Development Of An Objective Space Suit Mobility Performance Metric Using Metabolic Cost And Functional Tasks

Shane M. McFarland[1]
MEI Technologies/NASA-JSC, Houston, TX 77058

and

Jason Norcross[2]
Wyle Laboratories/NASA-JSC, Houston, TX 77058

Existing methods for evaluating EVA suit performance and mobility have historically concentrated on isolated joint range of motion and torque. However, these techniques do little to evaluate how well a suited crewmember can actually perform during an EVA. An alternative method of characterizing suited mobility through measurement of metabolic cost to the wearer has been evaluated at Johnson Space Center over the past several years. The most recent study involved six test subjects completing multiple trials of various functional tasks in each of three different space suits; the results indicated it was often possible to discern between different suit designs on the basis of metabolic cost alone. However, other variables may have an effect on real-world suited performance; namely, completion time of the task, the gravity field in which the task is completed, etc. While previous results have analyzed completion time, metabolic cost, and metabolic cost normalized to system mass individually, it is desirable to develop a single metric comprising these (and potentially other) performance metrics. This paper outlines the background upon which this single-score metric is determined to be feasible, and initial efforts to develop such a metric. Forward work includes variable coefficient determination and verification of the metric through repeated testing.

Nomenclature

AMA	=	[Reddit] "Ask Me Anything"
JSC	=	Johnson Space Center
NASA	=	National Aeronautics and Space Administration
PAO	=	Public Affairs Office
RFP	=	Request for Proposal
SSA	=	Space Suit Assembly
TRL	=	Technology Readiness Level

I. Background

Space suit mobility has historically been defined and characterized by a combination of range of motion and joint torque of the individual anatomical joints when performing isolated motions meant to drive that joint only in a given orthogonal plane[1,2,3]. While this has been the standard approach for several decades, there are numerous shortcomings that suit designers and engineers would like to see rectified. First, the lack of a standardized method for collecting both range of motion and joint torque of an individual joint by itself translates to many different test setups, procedures and methods of data analysis[1,3]. Second, all of these previously-used methods for data collection lack some degree of repeatability, even within the same test setup and the same conductor[3]. For example, the standard fish-scale method has been used for numerous range-of-motion and joint torque tests at Johnson Space Center. The

---

[1] Senior Space Suit Engineer, Space Suit and Crew Survival Systems Branch, NASA Johnson Space Center
[2] Lead Scientist, EVA Physiology Laboratory, Wyle Science, Technology and Engineering Group

results show high variability even within one test point of multiple joint articulations – much less the variability seen using a different fish scale or different test conductor[4]. In addition, attempts at higher fidelity data collection techniques, such as motion capture, require high overhead and cost with minimal improvement[1,5]. Lastly, and perhaps most importantly, isolated motions in standard anatomical planes are not representative of real-world tasks that a crewmember would be performing during an EVA (extra-vehicular activity), be it microgravity or surface exploration based[6].

To address these shortcomings, options are being explored within the Space Suit and Crew Survival Systems Branch to ascertain the feasibility of an alternative approach to defining mobility – one that is more repeatable, lower overhead, and more tied to functional EVA tasks. A feasibility study was conducted in 2013 which documented the first attempt at such an alternative option – one that looks at the metabolic energy-cost of a space suit. In other words, can we objectively quantify the mobility of a space suit by evaluating the metabolic cost of that suit to the wearer while performing a battery of functional EVA tasks? This attempts to address the issue of space suit mobility not at the individual joint level, but of the overall suit system while performing representative EVA tasks.

The 2013 feasibility assessment used three experienced suited subjects, performing eight functional tasks for two minutes in each of two suits – the Mark III and Z-1 planetary prototypes. Data was also collected of these subjects performing the same tasks unsuited using two different pacing techniques. CO_2 output was used as a common metric across all tests as a basis for comparison of metabolic load. Using flow rate and assumed metabolic characteristics, comparisons were made in liters of O_2 consumption, as well as liters of O_2 per kilogram system mass per task repetition as normalization schemes.

The results of this feasibility assessment demonstrated strong promise for the approach and a possible slight advantage of one suit over another in most tasks. Note that at the time, a standard metabolic cost metric was not determined, so two options were presented (VO_2, both in mL and mL normalized to system mass and repetition). Also, a significance threshold between suits was not determinable, so an arbitrary value of 10% was considered to be "significant" in this feasibility assessment.

The final report, as documented in a corresponding publication[7], also highlighted many improvements that could be made on the approach; namely:
- Elimination/modification of some tasks
- Standardize repetitions completed for each task instead of time
- Set the number of repetitions such that subjects take 4-5 minutes to complete the task
- Increase the subject count to at least five to improve statistical significance
- Collect metabolic performance data on each subject to improve accuracy of results
- Collect suited O_2 consumption directly, if possible, to improve accuracy of results
- Prioritize well-fitting subjects over experienced subjects in selection
- Consider the possibility of including a third suit of lower mobility but also lower mass, to test the extensibility of the metabolic cost technique

These and other improvements to the approach were included when the test team submitted a proposal to the NASA HRP Solicitation Omnibus Announcement, NNJ13ZSA002N-OMNIBUS for FY 2014. The proposal was awarded with a period of performance of one year starting on October 1, 2014.

The goals of this testing campaign were as follows:
- Assess the mobility of 3 different space-suit assemblies as characterized by metabolic cost when performing functional tasks
- Further characterize the variability associated with a single subject completing functional tasks
- Assess the consistency of performance trends across a pool of six subjects performing the same tasks in the same suits
- Evaluate the technique's sensitivity to suit assemblies with significantly differing masses

II. Testing Results Summary

The testing conducted in 2015 was comprised of six subjects (mixed suited experience), performing various tasks in three different suit assemblies – the Mark III, the Rear Entry I-Suit (REI) and the Demonstrator suit. Five tasks were completed: walking, side stepping, stair climbing, and upper body and full body object relocation tasks. These tasks are shown below in Figure 1.

International Conference on Environmental Systems
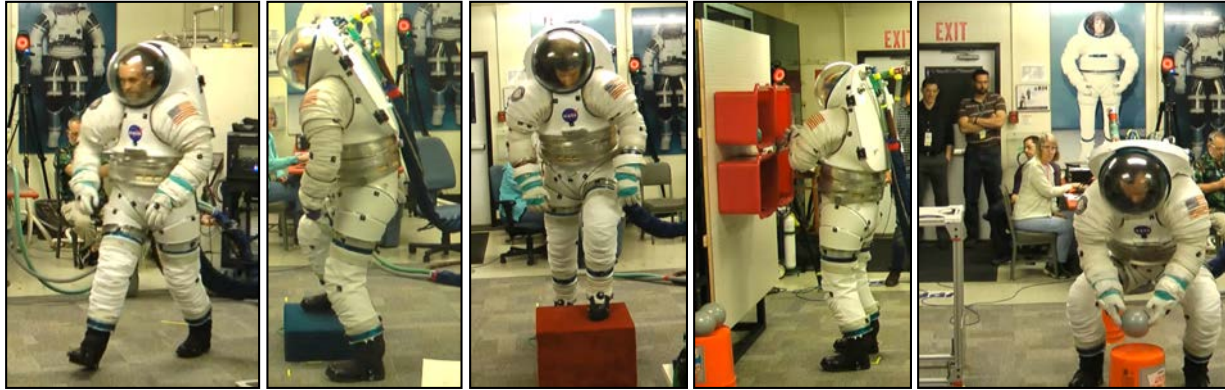
Figure 1. Walking, Side Stepping, Stair Climbing, Upper and Full Body Relocation tasks, respectively

The detailed results are not within the scope of this paper but will be available in an upcoming publication, as well as the final report for the NNJ13ZSA002N-OMNIBUS titled "Metabolic Assessment of Suited Mobility using Functional Tasks". However, a brief summary follows herein to provide sufficient context.

The testing demonstrated that method of characterizing space suit mobility and performance through metabolic cost continues to show great promise. The viability of showing statistically relevant differences between similar suit architectures performing functional tasks was demonstrated. When defining metabolic cost as BTU/rep for the second half of the trial with resting metabolic rate removed, the REI required slightly less metabolic cost to the user than the Mark III on tasks that required significant motion in the vertical plane (stair climbing, object relocation, side stepping). In addition, the Demonstrator suit, a low-mass, lower mobility design, was shown to require significantly higher metabolic cost for all tasks and subjects. These results were verified to be statistically significant in a mixed-effects regression analysis. A summary of all results is shown below in Figure 2.

Additionally, Rate of Perceived Exertion (RPE), recovery period data, resting metabolic rate, unsuited data, task completion time analysis, and mass normalization schemes were investigated. In nearly all cases, weaknesses in the data or method can be corrected in analysis or subsequent evaluations. Theorized "gaming techniques" such as low mobility/low mass suit and as-fast-as-possible task completion were evaluated and shown to be invalid or correctable in analysis.

3

International Conference on Environmental Systems
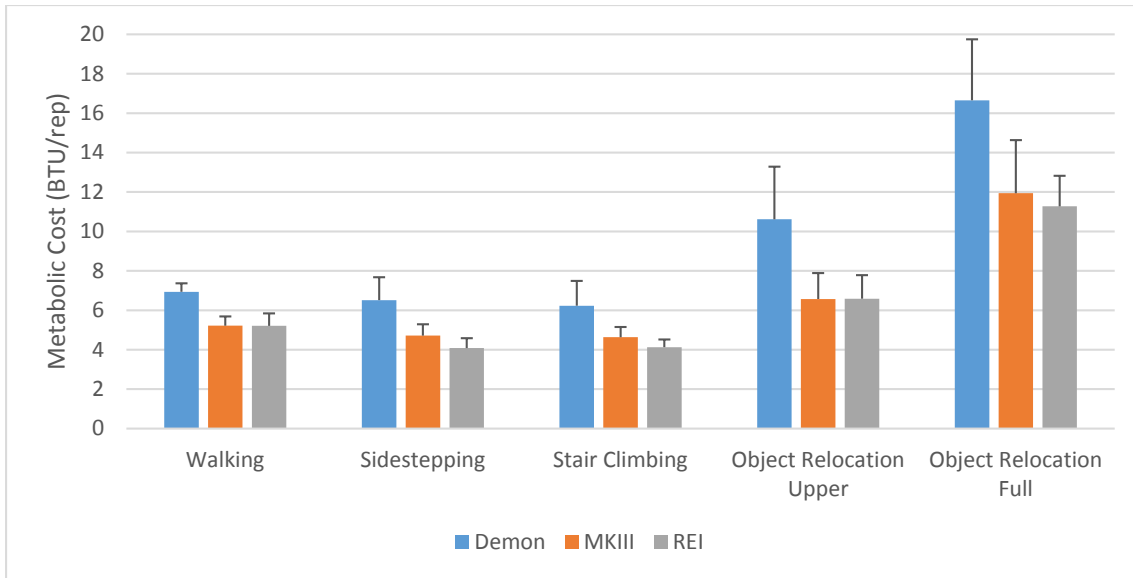
Figure 2. Summary of results from Energy Based Mobility testing. *Average of six completed subjects with three trials in each suit*

Lastly, subject variability, suit variability and learning effects were evaluated (note that three suits and six subjects allowed for a perfectly permutated test order). Outside of a few cases where an individual subject changed strategy between trials, there was little evidence of any learning effect indicating improved performance with repeated trials. Across the entire test series, the average improvement from the 1st to 2nd trial was only 1.6%, with the Demonstrator suit showing the largest improvement of 2.5%. However, these values were within the inherent variability of performing these tasks, so while there may be a very small learning effect on the aggregate, it is negligible and oftentimes, subjects performed best on their first run. Also note that when comparing the 2nd to 3rd trials the improvement was less than 0.1%. Additionally, the most experienced suited subject in the test series demonstrated a net negative improvement from the first trial across all tasks/suits. Therefore, it appears that for the functional tasks that were chosen, a fit-check and short familiarization session of the tasks were sufficient to get through any inherent learning curve that may exist.

A representative chart in Figure 3 demonstrates the inherent variability of repeated measures, and how the Demonstrator suit, while requiring additional metabolic cost, also exhibited much higher variability.
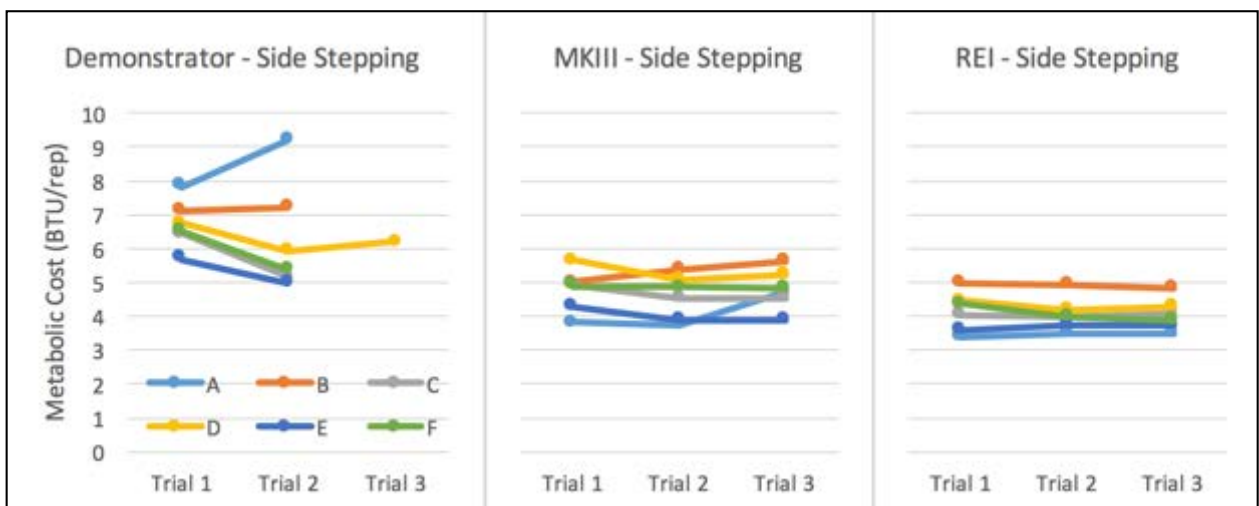
Figure 3. Trial-to-trial variability during side stepping task. *Note higher average cost and variability for Demonstrator suit.*

4

International Conference on Environmental Systems

III. Evaluation of Additional Variables

For the testing summarized above and corresponding report, the primary metric to define suit cost to the occupant was BTU/rep for the 2nd half of the trial, with the resting metabolic rate removed. While this metric worked well for the purpose, the significant amount of additional data that was collected and the detailed analysis thereof, provided the means to evaluate the feasibility of modifications to that metric, or the combination of said metric with additional metrics to develop a more comprehensive method for objectively evaluating suited performance using metabolic cost. This section serves to document these individual components on a piecewise basis for potential inclusion into a "test for score" metric.

A. Recovery Period

Metabolic cost is defined as the energy expenditure to complete a given task. We assumed that the overall metabolic cost to complete a given number of repetitions of a specific functional task would be the same even if the time to completion varied. To account for possible issues related to time to completion, we measured baseline energy expenditure for each subject after donning the suit and resting for 10 minutes or more until 5 minutes of steady data was collected. This baseline data was then removed from the metabolic cost calculations so that a subject who took longer would not be penalized for spending a greater portion of time just staying alive. In addition, subjects immediately returned to the donning stand or other resting position and two minutes of recovery data was collected to include as part of the metabolic cost accounting for some of the metabolic energy expenditure not readily measured by indirect calorimetry, which is typically defined as the O_2 deficit. An example of a typical metabolic profile for a task is shown in Figure 4.
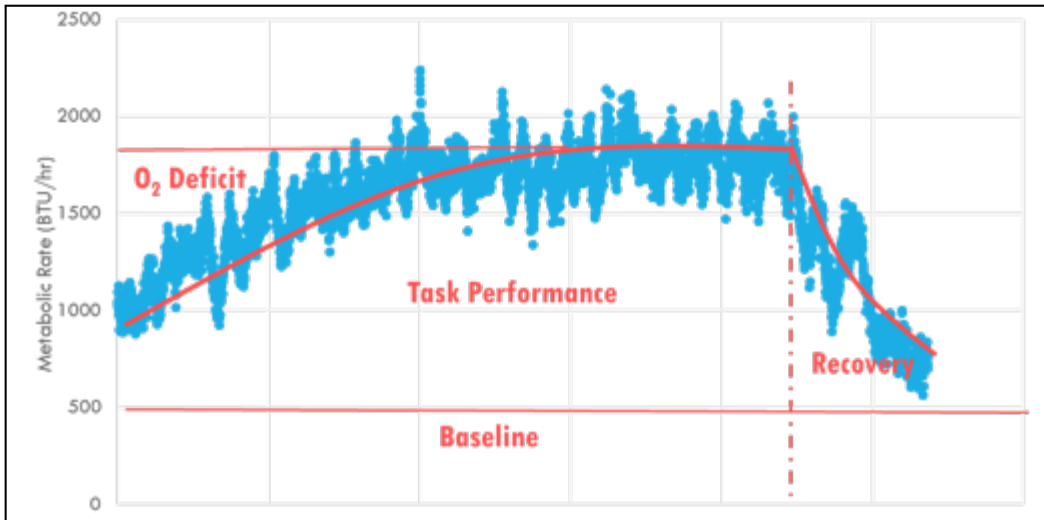

**Figure 4. Typical example of a metabolic profile.**

Due to data recording errors in the original test series, recovery data was only collected for 4 of 6 subjects, so although we expected to use the whole metabolic cost of the task including 2 minutes of recovery, we had to look for an alternate strategy to include all 6 subjects. In addition, it was often very difficult for subjects to complete the required number of repetitions in the Demonstrator suit, so a lower number of repetitions was permitted. Therefore, we also had to determine a strategy for accounting for the disparity in reps per trial. After detailed investigation, it was determined that using the BTU/rep for the 2nd half of the trial (minus rest) was the best approach, which had a high correlation ($r^2=0.8793$) with the ideal BTU metric.

Further review of the data was performed as it relates to the recovery period as shown in Figure 4. Specifically, in the context of Time to Completion [TTC], the assumption was that the metabolic cost associated with a task should be the same irrespective of how long it takes. To confirm this assumption, the TTC was plotted against the BTU/rep (2nd half, minus rest). As it turns out, for most tasks across all suits, the metabolic cost as defined was correlated with higher completion times. This is not ideal, as not only should metabolic cost be decoupled from time, it also allows a possible avenue for "gaming" the metabolic cost metric by completing the task as quickly as possible. Upon further investigation, the relationship between completion time and metabolic cost essentially disappeared when two minutes

International Conference on Environmental Systems

of post-task recovery metabolic cost was included in the analysis. A sample comparison is provided below in Figures 5 and 6.
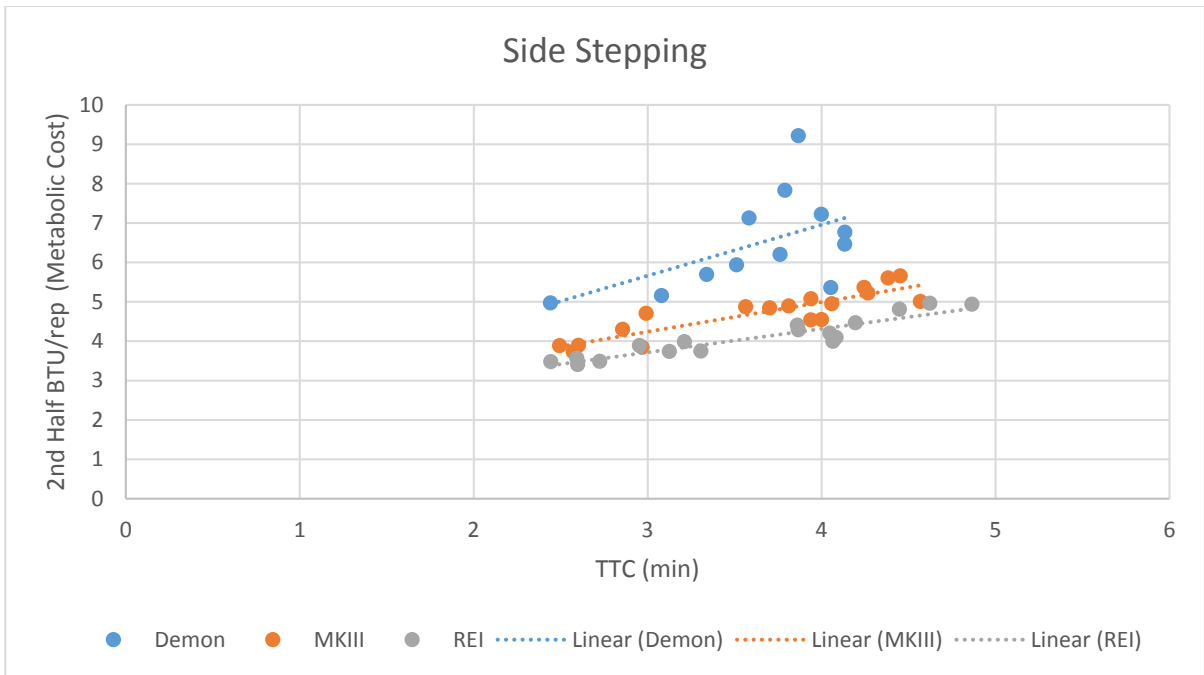
Figure 5. Completion time (TTC) vs. 2nd half BTU/rep. *Note slope of best fit, indicating relationship*
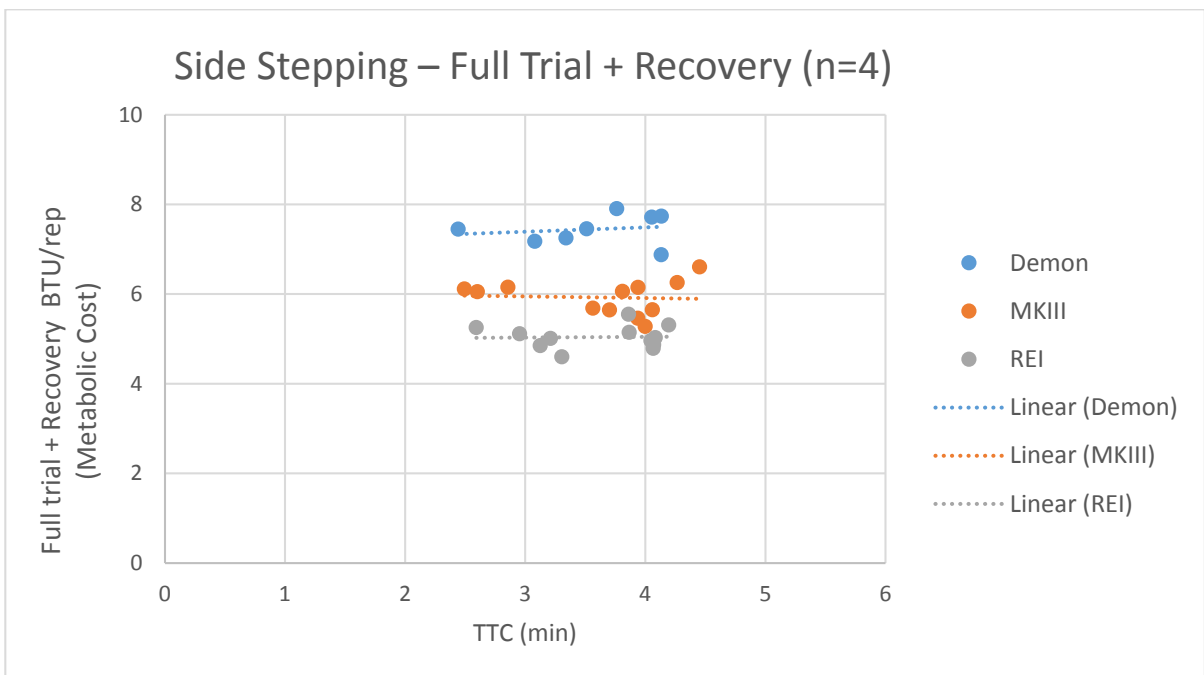
Figure 6. Completion time (TTC) vs BTU/rep including recovery period. *Note slope of best fit near zero.*

As a result of this analysis, it was determined that in future testing, the metabolic component of BTU for the full trial, plus two minutes recovery (normalized to repetition if necessary) should be used.

International Conference on Environmental Systems

B. Time to Completion

As discussed in the preceding section, it is possible to decouple the completion time of a task from the metabolic cost associated with performing it, which is preferable because it prevents the possibility of gaining an artifically preferable score by performing the task as quickly as possible. However, this raises the question of how completion time relates to performance. Everything else being equal, it is preferable to be able to complete a task quicker; note only does that speak toward efficiency of motion, but ultimately provides logistical benefits associated with more completed tasks in a given EVA or fewer EVA hours.

Therefore, going forward, when evaluating a possible comprehensive, objective suit performance metric, the efficiencies related to reduced task completion time should be considered.

C. Self-Reported Rate of Perceived Exertion

Borg's Rate of Perceived Exertion (RPE) of the subject was queried immediately following each trial, and was reported on a 6-20 scale as shown in Figure 7. The results are plotted in Figure 8 against the BTU/hr of the second half of each trial. There was no difference in how subjects rated RPE across different space suits indicating that subjects rated RPE consistently based on metabolic effort and there were no major differences based on suited configuration.

Howeve, when the same data is viewed for each individual subject, it is seen that RPE is a poor predictor of absolute metabolic rate, and is highly dependent on subject (R^2 ranging from 0.009 to 0.62). This underscores the importance of development of an objective performance metric decoupled from subjective feedback; however, the inclusion of RPE in a test-for-score type metric in conjunction with other factors such as BTU/rep, BTU•rep^{-1}•kg^{-1}, and time to completion could be something evaluated at a later date once the more objecgive components have been determined and weighted appropriately.

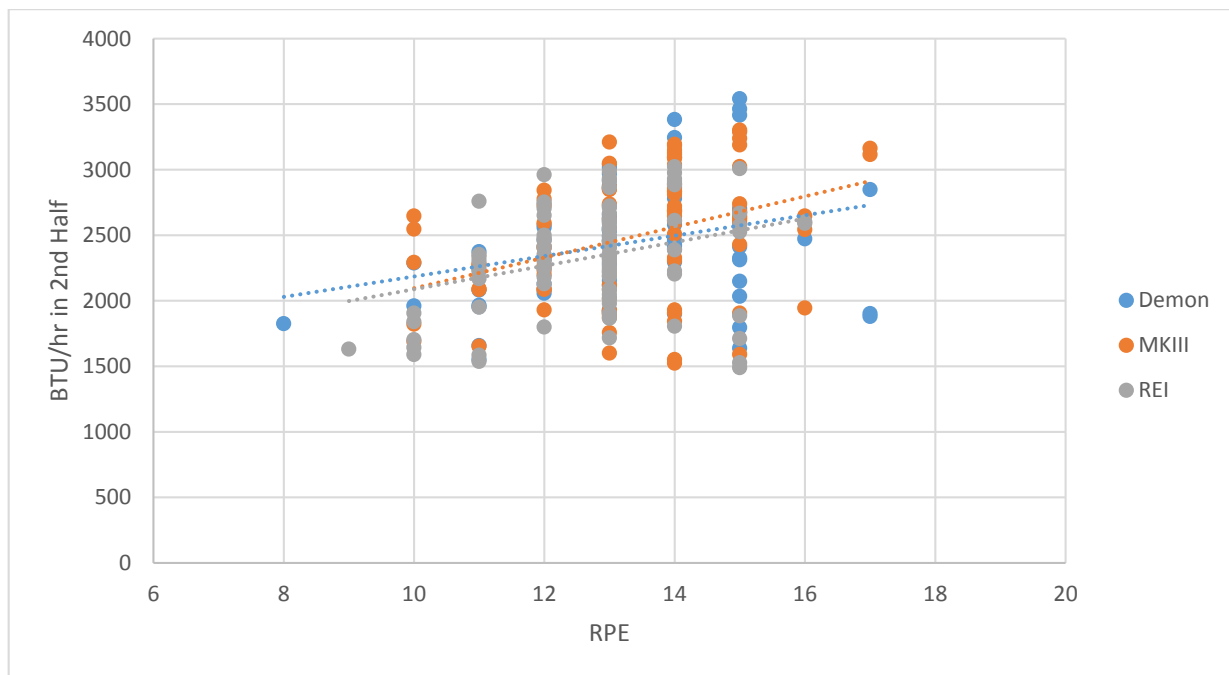| Borg Rating of Perceived Exertion	
6	No exertion at all
7	
8	Extremely light
9	Very light
10	
11	Light
12	
13	Somewhat hard
14	
15	Hard (heavy)
16	
17	Very hard
18	
19	Extremely hard
20	Maximal exertion

Figure 7. Borg RPE Scale

Figure 8. Regression of average metabolic rate in the 2nd half of the tasks with RPE across different space suits

D. Normalization to System Mass

While the primary metabolic cost metric was selected to be BTU/rep for the 2nd half of the trial, another way of comparing the suits can be to normalize the results to the mass of the suit and subject, thereby potentially "correcting" for the differences in the mass of the three suits, which is substantial; the increase in mass from the Demonstrator suit to REI is approximately 35 pounds, and the difference between the REI and Mark III is nearly an additional 50 pounds. In addition, the difference between subjects was up to 15 pounds in the greatest case. The normalization scheme is simply to divide the BTUrep for 2nd half metric by the mass of the suit and subject, resulting in a metric of $BTU\bullet rep^{-1}\bullet kg^{-1}$. For reference, this data (Figure 9) is graphed above the BTU/rep (Figure 10) metric for direct comparison.
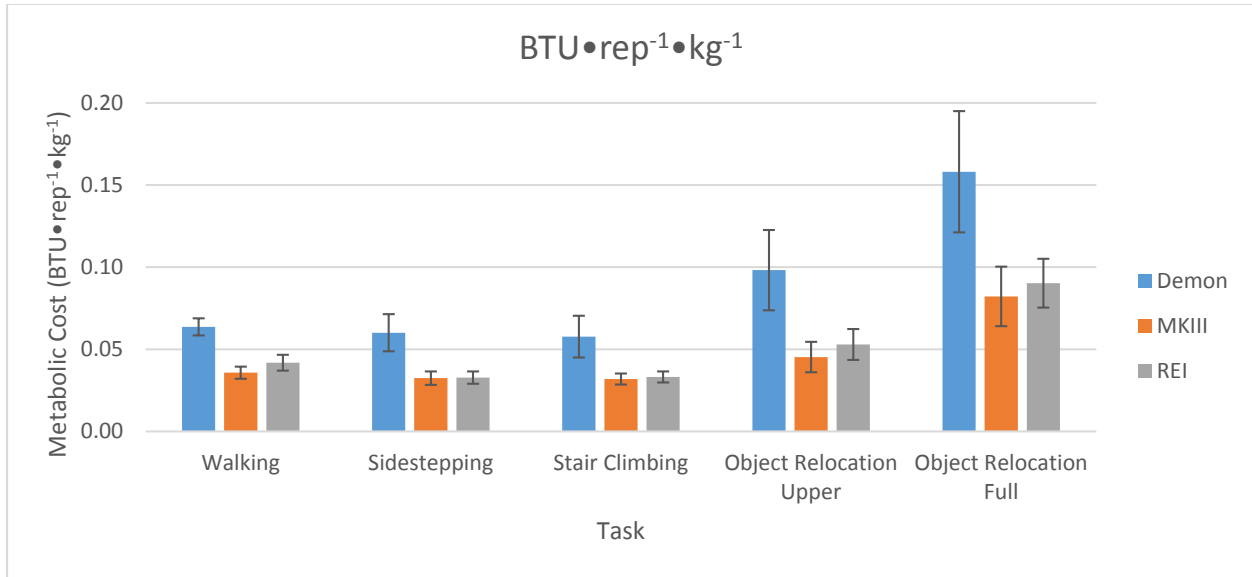
Figure 9. Mass-normalized metabolic cost (mean ± SD) comparison of difference space suits across functional tasks
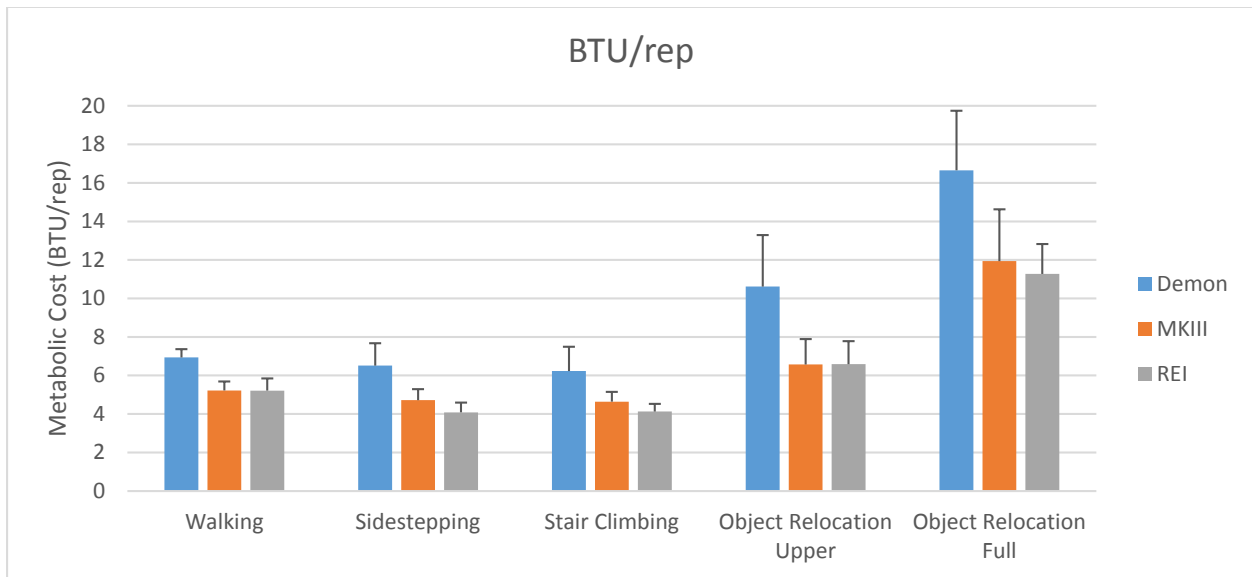
Figure 10. Metabolic cost (mean ± SD) comparison of difference space suits across functional tasks

When inspecting these two figures, it is clear that normalizing to system mass has an appreciable change to the comparison between the Mark III and REI suits. In the primary analysis using BTU/rep, the REI had the same or better metabolic cost than the Mark III, depending on task; when using $BTU\bullet rep^{-1}\bullet kg^{-1}$, this relationship reverses: now the Mark III has a same or better metabolic cost than the REI. This is not completely unexpected, as the Mark III

International Conference on Environmental Systems

weighs more than the REI. In a sense, the purpose of looking at the data this way is to account for the fact that this testing is completed in 1G. Given the fact that these suits are obviously designed to operate somewhere between micro-G and 3/8 G on Mars, one might argue that heavier suits are being "penalized" by having to be tested in a gravity field much higher than the design case. Normalizing to system mass essentially tries to characterize the additional metabolic cost associated with a heavier suit in 1G.

However, this metric is primarily being developed for evaluation of surface EVA suits – either 1/6 or 3/8 G, primarily. Therefore, it is not reasonable to try to compensate for the additional mass, only to reduce it commensurate with a specific surface gravity field. One could do this by by calculating weighted averages of the "mass-normalized" and "not-mass-normalized" metrics, if desired.

That being said, the most compelling argument one can make against using this mass normalization scheme is a simple hypothetical one. If two suits of different masses are able to achieve the exact same performance at the same metabolic energy in any given gravity field, should the heavier suit be rewarded simply for being heavier? Obviously not; if anything, the lighter suit should be rewarded for being lighter. The ultimate weight of a suit is driven by a host of design decisions and requirements; however, for this metabolic cost analysis to evaluate suited performance specifically, it should not penalize a suit that is able to achieve the same performance at a lower mass.

It is for this reason that this metric was not used as a primary means of analysis in previous testing.

However, because this testing is completed in 1G, a lighter suit is rewarded due to the fact that the subject has a lower metabolic response to carrying the suit weight compared to the heavier suit. This is good to an extent, because although in reduced gravity this benefit will be significantly less, there are other benefits to reduced mass, such as reduced system cost. Whether it is being rewarded too much or not enough is certainly up for debate, and again, higher mass is often a result of many other design drivers, such as durability, sizing, etc. There is significant open work in trying to determine how to equalize these two unrelated things; this would potentially create a hybrid of these two metrics: for example, reduce the lighter suit's "reward" and heavier suit's "penalty" by 20%. This could be an attempt at finding the balance between over-penalizing a heavier suit by testing in 1G (BTU/rep), and not accounting for the value of reduced suit weight through reduced metabolic impact and system cost (BTU•rep^{-1}•kg^{-1}).

This is not within the scope of this analysis, but it shows potential for being a superior metric of metabolic cost if one was able to define the value of both system mass/cost and suit metabolic cost in the same terms. In the final results of this testing, the BTU/rep metric rewards lighter suits, and that is better than rewarding heavier ones. By adding the additional layer of a weighted average to account for a specific gravity field, that would likely be the best technical approach until further analysis facilitates rewarding reduced system mass specifically.

As a side note, it is important to mention that the Demonstrator suit was included in testing for two primary purposes: One, to evaluate the feasibility that a lower mass, lower mobility suit could score artificially well in a metabolic metric simply because it is so light; two, to evaluate the metabolic cost versus more representative surface EVA prototypes. The latter is more of an academic exercise, as the Demonstrator is very similar to the Apollo A7LB suit architecture and having a comparison would be insightful. For the former, it is shown here that regardless of if you normalize to system mass or not, the Demonstrator suit does not compare favorably against the REI and Mark III; therefore, it does not seem possible that a lighter suit that is visibly weaker in pressurized mobility performance can "game" this metric simply by having the subject carry less suit mass during the task. To go a step further, when viewing technique for many of the tasks (side step, stairs, full body relocation) in the Demonstrator suit, it seemed quite clear that subjects relied on full gravity to facilitate completion of the task. For example, when climbing a single step, instead of simply raising their knee to bring their foot over the step (as would be done unsuited or in the other suits), they would first use the weight of their body to partially squat, and then spring their body and leg up using stored energy from the joints in the form of torque, while creating vertical inertia. For many subjects, this technique was vital to being able to perform the task at all. Therefore, on the Moon or Mars, one would expect the Demonstrator suit to perform even less favorably, as the reduced gravity field would preclude subjects from employing the same techniques.

International Conference on Environmental Systems

IV. Development of Preliminary Test for Score Metric

As mentioned throughout this paper, the next step in this work is to develop a test-for-score metric, where a single numeric score is calculated as a function of multiple metrics. This score could be tailored to a given gravity field and provide weighting factors to specifically reward shorter completion times and/or lower mass designs, if desired. By incorporating the metrics discussed above to be feasible as a component, determination of a preliminary formula follows. When attempting to account for metabolic cost (including recovery), completion time, the gravity of the design case, and value of system benefits to reduced mass:

$$S = \frac{1}{G \cdot (1 - V_m) \cdot C_s + (1 - G) \cdot V_m \cdot C_m + V_t \cdot t}$$

where:

- S is score
- G is gravity fraction of Earth
- V_m is the value of system benefits of a lighter suit between 0 and 1
- C_s is the standard metabolic cost in BTU/rep, suit-normalized
- C_m is the mass-normalized metabolic cost in BTU•rep^{-1}•kg^{-1}, suit-normalized
- V_t is the value of completion time between 0 (none) and 1 (same weight as metabolic cost)
- t is time to complete the task, suit-normalized

This illustrative example is by no means complete, but it does weight-average the metabolic metric commensurate with a specific gravity, allows for weighting the system benefits of a lighter suit, and allows for placing specific value on completing tasks quicker.

More work and testing would need to be done to determine the best possible candidate metric functions and constant values, and thoroughly vet them through repeated testing. Specifically, several defeciences with this draft are notable and could benefit from targeted investigation:

- The "system benefits to reduced mass" value V_m could work, as it provides a rough approximation of the value by assigning a value between 0 and 1. Currently, with the lack of further information, a value of 0 should be selected. Detailed system analysis, either generic or, ideally, specific to an operational concepts or mission, may be able to determine a basic scalar value to use for V_m. However, another likely scenario is that a detailed system analysis would drive a slightly more complicated means of accounting for reduced system mass value, and then the formula above would need to be modified appropriately to account for this additional complexity.
- Note that this metric currently requires normalization of the completion time and metabolic cost components which facilitates easier balance between them; however, appropriate selection of V_t and V_m (and other weighting factors, if added) could eliminate the need for this normalization scheme. That being said, normalization to a standard might be desirable at some point in the future, although it would require assurance that the suit selected as the standard provides sufficient fleet sizing implementation to ensure optimal fit before using it as a standard going forward.
- Furthermore on the issue of normalization, once the various weighting factors are determined, and depending on their variability, it may be possible or desireable to modify the overall equation such that S itself is also normalized to the selected standard suit (if/when one is determined and used as such).
- The consideration for additional metric components or variables (either subjective or objective) is entirely possible, and any additional thoughts or analysis on this topic are certainly welcome. Possibilities that have not yet been fully investigated include Electromyography (EMG) to monitor muscle activity during a task, kinematics of motion, subjective scoring, and additional advanced physiological variables that would indicate general fatigue or workload.

- Future testing in reduced gravity is further needed to verify that measurement in 1G translates into performance in reduced gravity. For instance, to complete tasks in the Demonstrator suit, subjects often had to use gravity to their advantage to squat down rapidly to pick up the balls during the full body object relocation. In reduced gravity, this technique may not work, but then again, another technique only possible in reduced gravity could possibly be used.

V. Conclusion

It may be a very long time, if ever, before we are able to eschew the subjective feedback that dominates the characterization of space suit performance for a more objective metric. Previous attempts at objectively measuring suit mobility or performance have yielded mixed results; this work is a look at a potential alternative that not only uses metabolic cost as the basis, but also provides the means of incorporating multiple objective (and possibly subjective) metrics into a single numeric score representing performance. While this work has been maturing for several years, and shows promise, it is still lacking in maturity and requires additional investigation to further develop.

This paper should not only serve as documentation of the work that has been done to date on the development of this metric, but also highlight the currently known deficiencies that warrant further investigation.

References

[1]Aitchison, L., "A Comparison of Methods for Assessing Space Suit Joint Ranges of Motion," AIAA-2012-3534, AIAA International Conference on Environmental Systems, 2012.

[2]Reinhardt, A., "Results and Applications of a Space Suit Range-of-Motion Study," SAE Technical Paper 891592, 1989, doi:10.4271/891592.

[3]Valish, D., Eversley, K. "A Comparison of Methods for Assessing Space Suit Joint Ranges of Motion," AIAA-2012-3532, AIAA International Conference on Environmental Systems, 2012.

[4]Matty, J. "Results and Analysis from Space Suit Joint Torque Testing," AIAA-2010-6211, AIAA International Conference on Environmental Systems, 2010.

[5]England, S., Benson, E., and Rajulu, S., "Functional Mobility Testing: Quantification of Functionally Utilized Mobility among Unsuited and Suited Subjects," NASA/TP-2010-216122, 2010.

[6]Aitchison, L., "Philosophies Applied in the Selection of Spacesuit Joint Range-of-Motion Requirements," SAE Int. J. Aerosp. 4(1):494-502, 2011.

[7]McFarland, S and Norcross, J. "A Novel Method for Characterizing Spacesuit Mobility Through Metabolic Cost," 2014-ICES-007, International Conference on Environmental Systems, 2014.

International Conference on Environmental Systems