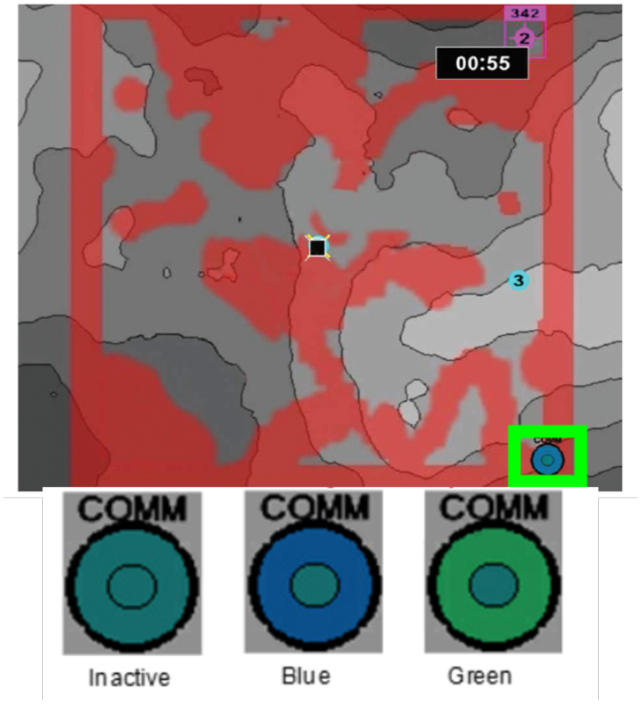


Figure 2-11: Horizontal Flight Display and Communication Indicator

craft. The communication indicator was located in the bottom right corner of the horizontal flight display.

Training for subjects included a powerpoint presentation describing the purpose of the simulation, detailing the individual aspects of the simulation display, and instructing the subject how to utilize the controls. The subjects then performed several ungloved practice runs under coaching from study personnel. The practice runs entailed learning how to fly the simulator and then incorporating each of the additional subtasks. Depending on the speed of learning as judged by the study personnel, the number of initial training trials ranged from 8-12. Additionally, after the initial pressurization of the first glove in the sequence, subjects were given 2-3 additional practice trials while pressurized to allow them to determine their preferred motor technique for the pressurized case. Subjects then flew 2 experimental trials for each gloved condition.

2.3 Statistical Analyses

Non-parametric Friedman tests were used to assess the average completion time per bolt for the u-bolt pegboard task, the average completion time for the EVA tether task, the differences in the percentage of bolts dropped for the u-bolt task, the RMSE in the pitch and roll axes for primary flight subtask, and the response time for the response time subtask, as the data were not normally distributed. Post-hoc Wilcoxon Sign Rank tests were conducted to examine the differences between specific gloved conditions. For the primary flight subtask in the lunar landing simulator, separate friedman tests and post-hoc comparisons of gloved conditions were conducted at each the probes where data was gathered (40 s, 50 s, 60 s, and 70s). Effect sizes for the aforementioned tasks were estimated using the r-value (r) where a small effect is a value between 0.1 – 0.3, a medium effect is between 0.3-0.5 and a large effect is greater than 0.5 [67].

The number of correct answers for the turntable task was also assessed using a non-parametric Friedman test, as the data were not normally distributed. Post-hoc Wilcoxon Sign Rank tests were conducted to examine the differences between specific gloved conditions.

The switchboard score was assessed using a repeated measure ANOVA with subject as a random factor and suited condition as a fixed factor with 7 levels (Baseline, size S unpressurized, size S pressurized, size P unpressurized, size P pressurized, size L unpressurized, size L pressurized). A Shapiro-Wilk test supported normality of the model residuals. Post-hoc dependent T-tests were conducted to examine the differences between specific suited conditions. Effect sizes for the switchboard task were estimated using Cohen's d, where a small effect is a value between 0.2 – 0.5, a medium effect is between 0.5 – 0.8, and a large effect is greater than 0.8.

The post-hoc pairwise comparisons for the turntable task, switchboard task, u-bolt pegboard task, EVA tether task, and the Lunar Landing Simulator primary flight subtask in the roll axis were corrected using the false discovery rate method [68]. For Lunar Landing Simulator primary flight subtask in the pitch axis and the response time subtask, the uncorrected post-hoc comparison results are presented with associated effect sizes.

Differences between section-wise scores on the switchboard were assessed using a Friedman test as the normalized section-wise scores were not normally distributed. Post-hoc Wilcoxon Sign Rank tests were conducted to examine the differences between specific section scores on the switchboard task.

A linear mixed effect model (LME) was fit to examine whether static fit was related to switchboard score. The model incorporated pressurization state (categorical variable of pressurized or unpressurized), the normalized direct static fit metric for hand length, where the absolute value of ease metric is divided by hand length (continuous variable), and subject (random variable). A Shapiro-Wilk test supported normality of the model residuals.

The Spearman correlation coefficient was estimated between the direct static fit measure for hand length and the fit perception measure for index fingertip and thumb tip to investigate whether subjects' perceptions of their fit were consistent with the direct measurements.

In the boxplots presented in the following results sections, the green line represents the median, the blue bar represents the interquartile range from the 25th percentile to the 75th percentile, and the blue dots represent outliers. On the figures which compare gloved conditions, the following notation is used: S-UP (small unpressurized), S-P (small pressurized), P-UP (prescribed unpressurized), P-P (prescribed pressurized), L-UP (large unpressurized), and L-P (large pressurized).

Chapter 3

Tactility Task and Subjective Survey Results and Discussion

This chapter presents the results for both the general turntable tactility task and the switchboard tactility task. These results are then discussed in the context of the literature to draw conclusions regarding the relationship between spacesuit glove fit and tactile performance. Additionally the results of the fit perception and comfort surveys are also discussed.

3.1 Results

The results of the statistical tests run to analyze the general turntable task and switchboard task are presented in the following section. The results of the fit perception and comfort are also presented.

3.1.1 General Turntable Tactility Task Results

For the turntable task, the first metric gathered was the number of correct answers for each configuration (Baseline, size S unpressurized, size S pressurized, size P unpressurized, size P pressurized, size L unpressurized, size L pressurized) (Figure 3-1).

A Friedman test supported an effect of suited condition on turntable score ($\chi^2(6)= 14.67$, $p = 0.023$). However, after corrections with the false discovery rate method,

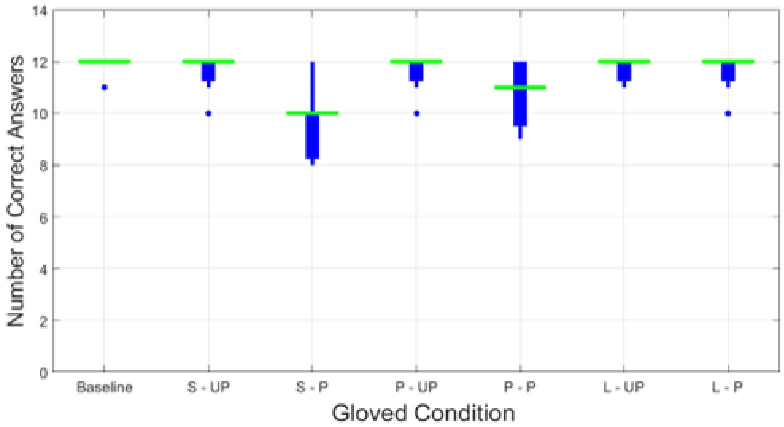


Figure 3-1: Number of correct answers per gloved condition. The maximum number of correct bumps was 12.

post-hoc comparisons did not support any pairwise significant differences.

Although for this smaller sample size, there were no significant differences in median, it can be qualitatively observed that there is a wider interquartile range of scores for the pressurized size S and size P case.

The cases which yielded an incorrect answer from the subjects (Figure 3-2) most commonly were bumps of 0.05 in height. It was also observed that the pads with no bump also presented a challenge to some subjects.

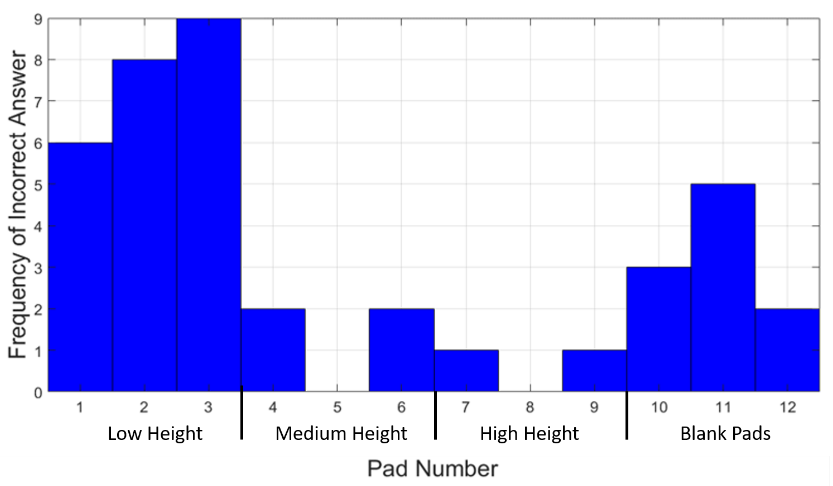


Figure 3-2: Frequency of incorrect answer per pad across all conditions using the numbering from Table 2.3. Heights of pads denoted on x-axis. See Figure 2-5 and Table 2.3 for information on bumps.

3.1.2 Switchboard Task Results

The ANOVA fit for switchboard score supported significant main effects of subject (F(6,36) = 5.61, $p < 0.001$) and gloved condition (F(6,36) = 13.05, $p < 0.001$). Post-hoc pairwise comparisons support that all conditions where a glove was worn yielded significantly lower scores than the baseline trial (d = 1.12 to 3.67) except for the size L, unpressurized case (Figure 3-3). Additionally, it was found that for each glove size, the score from the pressurized condition was significantly lower than the score from the unpressurized condition (d = 1.12 to 1.81). When examining the effects of sizing in the unpressurized gloves, it was found that the scores from the size L trials were significantly higher than those on the size S trials and size P trials (d = 1.31 to 1.96).

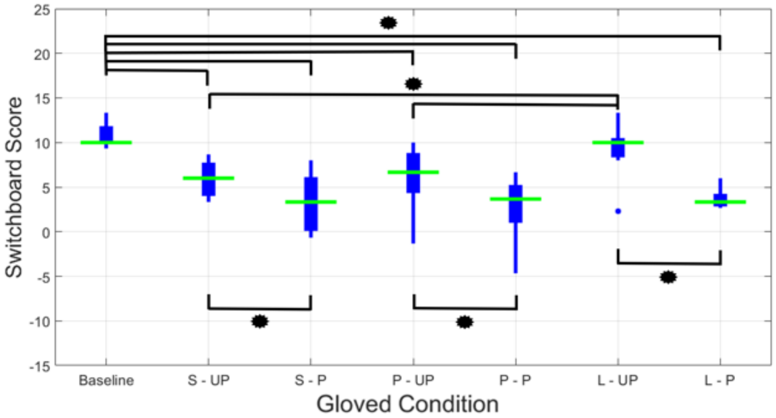


Figure 3-3: Switchboard scores for the suited conditions. Significant ($p<0.05$) FDR corrected post-hoc pairwise comparisons are notated with an asterisk

The scores were further parsed for each control component section of the switchboard task (Figure 3-4). A Friedman test supported an effect of control section on percentage of maximum section score ($\chi^2(3) = 61.25$, $p < .001$). Post-hoc comparisons confirmed that the encoder, toggle switch, and rotary switch all had significantly higher section-wise scores compared to the push button.

3.1.3 Direct Static Fit Metric Results

The normalized hand length ease metric ranged from 0.19 in size S for the subject with the largest hand length, to 0.32 in in size L for the subject with the smallest

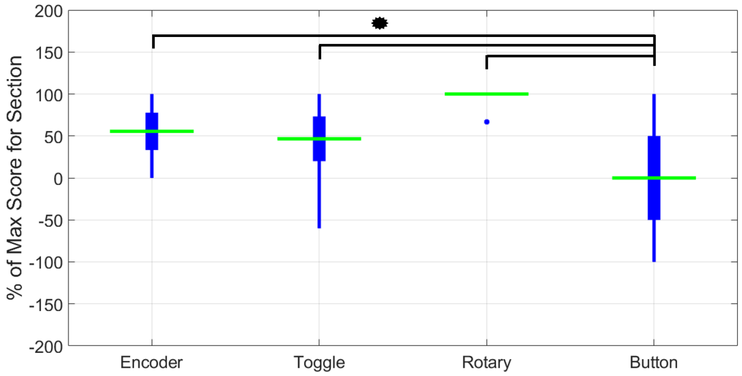


Figure 3-4: Distributions of percentage of maximum available points per control section on the switchboard task. Significant ($p<0.05$) FDR corrected post-hoc pairwise comparisons are notated with an asterisk.

hand length. The LME fit for switchboard score with predictors of normalized values of hand length ease and pressurization supported that both the normalized ease (95% CI [11.89, 71.20]) and pressurization state (95% CI [-5.11, -2.46]) had a significant relationship with switchboard score.

3.1.4 Subjective Survey Results

Using Figure 2-2 subjects were asked to rate their percieved fit in the glove for each condition. Figure 3-5 shows the results of these ratings averaged across all subjects for their right hand on Day 1 (when the tactility tasks were performed) of the study. For the purposes of Figure 3-5, the average perceived fit values are rounded to the nearest integer for simplicity. The numerical averages corresponding to to Figure 3-5 can be found in Appendix A.1. When unpressurized, several of the fingertip regions are tight for the small and prescribed gloves. This contrasts with the pressurized conditions, which see more loose ratings in the fingers, particularly for the large glove.

When calculating the Spearman correlation coefficient between the direct static fit measure for hand length and the fit perception ratings for the thumb and index finger tips, it was found that there was no significant correlation between these values for either the thumb (ρ = -0.009, p =0.96 or the index finger (ρ = -0.254, p = 0.10).

The raw results of the comfort survey from Figure 2-1 can be found in Appendix

A.2 for the unpressurized conditions and Appendix A.3 for the pressurized conditions.

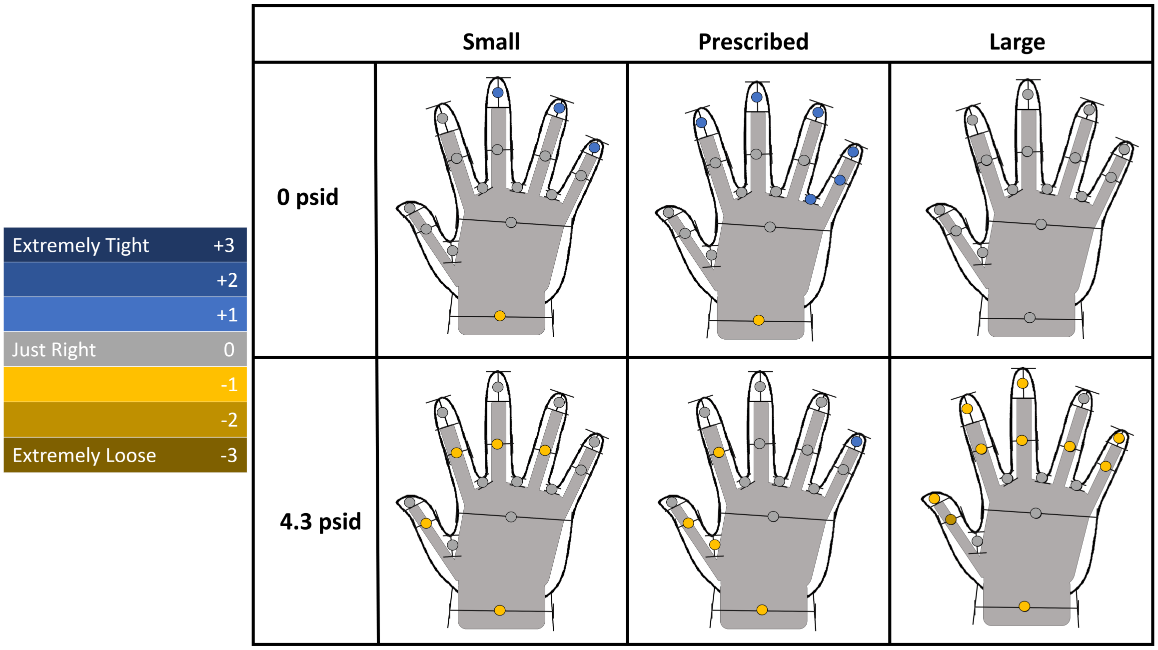


Figure 3-5: Results of the fit perception survey for day 1 in the right hand. Average ratings across subjects are rounded to the nearest whole number.

3.2 Discussion

In the following section the conclusions drawn from the analysis of the tactility tasks and subjective surveys are discussed.

3.2.1 General Turntable Tactility Task Discussion

The results of the turntable task show that for this dataset, gloved condition did have an effect on tactile performance although post-hoc tests did not find any significant pairwise differences. Figure 3-1 qualitatively shows a wider interquartile range for both the pressurized size S and pressurized size P case, suggesting that pressurization with limited easement did inhibit tactile performance for some subjects in this task.

3.2.2 Switchboard Tactility Task

The switchboard task was more operationally relevant to crewmember performance by design, and the results of the task provide several key insights into the effect of spacesuit gloves on performance. Figure 3-3 shows that 5 of the 6 gloved trials yielded significantly lower scores on the switchboard task over the ungloved baseline trial. This result agrees with previous literature that tactile performance is reduced by EVA spacesuit glove use over barehanded conditions, where Thompson et al. [56] found that the force required to detect resin bumps in a tactility task was higher for the gloved cases and O'Hara and et al. [54] found a substantial performance decrement on a two point discrimination task when comparing gloved cases to ungloved cases. Additionally, this finding is supported by subjective reports, where Shuttle era crews had issues with switches and other hand operated controls, which resulted in some crew members not wearing their gloves to improve their performance on necessary tasks [69]. These results support the importance of considering tactile performance when designing spacecraft cockpits and suggest the use of higher profile controls and increased control spacing.

Figure 3-3 also shows that within each glove size, switchboard score was significantly lower for pressurized trials than unpressurized trials. This result also agrees with findings by both O'Hara et al. [54] and Thompson et al. [56] that show that glove pressurization leads to decreased tactile performance. As with the previous findings, this result also indicates the importance of considering the effect of pressurization during spacecraft cockpit design. It is important to note that the aforementioned literature utilized EVA gloves, which are always pressurized during nominal operations outside the vehicle, whereas this study utilized IVA gloves which are used inside the vehicle, unpressurized during nominal operations and pressurized in contingency scenarios. IVA gloves therefore present a design challenge, as it is important that they allow crewmembers to achieve necessary tactile performance in both cases [39].

With respect to the effect of glove size on tactile performance, Figure 3-3 shows that the unpressurized case for glove size L had significantly higher scores than both those of glove size P and glove size S. While we are unaware of any previous literature that assesses tactile performance with respect to glove size, this result appears to be counter-intuitive upon initial examination. Tremblay-Lutter and Weihrer [12] found that increased ease, as one would expect in the transition from size S or size P to size L, led to a decrease in performance, although the direct comparison of dexterity tasks in that study to these tactility task results is unclear. One possible explanation is that the added ease in the index finger tip allowed for greater perception of the controls with subjects' sense of touch when the inner surface of the glove contacted the finger pad. While at the moment of contact between the glove and the control, the material stack-up would be the same across glove sizes, the easement in the larger glove size would have allowed for more room inside the glove for the finger to move around and contact and re-contact the internal glove surface. A similar phenomenon was theorized by Thompson et al. [56] to explain the finding that, for the case of resin bump height, glove pressurization slightly aided bump detection. Related to this explanation is the possibility is that with the smaller gloves, the fast adapting mechanoreceptors on the fingertips that provide key sensory feedback during object manipulation [70] were no longer transmitting, as they are insensitive to static forces. With the additional ease in size L, the finger had some internal motion and these mechanoreceptors were signaling the dynamic interactions. The finding that unpressurized performance increased with the larger size aligns with anecdotal evidence (conveyed via DCCI) from the Shuttle program, wherein crewmembers would routinely opt for one size bigger than their prescribed fit, in an effort to improve performance during nominal (unpressurized) operations. It is also important to note that increasing the ease beyond the threshold measured here could lead to decreased tactile performance.

3.2.3 Tactility Task Comparison

While some turntable trials did have lower scores than others, most scores were high, with the maximum number of incorrect answers on any trial being 4 out of 12. This result stands in contrast to the switchboard task (Figure 3-3), which had a greater range of scores and supported significant differences between suited conditions. These differences may result from the turntable task being simpler. The switchboard task required subjects to find and actuate controls in the correct manner, whereas the goal of the turntable task was only to provide a binary yes or no answer as to whether a bump was detected in a single location of interest.

Despite the differences between the protocols and outcomes of the two tasks, it is interesting to compare their results with respect to where subjects struggled on each task. For the turntable task (Figure 3-2), subjects missed the turntable bumps of height 0.05 in (most frequently for the bump of 0.05 in height and the largest diameter, 0.59 in) and had difficulty providing the correct answer to the blank pads. These results suggest that there is a height threshold for consistent tactile detection. Comparing these results to the switchboard (Figure 3-4), subjects had the worst performance on the push button portion of the task, which has the lowest height (0.047 in) of any of the controls. These results are in agreement with verbal feedback from subjects during the switchboard task, many of whom stated that they could not perceive the push buttons while gloved and instead were relying on muscle memory from the training trials to attempt to complete the task. These push button results are in contrast to the encoders, rotary switches and toggle switches, which all have larger height values and correspondingly higher median scores.

For keyboard buttons, which are most similar to the push buttons used in this experiment, the NASA Human Integration Handbook [71] lists a minimum height of 0.04 in and a preferred height of 0.08 in. Additionally, legend buttons, which are also similar to the switchboard push buttons, do not specify a minimum height (although the associated barrier depth is 0.2 in). The results of this experiment suggest that for push-button like controls that may require use with sense of touch (such as an in an emergency situation where visual attention may be focused on a display or other ongoing events, or be obscured due to smoke), it may be important to raise the minimum height above 0.05 in.

3.2.4 Fit Metric Discussion

Direct Static Fit Metric

The results of the LME incorporating the normalized direct static fit metric showed that static fit had a significant relationship with switchboard score, with increases in ease associated with increased score. This result supports the initial hypothesis that direct static fit measures are related to performance on an operational tactility task. The regression provides a finer interpretation of fit beyond glove size as the static fit measure incorporates the individual hand size, which varies, and glove size together. However, as the range of the ease was small, it is important not to extrapolate beyond the data that was used to create the model.

Subjective Surveys

When examining Figure 3-5 several qualitative observations can be made regarding subjects' percieved fit. It can be seen that across pressurized conditions and sizes no hand locations had rounded average perceived fit ratings outside of the -1 to +1 range (with the exception of the thumb knuckle in the large pressurized condition). This result could suggest that perceived fit was mostly satisfactory for subjects across a range of conditions. Alternatively, this could potentially indicate that providing a wider range of possible ratings (perhaps from -10 to 10 as opposed to -3 to 3) might prompt subjects to give a more refined rating of their perceived fit. Whereas directions given to the subject regarding the rating scale were intended to have subjects use a value such as -1 to identify an area that was slightly too loose, examination of the ratings distribution indicate that subjects may have instead utilized it simply indicate that the glove was loose, without regard to the degree of looseness. Providing subjects additional information on an appropriate reference guideline may yield improved perceptual responses.

With respect to specific areas of the hands, it can be seen that certain areas such as the hand width and index and middle finger crotches averaged ratings within the rounded "just right" range across all gloved conditions. A rating of looseness in the wrist area can also been seen across all gloved conditions except the large unpressurized case. When examining the effect of pressurization, all of the ratings that qualify as tight occur in the small and prescribed unpressurized case (which the exception of the tip of the pinky in the prescribed pressurized case). This contrasts with the loose ratings that can be seen in the fingers during the pressurized cases, with the most loose ratings occurring in the large unpressurized case. This makes intuitive sense, as when the gloves inflate with pressurization, they expand away from the hand, providing more ease to the user.

Another important concept to consider when evaluating the fit metrics is that the proportions of various anthrompometric measures vary across subjects. The inclusion criteria of falling into the range for the prescribed glove size was only based on two measurements (hand length and circumference). While subjects may have had similar hand length measurements in the proper range for the perscribed glove size used in this study, other measurements (such as digit length or girth) may have varied significantly between subjects and influenced their perceived fit.

The data did not support a significant correlation between perceived fit of the thumb and index finger and direct static fit measures with respect to hand length. There are several considerations when interpreting these results. For both the index finger and thumb only two subjects provided ratings outside of -1, 0 or 1, lining up with the aforementioned discussion that increasing the range of the survey scale could lead to increased fidelity. Another possibility is that while it was assumed that there would be a relationship between the direct measure of static fit with respect to hand length and these fit perception values for the index finger and thumb, that perhaps it would be more appropriate to calculate the direct static fit measure with respect to index finger and thumb length directly. It is also possible that the level of ease in the gloves used in this study was not easily perceptible.

The comfort survey was not quantitatively analyzed for this thesis, however a qualitative examination of the data yields several preliminary insights. The first is that for the majority of fields in Figure 2-1, no discomfort ratings were given, indicating that most areas of the glove were of satisfactory comfort. Additionally, the “O” field for the top of the hand (which refers to the knuckle area in Figure 2-1) had the clear plurality of recorded discomfort ratings, indicating an area of consistent discomfort across subjects. This lines up with verbal feedback from subjects as well as photo documentation of skin irritation on the side of the index finger knuckle. The raw results of this survey can be found in Appendix A.2 and Appendix A.3

Chapter 4

Dexterity Results and Discussion

The following section presents the results of the general u-bolt pegboard and tool-based EVA tether dexterity assessment tasks. The analysis of the data and the conclusions regarding the relationship between spacesuit gloves and dexterous performance are then discussed.

4.1 Dexterity Task Results

The results of the statistical tests for the general u-bolt pegboard task and tool-based EVA tether task are discussed in the following subsections.

4.1.1 U-bolt Pegboard Task Results

A Friedman test supported an effect of gloved condition on average completion time per bolt ($\chi^2(6) = 37.82$, $p < 0.001$). Post-hoc pairwise comparisons support that all pressurized cases had significantly higher average completion times per bolt than the baseline trial ($r = 0.42$ to $r = 0.63$) (Figure 4-1).

Additionally, it was found that for each glove size, the pressurized case had a significantly higher average completion time per bolt compared to the corresponding unpressurized case ($r = 0.63$). No significant differences between glove sizes were detected for either the unpressurized or pressurized conditions.

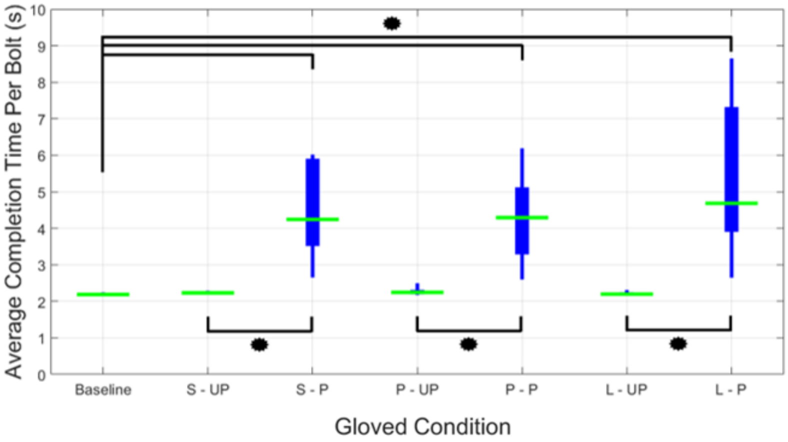


Figure 4-1: Average completion times per bolt for each gloved condition. Significant ($p<0.05$) FDR corrected post-hoc pairwise comparisons are notated with an asterisk.

The second metric collected for the u-bolt pegboard task was the number of dropped bolts. Bolts were only dropped during the pressurized trials (Figure 4-2). A Freidman test found no significant effect of glove size when comparing the percentage of dropped bolts between the pressurized trials.

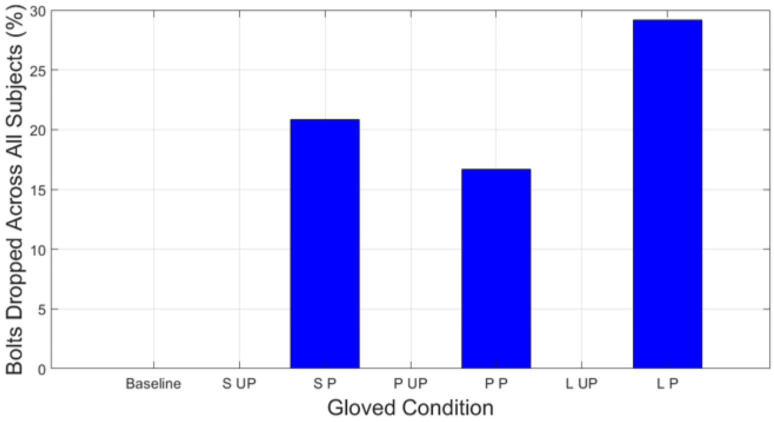


Figure 4-2: Percentage of dropped bolts for each gloved condition across all subjects. Number of bolts dropped for each individual subject ranged from 0-2 bolts (0-40% of the bolts per trial) for each gloved condition.

4.1.2 EVA Tether Task Results

A Friedman test supported an effect of gloved condition on average task completion time ($\chi^2(6) = 33.24$, $p < 0.001$). Similiarly to the u-bolt task, post-hoc pairwise comparisons support that all pressurized cases had significantly higher average task completion time than the baseline trial (r = 0.63) and that the pressurized case had a significantly higher average completion time compared to the unpressurized case for each glove size (r = 0.63) (Figure 4-3). No significant differences between glove sizes were detected for either the unpressurized or pressurized conditions.

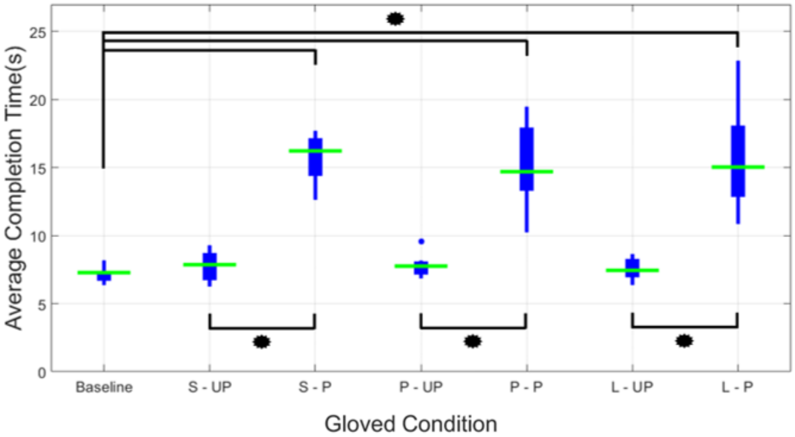


Figure 4-3: Average EVA tether task completion time for each gloved condition. Significant ($p<0.05$) FDR corrected post-hoc pairwise comparisons are notated with an asterisk.

4.2 Dexterity Task Discussion

The following subsections discuss the conclusions related to the effects of spacesuit gloves on dexterity in the context of the existing literature.

4.2.1 Effects of Pressurization on Dexterous Performance

The results of the u-bolt pegboard task and the EVA tether task demonstrate an effect of glove pressurization on dexterity. Figures 4-1 and 4-3 show that for both tasks, all of the pressurized trials yielded significantly higher completion times than the ungloved baseline trials. This is consistent with findings from O'Hara et al. [54] as well as Korona et al. [1] who found that performance on dexterity tasks was reduced by pressurized EVA gloves over barehanded conditions. Figures 4-1 and 4-3 also demonstrate that for each glove size, pressurization significantly reduced performance over the corresponding unpressurized trial. This result agrees with the findings of O'Hara et al. [54] who also found that glove pressurization leads to reduced performance on dexterity tasks over unpressurized gloved conditions. For the u-bolt pegboard task specifically, Figure 4-2 shows that only the pressurized trials led to dropped u-bolts further supporting the finding that pressurization led to decreased dexterity (however it should be noted that the number of dropped bolts ranged from 0-2 for each subject, with not all subjects dropping u-bolts). It is important to note that the aforementioned literature utilized EVA gloves which are always pressurized during nominal operations, whereas this study utilized IVA gloves which are used in an unpressurized state during nominal operations, but pressurize in contingency scenarios [39].

A qualitative examination of videos of subjects performing the dexterity tasks also provided insight into the effects of pressurization on task motor strategy. When examining the unpressurized trials for the u-bolt pegboard task, subjects varied in their motor strategy, with some subjects gripping the u-bolt using only their index and thumb, some adding the use of their middle finger, and others making use of their other digits and palm to stabilize the u-bolt as it was removed and reinserted. This stands in contrast to the pressurized cases, where all but one subject made use of only their thumb and index finger to grasp the u-bolt. This change may arise as the pressurization of the gloves makes the digits more bulky and strenuous to bend, making it more difficult to utilize other digits to grasp the ubolt. Examination of subject footage also showed that during the pressurized trials, subjects struggled to grasp and direct the u-bolt both during the removal and insertion of the bolt.

For the EVA tether task, subjects efficiently completed the task in the unpressurized conditions. When conducting the task in the pressurized case, subjects had difficulty picking up the tether hook off of the workspace and regaining control of the hook after pressing the timing button. However, once the subject had grasped the hook, the actuation and direction of the hook was not as difficult. This ease of actuation is to be expected as the tool was designed to be actuated in a pressurized spacesuit glove. In terms of variations in motor strategy, in the unpressurized case most subjects hit the timing button with their palm or outstretched fingers, whereas in the pressurized case, most subjects instead hit the button with the side of their hand. The preference for this motor strategy could arise due to the the rotation of the hand necessary to hit the button with the palm, which is more difficult in the pressurized condition due to friction in the wrist bearing. The friction in the glove bearing was not quantified, however verbal feedback indicated that some subjects noticed a difference in bearing friction between the unpressurized and pressurized conditions but that this did not prevent them from accomplishing the task.

4.2.2 Effects of Gloves on Dexterous Performance

While both dexterity tasks demonstrated differences between the ungloved condition and the pressurized condition, neither task demonstrated significant differences between the ungloved condition and the unpressurized gloved condition. These results conflict with O'Hara et al. [54] who found that unpressurized glove conditions led to significant decrements in dexterity performance over ungloved trials in EVA gloves, although it should be noted that EVA gloves are not designed to be utilized when unpressurized, which may play a role in the performance decrements observed in that study. Our results also conflict with Newton et al. [52] who found that use of an IVA glove from Final Frontier Designs led to a decrease in performance on a Lafayette Purdue Pegboard Task, Lafayette Hand Tool Dexterity Test, and rope-tying task over ungloved conditions. A similarity in task performance between the ungloved condition and unpressurized gloved condition, as found in our study, is a desired goal of an IVA glove and these results could support an improvement in IVA glove design for the selected gloves in context with the selected tasks. The results could also suggest that the dexterity tasks used in this study were not difficult enough to result in decreased performance in the unpressurized gloved condition. The u-bolt pegboard task was conducted at a metronome speed of 50 beats per minute in order to mitigate learning effects, however this pace may permit consistent completion across the ungloved and unpressurized conditions. In the O'Hara et al. [54] study a pacing was not provided, and subjects could select their own speed. For the EVA tether task, the tool itself was designed for gloved operation and only requires a simple squeezing motion to activate, rather than complex motions with fingers. An alternate explantion for these results could be that there are significant, but small effects sizes that would be observable with a larger sample size. Although, the effect sizes observered by O'Hara et al. [54] were found to be large. The similarity in performance between conditions could also arise if the subset of subjects that completed the task had a greater dexterity than previous studies.

4.2.3 Effect of Glove Size on Dexterous Performance

No significant difference among glove sizes was detected for either of the dexterity tasks. While no previous literature examines the effect of spacesuit glove sizing on dexterity tasks specifically, this result conflicts with Tremblay-Lutter et al. [12] who found that increased ease due to larger glove size led to reduced performance on dexterity tasks in chemical protective gloves. Although no there were no significant differences in medians detected, it can be qualitatively observed from figure 4-1 and 4-3 for both tasks, in the pressurized condition, the large glove size had the widest range of completion times indicating that for some subjects increased easement may have inhibited performance. It is unclear how the range of ease in Tremblay-Lutter et al [12] compares to the current study as underlying anthropometric measures are not provided to facilitate comparison with our normalized static fit measures, and the non-normalized measures could be not be compared because they related to different underlying body segments (digit length vs. hand length). As with the results for the comparison of ungloved to gloved unpressurized conditions, the results when comparing glove sizes could be due to the level of task difficult. For the u-bolt pegboard task, conducting the task at a faster metronome tempo or a self-selected maximum pace may allow for an evaluation for the limits of dexterity in a given gloved condition. The nuance of whether a task evaluates the limit of dexterity versus performance for a specific pacing is an important consideration for requirement development for spacesuits.

Chapter 5

Lunar Landing Simulator Results and Discussion

The following chapter presents the results of the lunar landing simulator dataset. Conclusions regarding the primary flight subtask and the secondary mental workload assessment task are discussed.

5.1 Lunar Landing Simulator Results

The following section presents the result of the statistical tests for the primary flight subtask assessing flight technical error in the roll and pitch axes, and the response time subtask that assesses mental workload via a reaction time task.

5.1.1 Primary Flight Subtask

A Friedman test supported an effect of gloved condition on RMSE in the pitch axis for the 40 second simulation time probe ($\chi^2(6) = 17.12$, $p = 0.0089$), but none of the subsequent probes (Figure 5-1).

For the 40 second probe, post-hoc pairwise comparisons support that the large unpressurized case had signficiantly higher RMSE than the baseline trial ($r = 0.37$) and the large pressurized condition had significantly lower RMSE than the baseline

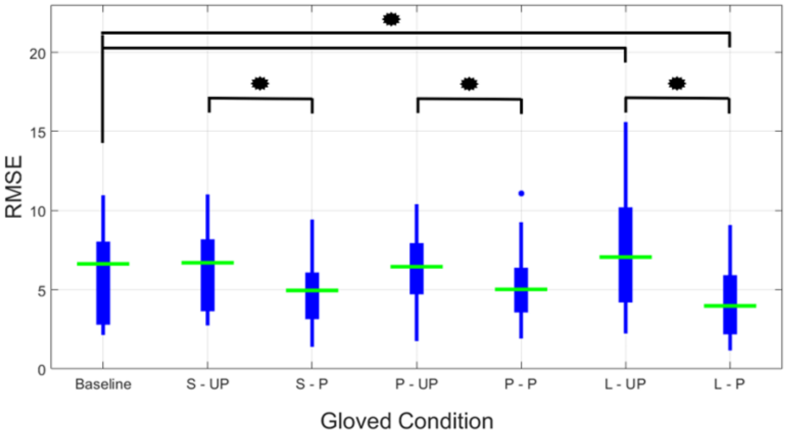


Figure 5-1: RMSE in pitch axis for each gloved condition at the 40 second probe. Significant uncorrected post-hoc pairwise comparisons are notated with an asterisk ($p < 0.05$), none are significant when corrected using FDR. The 40 second probe shows subject behavior when still actively nulling the error induced by the LPR.

trial in the pitch axis ($r = 0.41$). For each glove size, pressurized case had a significantly lower RMSE in the pitch axis compared to the unpressurized case ($r = 0.38$ to $r = 0.49$). No significant differences between glove sizes were detected for either the unpressurized or pressurized conditions.

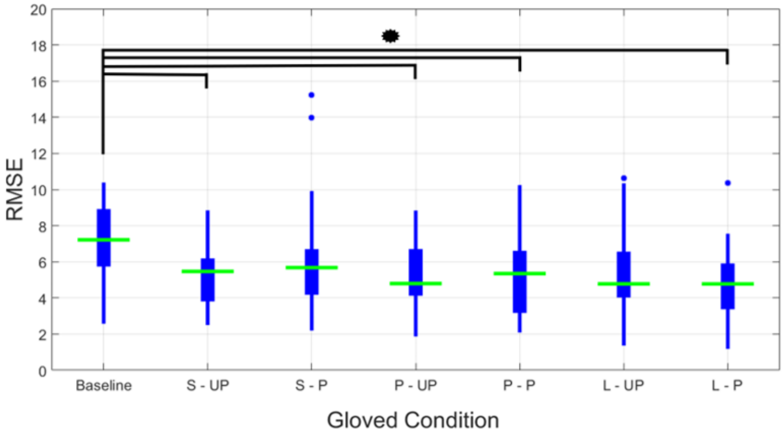


Figure 5-2: RMSE in roll axis for each gloved condition at the 40 second probe. Significant ($p<0.05$) FDR corrected post-hoc pairwise comparisons are notated with an asterisk.

In the roll axis, a Friedman test supported an effect of gloved condition on RMSE for the 40 second ($\chi^2(6) = 18.51$, $p = 0.0051$), 50 second ($\chi^2(6) = 14.46$, $p = 0.0249$), 60 second ($\chi^2(6) = 14.87$, $p = 0.0213$), and 70 second ($\chi^2(6) = 17.25$, $p =$

0.027) probes. Post-hoc pairwise comparisons support that the small unpressurized, prescribed unpressurized, prescribed pressurized, and large pressurized glove sizes had significantly lower RMSE than the baseline trial for the 40 second probe (Figure 5-2) (r = .48 to r =.59), the 50 second probe (r = .44 to r =.54) and the 60 second probe (r = .46 to r =.52). For the 70 second probe, the baseline condition had significantly higher RMSE than the small unpressurized, prescribed unpressurized, and large pressurized gloved conditions (r =.46 to r = .50). As with the pitch data, no significant differences between glove sizes were detected for either the unpressurized or pressurized conditions. The boxplots for the post-40 second probes in both axes can be found in Appendix B.1.

5.1.2 Response Time Subtask Results

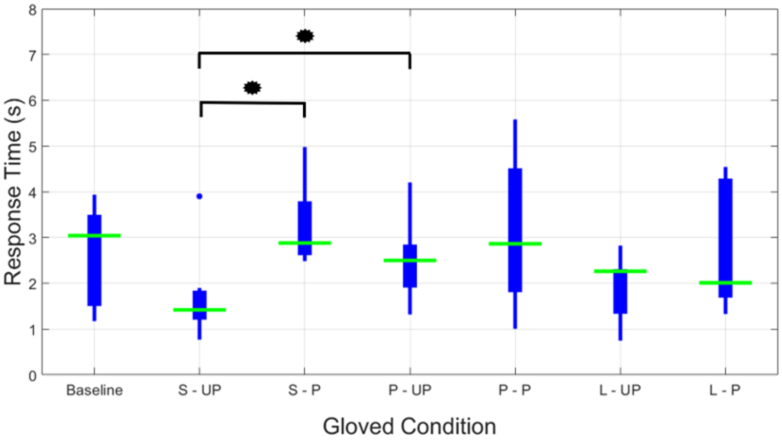


Figure 5-3: Response time for each suited condition. Significant uncorrected post-hoc pairwise comparisons are notated with an asterisk ($p < 0.05$), none are significant when corrected using FDR.

For the response time subtask, a Friedman test supported an effect of gloved condition on response time ($\chi^2(6) = 17.33$, p = 0.0082) (Figure 5-3). Post-hoc pairwise comparisons support that the pressurized condition had significantly higher response times than the unpressurized condition for the small glove size (r = .63) and that for the unpressurized condition, the small glove size had significantly lower response time than the prescribed glove size (r = .63). While not tested formally, it can be qualitatively observed that the pressurized cases have larger interquartile ranges than the unpressurized cases for the prescribed and large glove sizes.

5.2 Lunar Landing Simulator Discussion

The following section discusses the conclusions regarding the effects of spacesuit gloves on the performance of the aforementioned subtasks.

5.2.1 Primary Flight Subtask Discussion

The results of the primary flight subtask in the pitch axis support that the pressurized condition resulted in lower RMSE than the corresponding unpressurized condition for all glove sizes at the 40 second probe. This finding appears counter intuitive as pressurization led to decreased performance on both of the dexterity tasks in this study, as well as other dexterity tasks in the literature [1, 54]. A qualitative review of video footage of subjects conducting the lunar landing simulator task was conducted and it was noted that pulling back on the rotational hand controller to pitch the spacecraft upwards required more effort in the pressurized condition due to the added stiffness of the glove. Additionally, most subjects shifted their hand placement on the rotational hand controller from the stem of the joystick during the unpressurized trials to the head of the joystick during the pressurized trials. A potential explanation for the decreased RMSE when pressurized is that the added stiffness of the pressurized condition resulted in smaller corrections and thus less RMSE in the pitch axis due to limiting any overcorrections. Although no previous literature specifically examines the effect of glove stiffness on flight technical performance, work by Itaguchi and Fukuzawa [72] did find both lower constant and variable errors in the direction of higher arm stiffness on a multi-joint position reproduction task.

The results of the primary flight subtask in the roll axis indicate significant decreases in error between the baseline condition and 4 of the gloved conditions. Upon further examination of the unsorted glove trials (utilizing the random order each subject received glove sizes) it can be seen that for the 40 second probe (Figure 5-4) there is a gradual decrease in RMSE as the study progresses across subjects (this same trend is observed for subsequent probes) indicating the prescence of learning effects in the roll axis that were not observed in the pitch axis. Although subjects were given as many training trials as the 3-hour study period would allow, it appears that in future studies, more training runs are necessary for the flight performance subtask for this control axis. When examining the overall dataset in the roll axis, no significant effect of glove size or pressurization is observed. This finding could be influenced by the increased variability across the sorted trials due to the learning effects and the randomized fashion in which subjects received glove sizes.

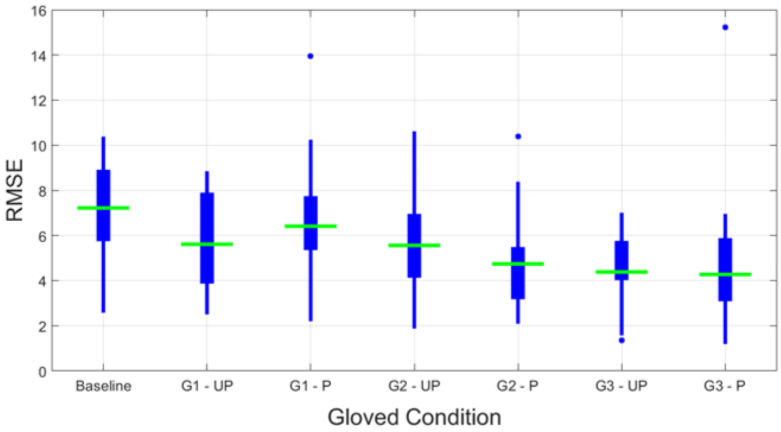


Figure 5-4: RMSE in the roll axis for the unsorted gloved conditions at the 40 second probe. Gradual decrease in error over the course of all trials indicates learning effects

5.2.2 Response Time Subtask Discussion

For the response time subtask, the results demonstrate that for the small glove size, pressurization led to significantly higher response time than the unpressurized condition. However, it can also be qualitatively observed from Figure 5-3 that there was more variation in response times for the pressurized cases for the prescribed and large glove sizes, indicating that different individual subjects experienced a range of mental workload during the task. Llanos et al. [63] previously found a wide range of mental workload ratings (ranging from 10 to 90 on a 100 point scale) as assessed by the NASA TLX scale for subjects conducting a flight simulator task in a pressurized IVA suit of fixed sizing. These data support the importance of individual differences on mental workload during a flight simulator task when wearing pressurized spacesuit gloves. These variations in response time could also be due to different motor strategies across subjects for responding to the communication indicator in the pressurized case, as different hand and finger placements were observed during a qualitative review of video footage of subjects completing the task.

With respect to glove fit on the response time task, it was found that for the unpressurized case, the prescribed glove size resulted in significantly slower response times than the small glove size, a result in-line with the initial hypothesis that change in static fit would effect performance on a mental workload assessment task; however, these timing results are not consistent with the ungloved value and the larger glove size. Further efforts are warranted to examine how sizing affects measures of workload derived from secondary performance measures. Tremblay-Lutter et al. [12] also found that increased easement led to decreased performance in chemical protective gloves, although it is unclear if these results are comparable as that work examined dexterous performance, not mental workload. Additionally, Llanos et al. [63] did find that when wearing a pressurized IVA suit while flying a spaceflight simulator, 28% percent of subjects reported mental demands higher than moderate levels as assessed by the NASA TLX scale indicating that spacesuit gloves can play a role in mental workload when performing a flight task.

Chapter 6

Conclusions, Limitations and Future Work

The final chapter of this thesis summarizes the key conclusions of the research effort. Additionally the limitations of the study and avenues for future work in this area are discussed.

6.1 Research Overview

As human spaceflight again ventures beyond LEO, there will be an accompanying increase in the frequency and complexity of suited operations. To facilitate this increase in diverse suited activity, it is important that spacesuits enable acceptable performance for a wide variety of tasks while also mitigating the risk of injury. Spacesuit fit is a key factor in enabling this acceptable performance, and while spacesuit fit has been incorporated into the suit development process in a subjective manner, knowledge gaps still exist in terms of qunatifying the relationships between spacesuit fit and functional task performance. This research effort sought to investigate the relationship between fit and performance towards the goal of enabling improved functional task performance on future space missions.

This study specifically investigated the relationship between static fit in IVA gloves and functional task performance. This human study utilized a battery of both generalizable and operationally relevant performance assessment tasks assessing the areas of tactility, dexterity, technical flight performance, and mental workload. To assess fit, direct measures of static fit and subjective surveys were employed. In order to mimic the pressurization of the full spacesuit, a glovebox approach was used for the evaluations.

6.2 Key Conclusions

6.2.1 Contributions to the Literature

- There is a positive relationship between static fit and switchboard tactility task score for the range of easement examined (with greater easement leading to higher scores) and the direct measure of static fit with respect to hand length proved useful for quantifying this relationship. Based on these results, it may be better to receive a larger glove size as limited easement can adversely affect tactile performance.

- The turntable task results showed the highest error for the bumps of low height and the switchboard task results showed the lowest accuracy scores for the section containing the low-profile push buttons. These tasks support a minimum height profile requirement above 0.05 in for low-profile push buttons.

- No effects of glove size on dexterity task performance were detected for the specified task pacing, however faster pacing could evaluate potential differences at the limits of dexterity.

- No effects of glove size on the lunar landing simulator task performance were detected, however the added stiffness from glove pressurization may have improved technical flight performance in the pitch axis compared to the unpressurized condition.

- For the switchboard tactility task, the gloved unpressurized case yielded worse performance than the ungloved baseline case, reaffirming the results previous of literature that the addition of spacesuit gloves cab inhibit tactile performance [54, 56]. For the switchboard tactility task and the u-bolt and EVA tether dexterity tasks, the pressurized case yielded decreased performance compared the unpressurized case for each glove size. This reaffirms the results of previous literature [1, 54, 56] showing that spacesuit glove pressurization can inhibit tactile and dexterous performance.

6.2.2 Tactility Performance Assessment

The most significant takeaway from the examination of tactile performance in this research effort is the relationship between static fit and performance on the switchboard tactility task. The results of the LME support that for the range of glove sizes and easement examined in this study, there is a positive relationship between direct measures of static fit with respect to hand length and the score on the switchboard tactility task, with greater easement leading to higher scores. These results agree with the post-hoc comparisons of the switchboard tactility results with respect to sizing, as in the unpressurized cases the large glove had significantly higher scores than both the small and prescribed gloves. These results indicate that it may be better to receive a larger glove size as limited easement can adversely affect tactile performance.

Another key conclusion that can be drawn from the tactility performance assessments is the importance of control height to tactile performance. The effect of height can be seen both on the turntable task (where the low profile pads resulted in the highest frequency of incorrect answers) and the switchboard task (where the low profile push buttons had significantly lower section scores than the other controls). These results demonstrate the importance of height in the design of controls that may have to be utilized via a crewmember's sense of touch (such as during an emergency scenario with reduced visibility). For low-profile keyboard buttons NASA lists a minimum height of 0.04 in (although the preferred height is listed as 0.08 in) [71]. For similarly low profile legend buttons, no minimum height is specified (although the associated barrier depth is 0.2 in) [71]. These tactility results indicate that it may be important to clearly specify a minimum height profile above 0.05 in for these types of controls. It should be noted that the displacement values for isolated push buttons have minimums listed above 0.05 in [71].

Lastly, the results from the switchboard tactility task show that for each glove size, the pressurization of the glove leads to a lower score when compared to the unpressurized case. The LME also shows that pressurization state has a significant relationship with switchboard score. These results align with previous literature [54, 56] showing that the pressurization of spacesuit gloves inhibits tactile performance.

6.2.3 Dexterity Performance Assessment

Unlike the switchboard tactility task, no significant relationship between static fit (as assessed by glove size) and performance was found in either the u-bolt pegboard task or the EVA tether task for dexterity assessment. However, it can be seen that in the pressurized case, there is a wider interquartile range of performance for the large glove across both tasks. This interquartile range indicates that factors such as increased easement or increased stiffness in the pressurized case may have inhibited performance for some subjects, but the same trends are not found in the unpressurized cases. As increased easement did not appear to negatively affect dexterity for the sizes evaluated, the recommendation of a larger size from the tactility studies is still appropriate.

For both dexterity tasks, performance in the unpressurized cases was significantly better compared to the respective pressurized case for each glove size. These results agree with previous literature [1, 54] that shows that the pressurization of spacesuit gloves have an adverse effect on dexterous performance.

However, these results show that for both dexterity tasks there was no significant difference in performance between ungloved conditions and unpressurized gloved conditions. This results differs from previous literature in both IVA and EVA gloves [52, 54]. It is important to note that tasks used in previous studies were self paced [1, 54], whereas the u-bolt pegboard task was limited by the metronome at 50 BPM, which influenced the pace selected and the outcomes observed. Equivalent performance between ungloved and gloved conditions is an operational goal of IVA gloves, and these results could represent a fufillment of that goal for these particular tasks.

6.2.4 Flight Performance and Mental Workload Assessment

Neither the primary flight performance assessment subtask nor the response time mental workload assessment subtask had a significant relationship with static fit (as assessed by glove size). As with the dexterity tasks, this result does not conflict with the recommendation of increased easement for better tactile performance. Further potential investigations into the relationship between both flight performance and mental workload and static fit are discussed in the future work section of this chapter.

The key trend observable in the lunar landing simulator data is that in the pitch axis for the 40 second probe, each pressurized case had lower RMSE than the corresponding unpressurized case for each glove size. This improved performance could potentially be due to the added stiffness of the pressurized gloves mitigating overcorrections during movement, as previous literature [72] has shown that arm stiffness led to increased performance on a position reproduction task. However, it is important to note that statistical results for this finding are only of medium effect size and do not remain significant after undergoing the FDR correction.

A final takeaway from the lunar landing simulator results is that further training is necessary to mitigate learning effects. Figure 5-4 shows that there was a steady improvement in RMSE as subjects spent more time on the task, indicating that further training is necessary to ensure steady-state baseline performance in terms of nulling errors and to isolate the effect of the study's independent variables. Additionally, utilizing a certain level of piloting experience as an inclusion criteria or incorporating a quantified metric of piloting experience into analyses (neither of which were done in this study) could have further impact on these results.

6.2.5 Static Fit Assessment

The significant relationship between static fit and switchboard tactility task score shows that the direct static fit metric with respect to hand length employed in this study is a useful measure for quantifying the relationship between fit and tactile performance. Additionally, the post-hoc comparisons results for glove size on the switchboard tactility tasks agree with the results of the LME. In this study, a certain range of hand anthropometry was used as an inclusion criteria, leading to a roughly similar relationship between anthropometry and each glove size across subjects, supporting the use of glove size as a coarse metric of fit for the other tasks.

The subjective surveys for fit and comfort were able to provide qualitative insights into aspects of fit. The fit perception survey was able to show that tightness in the fingers occurred mainly in the small and prescribed unpressurized conditions, whereas perceived looseness was more likely in the pressurized conditions, particularly the large pressurized case. The comfort survey was useful in identifying common areas of discomfort, such as the index finger knuckle. However, both of these surveys could use improvement to increase their fidelity, as is discussed in the following sections of this chapter.

6.3 Limitations and Future Work

The limitations of this study as well as areas for future improvement and further research are discussed in the following section.

6.3.1 Experimental Design

One limitation of this study is the relatively narrow range of static fit considered. For example, while the tactility assessment results provide interesting insight into the relationship between fit and performance, they only apply to the limited range of static fit examined in this study and cannot be generalized. Therefore an important area of future work would be to conduct further studies that utilize both more glove sizes and a wide range of human anthropometry. A more comprehensive examination of performance across a wider range of fit conditions would provide more insight than can be gained from our preliminary study.

An additional limitation of this study inherent to its design is the use of the glovebox. While the glovebox provides logistical benefits in terms of easily donning and doffing gloves, transitioning glove fits, and quickly pressurizing and depressurizing the chamber, there are benefits to performing testing in a full spacesuit. The use of a full suit provides a more representative view of spacesuit fit and biomechanics during real operations as it includes the interactions between the suit torso and arm components. While this study incorporated measures such as specifications for subject placement and posture to mitigate these limitations, it would still be a useful area of future work to repeat this same experimental protocol in a full spacesuit to compare the results.

6.3.2 Performance Assessment Tasks

For the assessment of tactility in this study, the turntable task proved limited in its ability to provide insights into the differences between gloved conditions. Revisions to this task, such as the addition of pressure sensor measurements to record the force with which the bumps are pressed, as well as expanding the range of sizes of the bumps, could be implemented to help identify differences in strategy between gloved conditions on this task.

As mentioned previously, no differences in performance with respect to glove size or between the ungloved and gloved unpressurized conditions were found utilizing the dexterity assessment tasks in this study. While this result supports the goals of a high-performance IVA glove, it is important to consider that the u-bolt task was intentionally designed with a specific pacing to mitigate learning effects and may not have been assessing the limits of dexterity. Particularly for the u-bolt pegboard task, the imposed pacing by the metronome may have inhibited the detection of differences between gloved conditions that could have been revealed at faster pacing. Therefore, for dexterity the tasks it may prove beneficial to increase pacing to examine the limit of dexterity in each condition.

For the flight simulator task, it is clear from the learning effects observed in the roll-axis that it is important to provide further training trials for this task. In future studies, it will be important to expand overall study time for the flight simulator and provide supervised training of an extended duration to ensure that steady-state baseline performance is achieved so that the study is not confounded by learning effects. Additionally, a limitation of this study was the wide range of piloting experience among the subjects (ranging from no previous experience with flight simulators to licensed pilots). Including piloting experience in the inclusion criteria for future studies could help improve the consistency of the results.

6.3.3 Fit Assessment

The static fit metric selected examined only one dimension. Future efforts should consider additional dimensions of static fit by calculating additional ease metrics for other anthropomorphic dimensions such as hand circumference, or digit length or circumference. Additionally, while a direct measure of static fit was used to assess the relationship between fit and performance for the tactility tasks, only the coarse metric of glove size was utilized for the dexterity and lunar landing simulator tasks. With respect to examining the correlation between static fit and performance, it is important to note that the direct measure of static fit provides a finer interpretation of fit beyond glove size as the direct ease metric incorporates the individual hand size, which varies, and glove size together. Future efforts should work to develop a linear mixed effect model that incorporates a direct static fit metric, as well as subject, to gain a more in-depth understanding of the relationship between fit and performance on the lunar landing simulator and dexterity assessment tasks. The fidelity of these models would be improved by increased task repetitions to increase sample size.

As mentioned in Chapter 3, the range of ratings that subjects provided on the fit perception survey was narrow, limiting the ability of this survey to provide insights into the relationship between perceived fit and performance, or the relationship between perceived fit and direct measures of static fit. Therefore, the fit perception survey could also be expanded to include a greater range of fit ratings to increase survey fidelity. Lastly, further quantitative analysis of discomfort ratings could be conducted to gain a more comprehensive understanding of subjects' perception of the gloves.

6.4 Final Thoughts

The future of human space exploration will entail extensive use of spacesuits in harsh planetary environments on missions that incorporate a variety of complex tasks. It is therefore important that spacesuits enable acceptable performance on the diversity of tasks that crewmembers will be required to perform, while also mitigating injury risk and discomfort. A key factor in ensuring this acceptable suited performance is spacesuit fit. While spacesuit fit has long been acknowledged as important and has been incorporated into the suit development process in a subjective manner, knowledge gaps exist in terms of quantifying the relationship between spacesuit fit and functional task performance.

This work sought to advance the goal of quantifying the relationship between fit and performance in IVA spacesuit gloves. The key performance areas of dexterity, tactility, flight technical error, and mental workload were all examined for the relationship to spacesuit glove fit. Both direct measures of static fit with respect to hand length and subjective fit perception and comfort surveys were employed to assess subject fit. For the operational switchboard tactility task, a relationship between direct static fit and task performance was quantified, showing that for the gloves considered, increased easement lead to better task performance in the unpressurized case. The other performance assessment tasks did not reveal definitive relationships between fit and performance, although they did yield insights into the effects of pressurization and spacesuit gloves on task performance and motor technique.

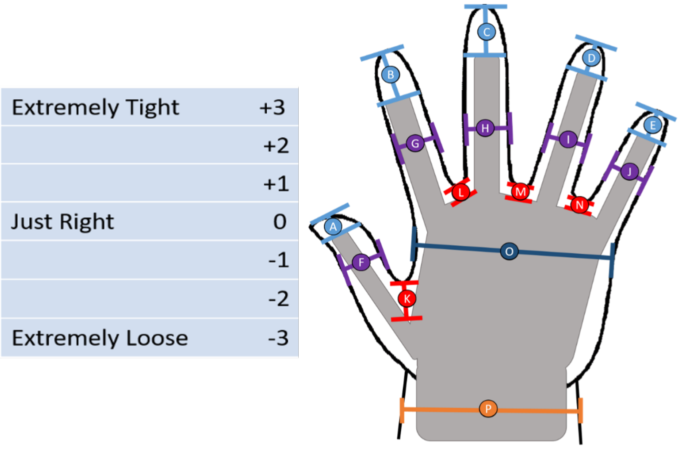
Going forward, there are improvements that can be implemented to this study in order to provide further insights into the relationship between IVA spacesuit glove fit and performance. These include expanding the range of glove sizes considered, augmenting the performance assessment tasks and surveys, and making use of full spacesuit testing. Beyond these improvements, future studies should seek to expand efforts to quantify the relationship between spacesuit fit and performance. These efforts could include examining other portions of the suit and analyzing EVA suits in addition to IVA suits. Future studies should also consider other domains of fit (such as dynamic fit) and alternate means of quantifying fit (such as wearable sensors). Continued efforts to quantify the relationship between spacesuit fit and performance will yield insight that improves the performance and safety of crewmembers as human spaceflight expands beyond LEO to more complex and extended planetary exploration missions.

Appendix A

Tables

Table A.1: Fit perception survey results in the right hand on Day 1. Results are averaged across subjects who completed tactility tasks. The range of possible fit perception ratings is -3 to 3. Figure A-1 below shows the fields laid out on the survey diagram.

	Gloved Condition					
Hand Location	S - UP	S - P	P - UP	P - P	L - UP	L - P
A - thumb tip	0	-0.43	0.14	-0.43	0.14	-1.14
B - index finger tip	0.43	0	0.57	0	-0.14	-1
C - middle finger tip	0.57	-0.43	0.71	-0.14	-0.14	-1
D - ring finger tip	0.57	0	0.57	0.43	0	-0.43
E - pinky finger tip	0.57	0.14	1	0.71	0	-0.86
F - thumb knuckle	-0.14	-1.43	-0.14	-1	-0.29	-1.71
G - index finger knuckle	0	-1.14	0.14	-0.57	0	-1.29
H - middle finger knuckle	0.14	-1.14	0.14	-0.14	0	-1.43
I - ring finger knuckle	0.29	-0.71	0.14	-0.29	-0.14	-0.71
J - pinky finger knuckle	0.14	-0.43	0.71	-0.29	0	-0.57
K - thumb crotch	0.14	-0.43	0	-0.57	-0.14	-0.14
L - index finger crotch	0	-0.29	0.29	0.14	0.14	0
M - middle finger crotch	0.14	-0.14	0.14	-0.14	0.14	0
N - ring finger crotch	0.14	-0.29	0.57	-0.14	0	-0.14
O - hand width	0.14	0.14	0.29	0.29	0.14	0
P - wrist	-0.71	-0.86	-0.57	-0.86	-0.43	-1.29



Extremely Tight	+3
	+2
	+1
Just Right	0
	-1
	-2
Extremely Loose	-3

Figure A-1: Subjective fit survey with areas of evaluation and range of possible fit ratings.

Table A.2: Comfort survey results for unpressurized gloves in the right hand on day 1. Blank fields (indicating no discomfort) are not listed.

Gloved Condition	Subject	Sensation	Location	Intensity
S - UP	1	chafing	E	1
	1	touching	O	1
	3	touching	B	1
	3	pressure point	O	1
	4	pressure point	O	1
	7	chafing	C	1
	7	touching	O	2
P - UP	3	touching	A	1
	3	touching	B	1
	3	touching	C	1
	3	touching	D	1
	3	pressure point	O	2
	4	pinching	O	1
	5	pinching	B	1
	5	pressure point	E	1
	6	pressure point	O	1
	7	chafing	C	3
	7	touching	O	3.5
L - UP	1	chafing	E	1
	4	pressure point	K	1
	5	pressure point	D	1
	5	pressure point	E	1
	6	pressure point	C	1
	6	chafing	C	1
	7	touching	O	1

Table A.3: Comfort survey results for pressurized gloves in the right on hand day 1. Blank fields (indicating no discomfort) are not listed.

Gloved Condition	Subject	Sensation	Location	Intensity
S - P	1	pressure point	E	1
	1	pressure point	O	1
	1	chafing	O	2
	3	touching	B	1
	3	pressure point	O	2
	4	pressure point	O	1
	5	pressure point	B	1
	5	chafing	O	1
	6	pressure point	C	2
	6	pressure point	L	2
	6	pressure point	O	2
	7	chafing	E	1
	7	touching	O	2.5
P - P	1	chafing	E	0.5
	1	pressure point	O	2
	2	pressure point	O	2
	3	touching	B	1
	3	touching	D	1
	3	pressure point	O	3
	4	pinching	O	2
	5	pinching	E	1
	5	chafing	L	1
	6	pressure point	D	3
	6	chafing	D	3
	6	pressure point	L	2
	6	pressure point	O	2
	7	pressure point	O	3.5
L - P	1	chafing	E	1
	1	pressure point	O	1
	2	pressure point	O	1
	3	pressure point	O	1
	4	pressure point	K	2
	5	chafing	O	1
	6	chafing	D	2
	6	pressure point	L	2
	7	chafing	C	2
	7	touching	O	3

Appendix B

Figures

Figure B-1: RMSE in pitch axis for each gloved condition at the 50 second probe. Significant uncorrected post-hoc pairwise comparisons are notated with an asterisk ($p < 0.05$), none are significant when corrected using FDR.

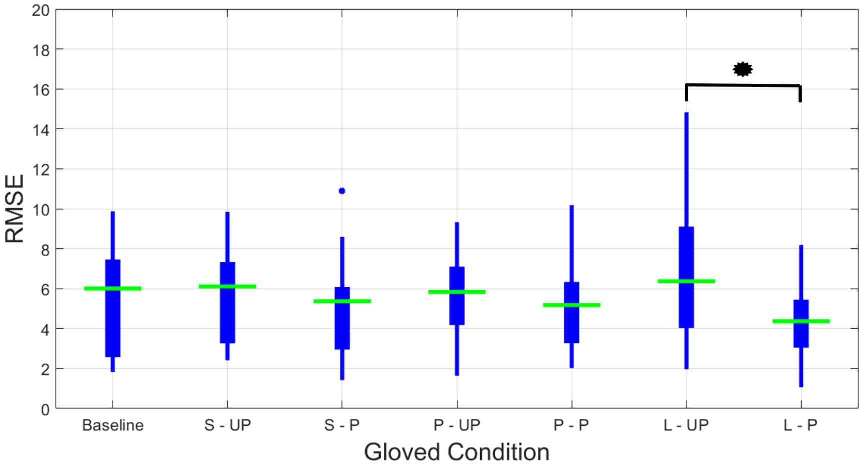


Figure B-2: RMSE in pitch axis for each gloved condition at the 60 second probe. Significant uncorrected post-hoc pairwise comparisons are notated with an asterisk ($p < 0.05$), none are significant when corrected using FDR.

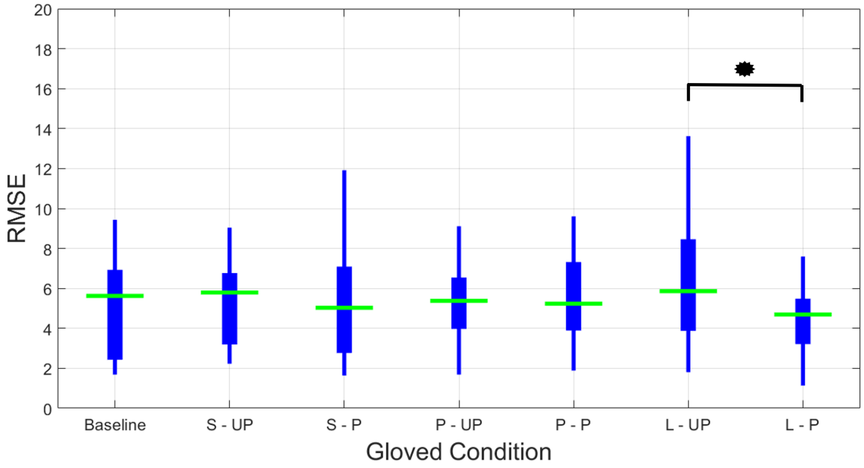


Figure B-3: RMSE in pitch axis for each gloved condition at the 70 second probe. Significant uncorrected post-hoc pairwise comparisons are notated with an asterisk ($p < 0.05$), none are significant when corrected using FDR.

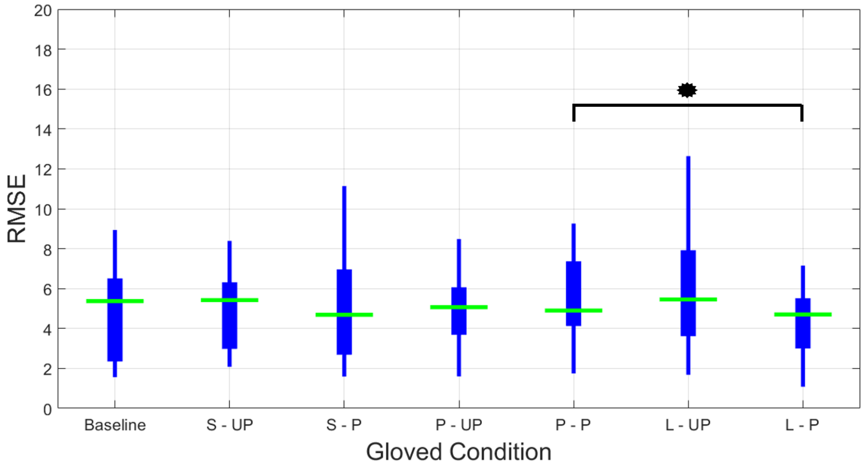


Figure B-4: RMSE in roll axis for each gloved condition at the 50 second probe. Significant ($p<0.05$) FDR corrected post-hoc pairwise comparisons are notated with an asterisk

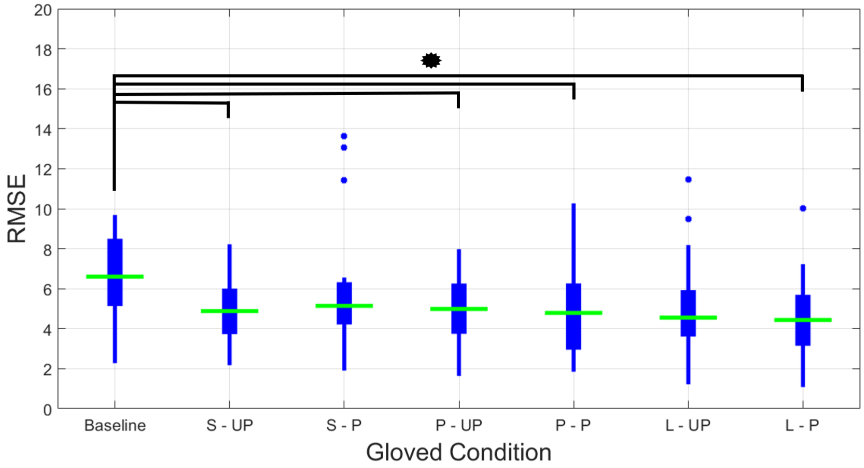


Figure B-5: RMSE in roll axis for each gloved condition at the 60 second probe. Significant ($p<0.05$) FDR corrected post-hoc pairwise comparisons are notated with an asterisk

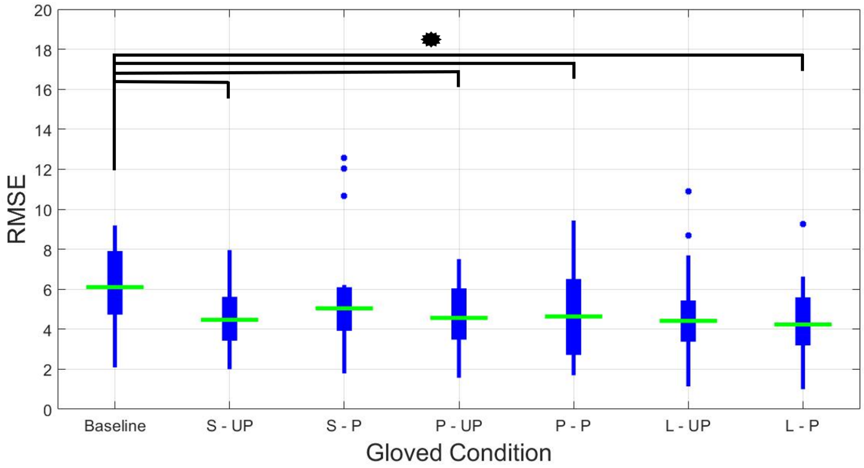


Figure B-6: RMSE in roll axis for each gloved condition at the 70 second probe. Significant ($p<0.05$) FDR corrected post-hoc pairwise comparisons are notated with an asterisk

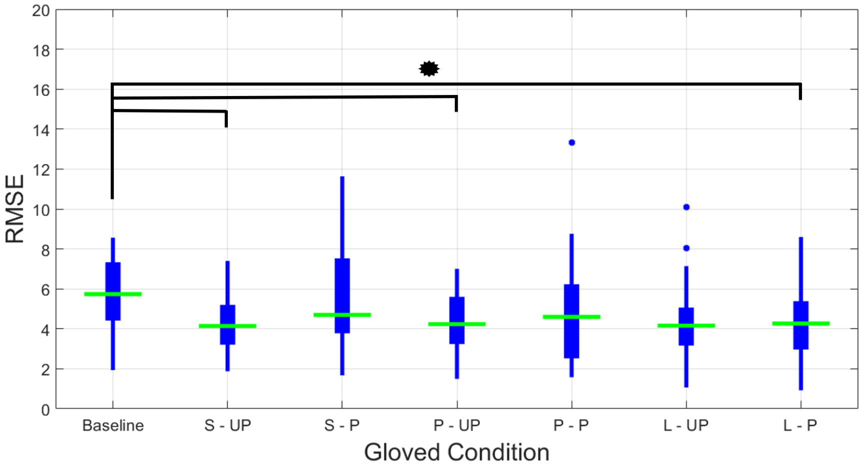
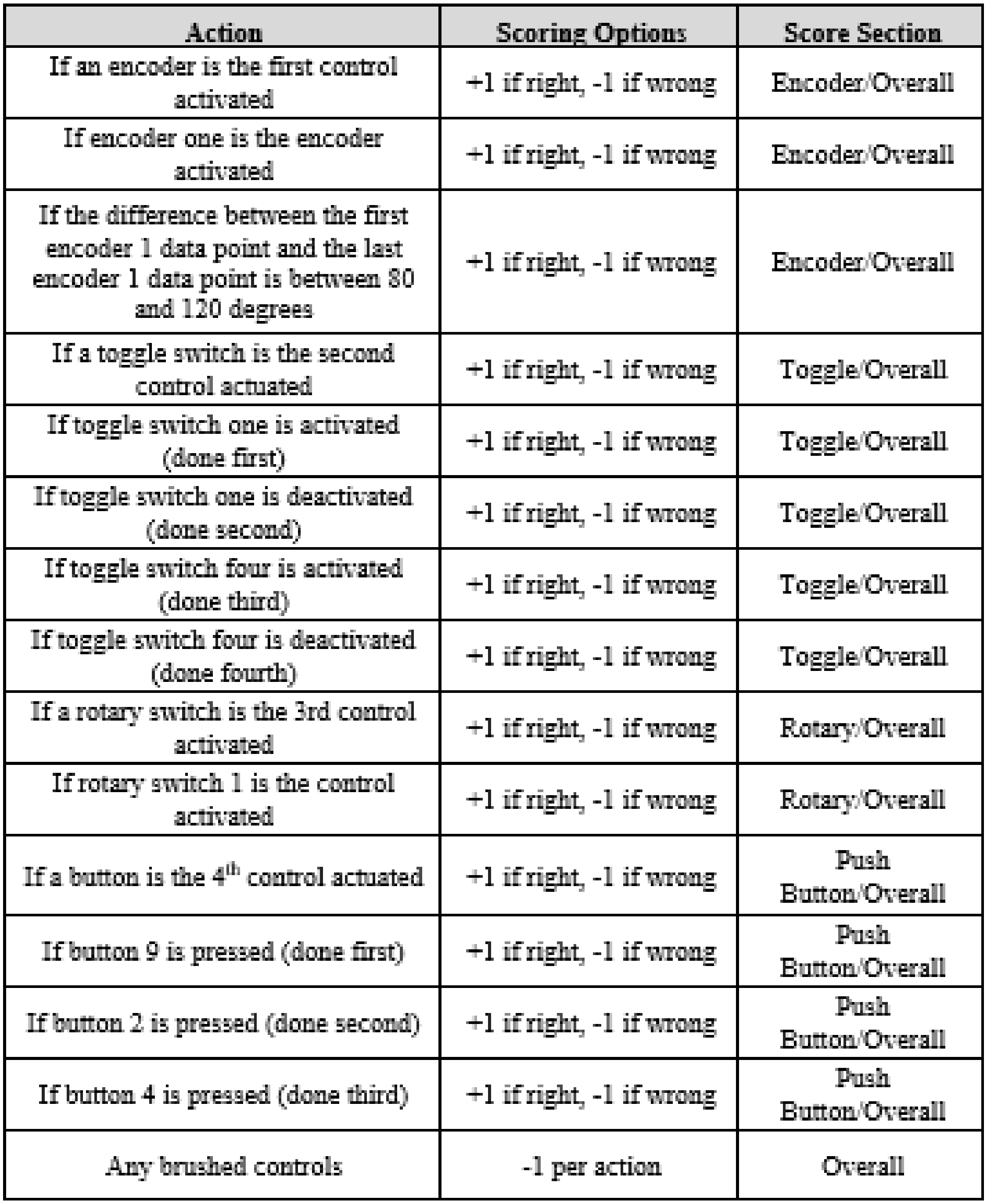


Figure B-7: Switchboard scoring rubric. Maximum possible score is 14.



Action	Scoring Options	Score Section
If an encoder is the first control activated	+1 if right, -1 if wrong	Encoder/Overall
If encoder one is the encoder activated	+1 if right, -1 if wrong	Encoder/Overall
If the difference between the first encoder 1 data point and the last encoder 1 data point is between 80 and 120 degrees	+1 if right, -1 if wrong	Encoder/Overall
If a toggle switch is the second control actuated	+1 if right, -1 if wrong	Toggle/Overall
If toggle switch one is activated (done first)	+1 if right, -1 if wrong	Toggle/Overall
If toggle switch one is deactivated (done second)	+1 if right, -1 if wrong	Toggle/Overall
If toggle switch four is activated (done third)	+1 if right, -1 if wrong	Toggle/Overall
If toggle switch four is deactivated (done fourth)	+1 if right, -1 if wrong	Toggle/Overall
If a rotary switch is the 3rd control activated	+1 if right, -1 if wrong	Rotary/Overall
If rotary switch 1 is the control activated	+1 if right, -1 if wrong	Rotary/Overall
If a button is the 4th control actuated	+1 if right, -1 if wrong	Push Button/Overall
If button 9 is pressed (done first)	+1 if right, -1 if wrong	Push Button/Overall
If button 2 is pressed (done second)	+1 if right, -1 if wrong	Push Button/Overall
If button 4 is pressed (done third)	+1 if right, -1 if wrong	Push Button/Overall
Any brushed controls	-1 per action	Overall

Bibliography

[1] A. Korona, S. Walsh, and S. McFarland. Nasa EVA glove characterization protocol development. In *International Conference on Environmental Systems*, Albuquerque, NM, 2018.

[2] S. Chappell, J. Norcross, A. Abercromby, O. Bekdash, E. Benson, and S. Jarvis. Risk of injury and compromised performance due to EVA operations. Technical report, Houston, Texas, 2017.

[3] Richard Scheuring, Charles Mathers, Jeffrey Jones, and Mary Wear. Musculoskeletal injuries and minor trauma in space: Incidence and injury mechanisms in us astronauts. *Aviation, space, and environmental medicine*, 80:117–24, 03 2009.

[4] NASA. Eva gap 7b: How does suit sizing and fit affect crew health and performance in the exploration environments? https://humanresearchroadmap.nasa.gov/gaps/gap.aspx?i=653, 2019.

[5] S. Jarvis, J. Norcross, and A. Abercromby. Subjective suit fit assessment. In *Human Research Program Investigators' Workshop*, Galveston, Texas, 2017.

[6] H. J. Choi, K. B. Mitchell, T. N. Garlie, and L. DeSimone. Effects of body armor fit on Warfighter mobility as measured by range of motion (ROM). *Advances in Intelligent Systems and Computing*, 789, 2019.

[7] J. Huck, O. Maganga, and Y. Kim. Protective overalls: evaluation of garment design and fit. *International Journal of Clothing Science and Technology ; Bradford*, 9(1):45–61, 1997.

[8] E. Benson and S. Rajulu. Complexity of sizing for space suit applications. In *Digital Human Modeling: Second International Conference*, San Diego, California, 2009.

[9] Matthew Schwartz. Nasa scraps first all-female spacewalk for want of a medium-size spacesuit. https://www.npr.org/2019/03/26/706779637/nasa-scraps-first-all-female-spacewalk-for-want-of-a-medium-sized-spacesuit, 2019.

[10] H. J. Choi, G. F. Zehner, and J. A. Hudson. A manual for the performance of protective equipment fit-mapping. Technical Report AFRL-RH-WP-SR-2010-0005, 2009.

[11] L. Stirling, D. Kelty-Stephen, R. Fineman, M. Jones, B. K. Park, M. Reed, J. Parham, and H. Choi. Static, dynamic, and cognitive fit of exosystems for the human operator. *Human Factors*, 31, January 2020.

[12] J. Tremblay-Lutter and S. Weihrer. Functional fit evaluation to de-termine optimal ease requirements in chemical protective gloves. *Performance of Protective Clothing*, 5:367–383, 1996.

[13] J. Tremblay-Lutter, E. Crown, and K. Rigakis. Evaluation of functional fit of chemical protective gloves for agricultural workers. *Clothing and Textiles Research Journal*, 14(3), 1996.

[14] H. J. Choi, G. F. Zehner, and T. Metzger. F-35 protective equipment fit assessment: Light weight coverall. Technical Report AFRL-RH-WP-TR-2011-0055, 2011.

[15] R. A. Fineman, T. M. McGrath, D. G. Kelty-Stephen, A. F. Andrew, and L. A. Stirling. Objective metrics quantifying fit and performance in spacesuit assemblies. *Aerospace Medicine and Human Performance*, 89:985–995, November 2018.

[16] L. Stirling, P. Arezes, , and A. Anderson. Implications of space suit injury risk for developing computational performance models. *Aerospace medicine and human performance*, 90(10):553–565, 2019.

[17] C. R. Reid, L. R. Harvill, J. R. Norcross, E. A. Benson, S. A. England, K. Young, and S. L. Rajulu. An ergonomic evaluation of the Extravehicular Mobility Unit (EMU) spacesuit Hard Upper Torso (HUT) size effect on mobility, strength, and metabolic performance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 58, 2014.

[18] Particia Schmidt. *An Investigation of Space Suit Mobility with Applications to EVA Operations*. PhD thesis, Massachusetts Institute of Technology, 2001.

[19] L. L. M. Bossi, M. L. H. Jones, A. Kelly, and D. W. Tack. A preliminary investigation of the effect of protective clothing weight, bulk and stiffness on combat mobility course performance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60, 2016.

[20] C. R. Cullinane, R. A. Rhodes, and L. A. Stirling. Mobility and agility during locomotion in the Mark III space suit. *Aerospace Medicine and Human Performance*, 88, 2017.

[21] A. P. Anderson and D. J. Newman. Pressure sensing for in-suit measurement of space suited biomechanics. *Acta Astronautica*, 115, 2015.

[22] C. Compton and L. Dunne. Pilot investigation of a novel technique for measuring dynamic body-garment contact. In *International Conference on Environmental Systems*, Vienna, Austria, 2016.

[23] S. Lombardo, R. Fineman, T. McGrath, and L. Stirling. Evaluating Human Dynamic Fit within the EMU During an Upper Body Cycle Ergometer Task. In *International Conference on Environmental Systems*, Boston, MA, 2019.

[24] J. Bell-Krotoski and E. Tomancik. The repeatability of testing with Semmes-Weinstein monofilaments. *The Journal of Hand Surgery*, 12, 1987.

[25] M. Casadio, R. Iandolo, S. Nataletti, F. Marini, P. Morasso, V. Pnassi, and R. A. Scheidt. Robotic techniques for the assessment of proprioceptive deficits and for proprioceptive training. *Rehabilitation robotics: Technology and application*, pages 289—-303, 2018.

[26] L. Stirling, H. C. Siu, E. Jones, and K. Duda. Human factors considerations for enabling functional use of exosystems in operational environments. *IEEE Systems Journal*, 13, 2019.

[27] D. Kahneman. *Attention and effort.* Englewood Cliffs, N, J, 1973.

[28] C. D. Wickens. Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3, 2002.

[29] G. Ogden, J. Levine, and E. Eisner. Measurement of workload by secondary tasks. *Human Factors*, 21(5):529–548, October 1979.

[30] R. O'Donnell and T. Eggemeier. Workload assessment methodology. *Handbook of Perception and Human Performance*, II:42–49, 1994.

[31] S. G. Hart and L. E. Staveland. Development of the NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 52, 1988.

[32] Mica Endsley. Endsley, m.r.: Toward a theory of situation awareness in dynamic systems. human factors journal 37(1), 32-64. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37:32–64, 03 1995.

[33] M. Fracker. Measures of situation awareness: Review and future directions. Technical report, Wright-Patterson Air Force Base, Ohio, 1991.

[34] Ruiqi Ma and David Kaber. Situation awareness and workload in driving while using adaptive cruise control and a cell phone. *International Journal of Industrial Ergonomics*, 35:939–953, 09 2005.

[35] C. J. Hainley, K. R. Duda, C. M. Oman, and A. Natapoff. Pilot performance, workload, and situation awareness during lunar landing mode transitions. *Journal of Spacecraft and Rockets*, 50(4):793–801, 2013.

[36] B. Bequette, A. Norton, E. Jones, and L. Stirling. The effect of a powered lower-bodyexoskeleton on physical and cognitive warfighter performance. In *Human Factors and Ergonomics Society Annual Meeting*, Philadelphia, PA, 2018.

[37] Matthew S. Cowley, Sarah Margerum, Lauren Hharvill, and Sudhakar Rajulu. Model for predicting the performance of planetary suit hip bearing designs. In *International Conference on Applied Human Factors and Ergonomics*, San Francisco, CA, 2012.

[38] Matthew Gast and Sandra Moore. A glimpse from the inside of a space suit: What is it really like to train for an eva? *Acta Astronautica*, 68:316–325, 02 2011.

[39] S. E. Jacobs, D. B. Tufts, and D. M. Gohmert. Space suit development for orion. In *International Conference on Environmental Systems*, Albuquerque, New Mexico, 2018.

[40] O. Bekdash, J. Norcross, J. Fricker, M. Young, I. Meginnis, and A. Abercromby. Validation of inspired carbon dioxide measurement methods in the extravehicular mobility unit space suit. In *International Conference on Environmental Systems*, Albuquerque, NM, 2018.

[41] S. Anderson, C. Kaufman, and K. Johnson. Urine removal from suited crew in orion vehicle depressurization scenario. In *International Conference on Environmental Systems*. Boston, MA, 2019.

[42] C. Watts and M. Vogel. Space suit portable life support system 20 human-in-the-loop testing,. In *International Conference on Environmental Systems*, Vienna, Austria, 2016.

[43] Christopher Carr. *The Bioenergetics of Walking and Running in Space Suits*. PhD thesis, Massachusetts Institute of Technology, 2005.

[44] P. J. Bertrand, A. Anderson, A. Hilbert, and D. Newman. Feasibility of spacesuit kinematics and human-suit interactions. In *International Conference on Environmental Systems*, Tucson, Arizona, 2014.

[45] ILC Dover. Ilc dover suit modeling, sizing and performance evaluations. In *NASA EVA Workshop*. Galveston, TX, 2017.

[46] R. Opperman, J. Waldie, A. Natapoff, D. Newman, and J. Jones. Probability of spacesuit-induced fingernail trauma is associated with hand circumference. *Aviation Space and Environmental Medicine*, 81:907–913, 2010.

[47] A. Anderson, Y. Menguc, R. J. Wood, and D. Newman. Development of the polipo pressure sensing system for dynamic space-suited motion. *IEEE Sensors Journal*, 15, 2015.

[48] Young-Young Shen, Abhishektha Boppana, Katya Arquilla, and Allison Anderson. Wearable sensor suit system for quantifying human-spacesuit interactions. In *IEEE Aerospace Conference*, pages 1–13, Big Sky, MT, 2018.

[49] I. Dianat, C. M. Haslegrave, and A. W. Stedmon. Methodology for evaluating gloves in relation to the effects on hand performance capabilities: A literature review. *Ergonomics*, 55(17):1429–51, August 2012.

[50] O. Bekdash, J. Norcross, and S. Mcfarland. Monitoring Human Performance during Suited Operations: A Technology Feasibility Study Using EMU Gloves. In *International Conference on Environmental Systems*. Bellevue, Washington, 2015.

[51] E. A. Benson, S. A. England, M. Mesloh, S. Thompson, and S. Rajulu. Use of traditional and novel methods to evaluate the influence of an EVA glove on hand performance. In *International Conference on Environmental Systems*, Barcelona, Spain, 2010.

[52] C. E. Newton, J. P. Kring, T. Southern, and N. Moiseev. Developing a standardized testing protocol for space suit gloves. In *International Conference on Environmental Systems*, Bellevue, Washington, 2015.

[53] R. Heus, H. A. M. Daanen, and G. Havenith. Physiological criteria for functioning hands in the cold. *Applied Ergonomics*, 26, 1995.

[54] J. M. O'Hara, M. Briganti, J. Cleland, and D. Winfield. Extravehicular activities limitations study. Technical report, NASA, 1988.

[55] D. C. Buhman, J. A. Cherry, L. Bronkema-Orr, and R. Bishu. Effects of glove, orientation, pressure, load, and handle on submaximal grasp force. *International Journal of Industrial Ergonomics*, 25, 2000.

[56] S. Thompson, K. Holden, S. England, and E. Benson. The effect of pressurized space gloves on operability of cursor controls, mobility, and strength. *Proceedings of the Human Factors and Ergonomics Society*, 55(1):1587–1591, 2011.

[57] D. L. Dempsey and I. Barshi. Applying research-based training principles: Towards crew-centered, mission-oriented space flight training. In *International Conference on Applied Human Factors and Ergonomics*, Orlando, FL, 2019.

[58] S. Nassiff and N. Armstrong. Apollo flightcrew training in lunar landing simulators. In *Simulation and Support Conference*, Los Angeles, CA, 1968.

[59] T. Southern, N. Moiseev, and M. Iturmendi. Operational testing of a novel IVA space suit. In *International Conference on Environmental Systems*, Bellevue, Washington, 2015.

[60] T. Southern, N. Moiseev, A. Persad, and M. Anand. Pressurized testing of a commercial intra-vehicular activity space suit. In *International Conference on Environmental Systems.* Albuquerque, NM, 2018.

[61] B. Kantowitz. Defining and measuring pilot mental workload. Technical report, Mental State Estimation, NASA Langley Research Center, Hampton, VA, 1987.

[62] R. Taylor and J. Berman. Ergonomic aspects of aircraft keyboard design: the effects of gloves and sensory feedback on keying performance. *Ergonomics*, 25(11):1109–1123, 1982.

[63] P. Llanos, V. Kitmanyen, E. Seedhouse, and R. Kobrick. Suitability Testing for PoSSUM Scientist-Astronaut Candidates using the Suborbital Space Flight Simulator with an IVA Spacesuit. In *International Conference on Environmental Systems*, Charleston, South Carolina, 2017.

[64] M. McQuerry. Effect of structural turnout suit fit on female versus male firefighter range of motion. *Applied Ergonomics*, 82, 2020.

[65] H. Park, R. S. Kakar, J. Pei, J. M. Tome, and J. Stull. Impact of Size of Fire boot and SCBA Cylinder on Firefighters' Mobility. *Clothing and Textiles Research Journal*, 37(2), 2019.

[66] NASA. Man-systems integration standards - 14, eva. Online - URL: https/msis.jsc.nasa.gov/sections/section14.html_14.6_EVA_TOOLS.

[67] M. Tomczak and Ewa T. The need to report effect size estimates revisited. *An overview of some recommended measures of effect size, Trends in Sports Sciences*, 1(21):19–25, 2014.

[68] B. Yoav and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, 57:289–300, 1995.

[69] NASA. Columbia crew survival investigation report. Technical Report NASA/SP-2008-565, 2008. pp. 3-30.

[70] R. S. Johansson and J. R. Flanagan. Coding and use of tactile signals from the fingertips in object manipulation tasks. 10:345–359, 2009.

[71] NASA. *Human Integration Design Handbook*. 2014.

[72] Y. Itaguchi and K. Fukuzawa. Effects of arm stiffness and muscle effort on position reproduction error in the horizontal plane. *Perceptual and Motor Skills*, 114(3):757–773, 2012.