ARL-TR-8444 ● AUG 2018

US Army Research Laboratory

Methodologies for Evaluating the Effects of Physical Augmentation Technologies on Soldier Performance

by Harrison P Crowell, Gregory B Kanagaki, Meghan P O'Donovan, Courtney A Haynes, Joon-Hyuk Park, Jennifer M Neugebauer, Edward R Hennessy, Angela C Boynton, K Blake Mitchell, Andrew J Tweedell, and Henry J Girolamo

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

ARL-TR-8444 ● AUG 2018

Natick Soldier RD&E Center

US Army Research Laboratory

Methodologies for Evaluating the Effects of Physical Augmentation Technologies on Soldier Performance

by Harrison P Crowell, Courtney A Haynes, Jennifer M Neugebauer, Angela C Boynton, and Andrew J Tweedell
Human Research and Engineering Directorate, ARL

Gregory B Kanagaki, Meghan P O'Donovan, Edward R Hennessy, K Blake Mitchell, and Henry J Girolamo
Natick Soldier Research, Development & Engineering Center

Joon-Hyuk Park
National Research Council, Washington, DC

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
August 2018	Technical Report	April 2015 – September 2017

4. TITLE AND SUBTITLE
Methodologies for Evaluating the Effects of Physical Augmentation Technologies on Soldier Performance

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Harrison P Crowell, Gregory B Kanagaki, Meghan P O'Donovan, Courtney A Haynes, Joon-Hyuk Park, Jennifer M Neugebauer, Edward R Hennessy, Angela C Boynton, K Blake Mitchell, Andrew J Tweedell, and Henry J Girolamo

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
US Army Research Laboratory
ATTN: RDRL-HRF-B
Aberdeen Proving Ground, MD 21005-5425
US Army NSRDEC; ATTN: RDNS-SEO, Natick, MA 01760

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-8444

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The loads that US Army Soldiers carry can decrease their effectiveness by reducing their mobility and agility and increasing their fatigue and risk of injury. A potential solution to these problems is to develop and implement physical augmentation systems such as exoskeletons and exosuits to assist with load carriage and other Soldier tasks. An integral part of developing physical augmentation systems involves evaluating their performance at various stages of the development process. The US Army Natick Soldier Research, Development & Engineering Center (NSRDEC) and the US Army Research Laboratory (ARL) have experience evaluating more than 20 physical augmentation systems, from prototypes to commercially available systems. This report discusses the methods and metrics used by researchers at NSRDEC and ARL to evaluate physical augmentation systems designed for military applications. It covers terminology, considerations for protocol development, and performance metrics. The report is intended to be a guide for planning evaluations and measuring performance in a consistent manner. It is meant to promote the development of innovative, safe, and highly effective physical augmentation systems.

15. SUBJECT TERMS
human augmentation, physical augmentation, exoskeleton, exosuit, mobility, soldier performance

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Harrison P Crowell
Unclassified	Unclassified	Unclassified	UU	78	19b. TELEPHONE NUMBER (Include area code) (410) 278-5986

Contents

List of Tables v

Acknowledgments vi

1. Introduction 1

2. Purpose 2

3. Scope 3

4. Terminology 5

5. Considerations for Protocol Development 6

5.1 Introduction 6

5.2 Elements of Protocol Development 7

5.2.1 Research Questions and Hypotheses 7

5.2.2 Subject Population 8

5.2.3 Methods for Evaluation 9

5.2.4 Metrics 16

5.2.5 Resources, Training, and Scheduling 17

5.2.6 Safety and Risk Mitigation 20

5.3 Protocol Development Summary and Recommended Best Practices 22

6. Performance Metrics: Biomechanical, Physiological, Operational, and Human Factors Metrics 23

6.1 Biomechanical Metrics 24

6.1.1 Postural Stability Metrics 24

6.1.2 Spatiotemporal Gait Metrics 25

6.1.3 Kinematic Metrics 27

6.1.4 Kinetic Metrics 28

6.2 Physiological Metrics 30

6.2.1 Metabolic Metrics 30

6.2.2 Muscle Function Metrics – Electromyography (EMG) 32

6.2.3 Muscle Function Metrics – Muscle Strength and Endurance 33
6.2.4 Psychophysiological Metrics 34
6.3 Operational Metrics 35
6.3.1 Strength Metric 36
6.3.2 Movement Metrics 36
6.3.3 Marksmanship Metrics 40
6.4 Human Factors Metrics 41
6.4.1 Fitting and Sizing Metrics 41
6.4.2 Usability Metrics 42
6.4.3 Equipment Compatibility Metrics 44
6.4.4 Training Metrics 45

7. Conclusion 46

8. References 48

Appendix A. Rating of Perceived Exertion Scale 60

Appendix B. Rating of Pain, Soreness, and Discomfort Scale 62

Appendix C. Comfort Affected Labeled Magnitude Scale 65

List of Symbols, Abbreviations, and Acronyms 67

Distribution List 69

List of Tables

Table 1 Suggested evaluation environments, tasks, and metrics based on technical maturity..... 11

Acknowledgments

Gregory B Kanagaki and Harrison P Crowell contributed equally to this report.

The authors acknowledge the following people for their help with the technical review and editing of this report: Dr J Cortney Bradford, Ms Karen N Gregorczyk, Dr William G Billotte, Mr John J Christopher, and Ms Amber L Bennett.

In addition, the authors acknowledge Dr F Matthew Kramer, and Dr Michael E LaFiandra for their role in guiding the Technology Program Agreement that led to the production of this report.

1. Introduction

The US Army Soldier is physically overburdened, which increases fatigue, reduces movement and maneuver, increases the likelihood of acute and chronic injuries, and damages mission effectiveness and readiness (Knapik et al. 2004, 2007; Bachkosky et al. 2007). Wearable physical augmentations systems (e.g., exoskeletons and exosuits) are a promising approach to addressing these problems. The Army's Robotic and Autonomous Systems Strategy (US Army Training and Doctrine Command 2017) outlines the path forward for "ensuring overmatch against increasingly capable enemies" and further states, "To continue to lighten the Soldier load in the future, the Army invests in exoskeleton technology." This is the near-term (2017–2020) strategy (i.e., having exoskeletal systems bear some of the Soldiers' burden). The mid-term (2021–2030) and far-term (2031–2040) strategies involve introducing exoskeletons that would also carry armor and innovative firepower solutions. These will be further developed into systems that provide the user with a common operating picture, updated intelligence, and the ability to integrate direct and indirect fire weapon systems. The Army's long-term strategy for using exoskeletons implies that a number of different kinds of exoskeleton systems will be built. These systems will need to be evaluated during the research and development cycle. It is important that they be evaluated in a consistent manner so that the technical progress is properly assessed and documented. The aim of this report is to initiate the use of common methods and metrics for the evaluation of physical augmentation systems.

Numerous organizations are publicly pursuing research and development of exoskeletons for military, industrial, and medical purposes. This research and development ranges from fundamental science to enabling component technologies to full systems. There are multiple military applications for physical augmentation systems. These include, but are not limited to, the following:

- Mobility Augmentation – allow Soldiers to carry loads farther and faster, jump higher, and enhance agility.
- Manual Materiel Handling – allow Soldiers to lift and carry heavy objects quickly and safely.
- Tool Operation – allow Soldiers to hold and operate heavy tools overhead or away from their bodies for extended periods of time during assembly, maintenance, and repair tasks.

- Load Distribution – adjust the distribution of torso-borne loads (e.g., body armor, rucksack, and other equipment) onto different parts of the torso (e.g., from shoulder to hips and vice versa).
- Medical – automated casualty first-aid and evacuation.
- Capability Platform – support the weight of additional armor, cooling systems, visual augmentation systems, weapons, explosive ordnance disposal suits, and life support systems.
- Postural Support – allow Soldiers in facility construction and equipment maintenance and repair to work in crouched postures for extended periods of time.
- Shock and Vibration Damping – protect the lower extremity joints from injury.
- Training – reduce the risk of injury and aid in recovery from an injury.

It is unlikely that one type of physical augmentation system will meet this broad array of military applications and associated disparate task requirements. In effect, developers will almost certainly optimize their systems specific to a single application or a small set of applications. Different applications will require different metrics to measure performance of the user and the system. Thus, it is imperative that there be discussion, collaboration, and consensus within the Army to understand how to adequately assess the effects of augmentation systems on Soldier performance in varied environments and applications.

2. Purpose

This document is the culmination of the April 2015–September 2017 Technology Program Agreement (TPA: NA-HR-2015-04) between US Army Natick Soldier Research, Development & Engineering Center (NSRDEC) and the US Army Research Laboratory Human Research and Engineering Directorate (ARL-HRED). The focus of the TPA was to facilitate an understanding of how to evaluate the effects of physical augmentation systems on Soldier performance. The overall intent was to promote standard evaluation methods and metrics and to ensure data compatibility to foster ongoing and future collaboration. However, this report is also applicable to physical augmentation research and development being conducted by other military and non-military organizations.

This document is based on significant combined experience evaluating over 20 systems for a range of military populations and applications. It is intended to be a

guide for planning evaluations and measuring performance in a consistent manner. Three sections form the core of this report:

- Terminology – This section defines the technical terms used throughout this document while leveraging standard terminology where applicable.
- Considerations for Protocol Development – This section highlights points to be considered as the evaluation of physical augmentation systems is being planned. Key considerations include the overall goal of the evaluation, intended use of the system, and technical maturity of the system.
- Performance Metrics – This section summarizes biomechanical, physiological, operational, and human factors metrics. In addition, the lab and/or field capabilities necessary to assess and quantify those metrics are identified.

The terminology, protocol development, and metrics information provided here are recommendations only, not standards. They are primarily for an audience with a technical background and foundational understanding of the human performance sciences for military applications. Similarly, this content complements, and does not replace, existing standards that may be related to physical augmentation systems.

From a longer term perspective, this report broadly supports the generation of voluntary consensus standards by standards developing organizations (e.g., ASTM International) in conjunction with research efforts led by the National Institute of Standards and Technology (NIST). These collective efforts ultimately aim to develop common terminology, methods, and metrics with government, industry, academia, and international stakeholder buy-in.

3. Scope

This report specifically focuses on wearable, physical augmentation technologies designed to improve physical performance (such as strength, endurance, and agility) of users during various military activities. These augmentation systems may also be intended to maintain or improve user safety. Multiple system designs and configurations are possible, and therefore this report may apply to powered, unpowered, lower body, upper body, and full body systems. There are some forms of physical augmentation that are not considered in this report. For example, physical performance enhancement such as performance enhancing drugs, nutritional supplements, and body temperature control are considered outside of the scope of this report.

The focus of this report is on systems to assist able-bodied populations for military applications, although this information may also be useful for industrial or medical applications. While the terminology and lab-based metrics are largely consistent across military, industrial, and medical user communities, there are military-unique aspects of physical augmentation systems that are addressed in this report. These aspects include potentially extreme environments (e.g., extreme temperatures [hot and cold]; humidity; mud; dust; sand; water [fresh and salt]; and chemical, biological, radiological, nuclear, and explosive [CBRNE] exposures), broad mission/task variability (duration and type), signature (noise, visual, thermal, and electronic), and limited availability of power sources. Desired military operational goals for a physical augmentation system include:

- Enhancing Soldier and small unit readiness – ensuring the highest possible percentage of Soldiers are deployable, such as through mitigating injuries during training and in theatre.
- Enhancing mobility and cross-domain maneuver for overmatch – enabling Soldiers to quickly transition across domains (land, air, space, sea) faster than their adversaries can react.
- Enabling force multiplication – utilizing fewer Soldiers but with the same warfighting power or better.
- Optimizing Soldier and small unit performance – increasing Soldiers' abilities through enhanced training, education, and physical performance approaches.
- Enhancing lethality and survivability – increasing weapon lethality and mobility and increasing Soldier protection against threats (e.g., ballistic, blast).
- Enhancing situational awareness – providing information needed to make rapid tactical decisions at the individual Soldier level.

Within the domain of technology development, there exists the technology readiness level (TRL) scale used to classify the technological maturity of developing systems (Department of Defense 2010). The scale ranges from TRL 1, which describes basic principles that have been observed and reported, to TRL 9, which describes systems proven through successful mission operations. Note that low technical maturity of a system does not preclude evaluation, as evaluations can begin even at the concept development phase and prior to building any physical prototypes. Evaluations of initial concepts can be based on feedback collected from potential users and subject matter experts. To advance to TRL 9, systems must undergo an iterative process of evaluation and redesign. This report focuses on

evaluation methods and metrics for prototypes spanning TRL 3 (analytical and experimental critical function and/or characteristic proof of concept) to TRL 6 (system/subsystem model or prototype demonstration in a relevant environment).

Currently, there is a significant proliferation of physical augmentation systems emerging from government laboratories, industry, and academia to meet an array of user applications. The methodologies and metrics herein are intended to be tailored by the evaluation team depending on the intended military environment and application. As such, as many or as few of these metrics may be applied as desired by the evaluation team.

4. Terminology

The purpose of this section is to define the commonly used terms in this report. The field of physical augmentation is an emerging and rapidly evolving field, which creates a critical need for common terms when discussing standards, testing, and evaluation protocols. It should be noted that the terminology included in this section is intended to complement current efforts (as of 2017) by NIST to establish definitions of commonly used terms. Definitions published by ASTM International and the International Organization for Standardization (ISO) are included and adapted as appropriate.

Term	Definition
Active	Powered; having a power source, such as batteries, that drives actuators, which move segments of the augmentation system.
Actuation	Movement of the physical augmentation system in order for it to perform its function; often accomplished by motors, springs, hydraulics, or pneumatics.
Adaptive control	Control scheme whereby the control system parameters are adjusted from conditions detected during the process (ASTM F3200-16: Standard Terminology for Driverless Automatic Guided Industrial Vehicles).
Augmentation	Performance enhancement beyond inherent human capabilities.
Augmentation system	Equipment designed to enhance performance; exoskeleton, wearable robot, wearable augmentation, and augmentation system are commonly used interchangeably.
Control system	Set of logic control and power functions that allow monitoring and control of the mechanical structure of the exoskeleton and communication with the environment (equipment and users). (Adapted from ISO 8373:2012: Robotics and Robotic Devices – vocabulary.)
Exoskeleton	Rigid human wearable device that augments, enables, or enhances motion or physical activity. In the context of augmentation, an exoskeleton is commonly powered or unpowered. Powered systems are commonly active or semi-active systems and include actuators and power supply. Unpowered systems are commonly passive or semi-passive systems. Note that exoskeleton, wearable robot, wearable

	augmentation, and augmentation system are commonly used interchangeably. (Adopted from NIST Terminology working group.)
Exosuit	Non-rigid human wearable device that augments, enables, or enhances motion or physical activity.
Full body augmentation	Augmentation designed to aid both the upper and lower body within the same system.
Lower body augmentation	Augmentation designed to aid at the hip, knee, and/or ankle joint.
Passive	Unpowered. Commonly used mechanisms include springs, elastic elements, and dampers.
Pilot Testing	Practicing the data collection procedures, gathering initial data for analysis, and finalizing the protocol.
Powered	Uses a power supply, commonly batteries, in order to function.
Protocol	Precise and detailed procedures and design to be followed for a research study or an evaluation.
Robot	Actuated mechanism, programmable, with a degree of autonomy, moving within its environment, to perform intended tasks. (Adapted from ISO 8373:2012: Robotics and Robotic Devices – vocabulary.)
Semi-active	System that is sometimes active and transparent during other times. For example, a system may be actively assisting during a walking task, but during any other motion there is no active augmentation occurring.
Semi-passive	System that is sometimes passive and transparent during other times. For example, a system may be passively assisting during a stair climbing task, but during any other motion there is no augmentation occurring.
Technology readiness level (TRL)	One level on a scale from 1 (lowest technology maturity) to 9 (highest technology maturity). Adopted by the Department of Defense (DoD) as a method of estimating technology maturity during the acquisition process. (TRLs described in the DoD Technology Readiness Assessment Guidance: https://www.acq.osd.mil/chieftechnologist/publications/docs/TRA2011.pdf. Accessed 31 January 2018.)
Unpowered	No external source of power such as batteries, compressed gas, or engines.
Upper body augmentation	Augmentation designed to aid the torso, arms, head, and/or neck.
Wearable robot	Exoskeleton or exosuit that is programmable in one or more axes with a degree of autonomy, moving within its environment, to perform intended tasks. (Adopted from NIST Terminology working group.)

5. Considerations for Protocol Development

5.1 Introduction

There are a number of considerations that should be taken into account when developing the protocol for a physical augmentation evaluation. This section addresses those considerations so that the most appropriate methods and metrics can be used in the evaluation. These considerations include the goals of the evaluation, the intended uses for the system, and the system's level of development.

In addition, the resources available, the training required for users, and safety of the user and evaluators must be considered.

This section is based on the process used to develop protocols for physical augmentation system evaluations conducted over roughly the last decade at ARL and NSRDEC. The physical augmentation systems evaluated were mostly prototypes in the early stages of development. As physical augmentation systems evolve, the protocols for their evaluations are also likely to evolve. In fact, to a certain degree, evaluations for a particular prototype will have to be somewhat customized to the functionality and purpose of that specific system. However, the general guidelines and considerations discussed here may be applied to the evaluation of systems in any stage of development.

5.2 Elements of Protocol Development

Although each institution may have their own requirements for writing a research protocol, most protocols will need to include a description of the following:

- Research questions to be addressed and associated hypotheses
- Subject population – influenced by expectations for the evaluation and system technical maturity
- Methods for evaluation – including system conditions, clothing and equipment, test environment, and tasks
- Metrics used to quantify human and/or system performance
- Resources, training, and schedule
- Safety and risk mitigation

Each of these elements is influenced to varying degrees by the technical maturity of the system, the goals for the evaluation, and the intended function of the system.

5.2.1 Research Questions and Hypotheses

In any protocol development, it is first important to understand the overall goals for the evaluation. Often there is an underlying research question that initially prompted the development of the system. For example, a system that has been developed to assist with load carriage is attempting to address the question of whether it is possible to aid load carriage with an external system without burdening the user. The specificity of the research questions addressed within a protocol will depend on, and may be limited by, the maturity of the system. Initial evaluations for a new prototype may focus on relatively small and informal tests to address

form and fit of the system and to determine if a prototype system is a good candidate for further development. To continue the load carriage example, an initial evaluation may simply compare the users' aided (with the system) and unaided walking with a light load in a controlled laboratory environment. Hypotheses may focus on changes in gait patterns or physiological demand with and without the system (e.g., knee flexion, body lean angle, walking speed, cadence, oxygen consumption) to understand the changes elicited by the presence of the system. These results can then be used to improve system design and control to accommodate user comfort and/or to better accommodate the user's natural gait. As the system is further developed, subsequent evaluations should address the system's ability to aid with carrying heavier loads or negotiating uneven terrain as would be required in a real-world scenario. Hypotheses may now focus on performance-based measures (e.g., course completion time) to determine more holistically whether the system provides a benefit to the user in a real-world environment. These more stringent evaluations can be targeted to produce generalizable results, which will serve as a baseline of performance needed before a decision is made for the augmentation system to move through the Army acquisition lifecycle.

Other considerations for research questions and hypotheses may be identified by the system developers or organizations sponsoring development. Often developers have secondary research questions they would like to address to improve system design or explore alternative uses for the system but lack necessary facilities or access to a specific subject population to conduct their own internal evaluation. If these questions can be accommodated in the protocol without substantially affecting the time, resource, and personnel requirements or increasing risk to the subjects, consider addressing these additional research questions. Often, it requires no extra data collection procedures, but only an alternative analysis of data already being collected. If no conflict of interest between the goals of the research staff and the system developers exists, evaluations that accommodate general research questions and also address specific secondary questions posed by the developers result in a wealth of data that can accelerate system development.

5.2.2 Subject Population

The number and type of participants needed for an evaluation are generally a function of the intended user group for the system, technical maturity of the system, the goals for the evaluation, and whether or not statistical significance in the findings is required. The first decision to be made is whether it is essential to recruit within a specific population. For evaluations early in the development cycle that are assessing basic motor functions with and without the system, recruitment from

the general population is appropriate. As the system becomes more mature and the protocol tasks become more specific, it may be necessary to recruit subjects with particular skills or qualifications. Specific subject selection may be done for both expertise and safety reasons. For the load carriage example, later evaluations involving full military load carriage may necessitate the recruitment of Soldiers experienced in load carriage. These subjects possess expertise in military load carriage and thus can provide the best feedback regarding the military utility of the system being tested. These individuals are also well-trained and capable of the physically demanding load carriage task used for the evaluation, ensuring that the protocol can be conducted safely.

For early system evaluations, when expectations for system performance are uncertain, it may be unrealistic to seek statistical significance in the results. Often, early evaluations are exploratory. Collecting data with a few capable subjects can be sufficient to provide data to inform further development of the system and generate trends for anticipated performance changes in future evaluations. For more mature systems, a greater number of subjects is typically desired. At this stage, the system has been refined and researchers will have better developed hypotheses and expectations for system performance. As with most research efforts, data from earlier or related studies may be used to conduct a power analysis to determine an appropriate sample size for seeking statistical significance in the results. It is perhaps more critical in the evaluation of physical augmentation systems to adjust the required subject number to account for possible subject attrition. As will be discussed in a forthcoming section of this report, training subjects with an augmentation system and with the tasks developed for the evaluation prior to formal data collection is very important for conducting a fair evaluation of the system. Such training may involve having subjects complete portions of the experimental protocol with and without the system during which any necessary adjustment of the system can be made for each participant. During training, injuries, abrasions, blisters, and other complications may arise that result in maladaptation to the system or subject dropout. If data collection on a particular number of subjects is required, it is recommended that more subjects are recruited and trained on the system to serve as alternates for the formal evaluation.

5.2.3 Methods for Evaluation

Developing the methods for a system evaluation requires careful selection of the specific tasks to be performed, the environment in which the evaluation will take place, the instrumentation to be used, and the appropriate system conditions.

5.2.3.1 Task Selection

Understanding the intended use of the system is the first step in selecting appropriate tasks. Physical augmentation systems for the military could include systems that aid with a variety of tasks including walking, running, load carriage, manual materiel handling, tool use, injury prevention, and many others. Tasks used to evaluate the system should simulate as closely as possible the task that the system has been designed to support. The task should be scaled, however, for the system's current level of technical maturity. For example, the approach described by Mudie et al. (2018) initially uses simple laboratory tasks and then progresses to complex field tasks. In early evaluations, it is good practice to evaluate the system first using a modified, low-difficulty version of the task and gradually increasing difficulty to identify changes in system performance. For example, if a system is intended to aid in carrying excessive loads, it is recommended to begin with a low or moderate load to ensure the system is reliable during these easier tasks. If a system is intended to augment sprinting, ensure that the system reliably permits walking or jogging first and gradually increase speed. Varying conditions or level of task difficulty during an evaluation can be an effective means of identifying the benefits and limitations of a specific system. Tasks that have been used in previous evaluations can be found in Table 1. When developing a protocol, consider how these tasks may be scaled for the system's TRL. If the system will be evaluated multiple times as it is developed, it may be a good approach to use scaled versions of the same tasks to gradually evaluate and challenge the development of the system.

Table 1 Suggested evaluation environments, tasks, and metrics based on technical maturity

TRL	State of system	Environ-ment	Tasks	Metrics
3-4	Proof-of-concept	Lab	• Tasks performed under loads that may or may not be operationally relevant • Tasks limited to laboratory environment o Static standing o Functional range of motion o Walking on treadmill, slow to moderate pace, 0% grade o Stepping up/down, stepping over obstacles	• Biomechanical (Section 6.1) o Postural stability o Spatiotemporal gait o Kinematic o Kinetic • Physiological (Section 6.2) o Metabolic o Muscle function o Psychophysiology • Operational (Section 6.3) o Movement ▪ Preliminary mobility • Human factors (Section 6.4)
4-5	Early prototype	Lab & field	• Tasks performed under loads approaching operational relevance • Tasks in laboratory and field environments In addition to the tasks specified under TRL 3-4, o Firing positions o Vertical jumps, squats, lunges o Walking on treadmill, moderate pace o Walking overground, self-selected pace o Stepping up/down, stepping over obstacles in the field environment	In addition to the metrics under TRL 3-4, • Operational (Section 6.3) o Movement ▪ Traverse natural terrain ▪ Road march

Table 1 Suggested evaluation environments, tasks, and metrics based on technical maturity (continued)

TRL	State of system	Environ-ment	Tasks	Metrics
5-6	Advanced prototype	Lab & field	• Tasks performed under operationally relevant loads for longer durations • Tasks in laboratory and field environments with more emphasis on outdoor and operationally relevant tasks In addition to the tasks specified under TRL 3-4 and TRL 4-5, o Long-jumps, stairs, underpass, hurdles o Jogging and running on treadmill o Walking overground, self-selected pace, paved and rugged terrain, 3–6 miles o Obstacle and/or urban terrain course • Tasks performed by secondary users (i.e., maintenance personnel)	In addition to the metrics under TRL 3-4 and TRL 4-5 • Operational (Section 6.3) o Strength o Movement o Marksmanship

The difficulty posed by the selected tasks should be appropriate for the system's technical maturity level. Early in development, it is common for augmentation systems to be bulky or heavy. Often the initial emphasis is on proving the efficacy of the technology. Only once they have achieved a "proof of concept" for system functionality will developers focus on streamlining and miniaturizing components. A bulky or heavy system may be an indication of low technical maturity. With these types of systems, it may be inappropriate to evaluate the system using highly dynamic physical tasks. Lab-based functional tasks, range-of-motion tasks, or ambulation restricted to the treadmill are recommended. These controlled lab-based scenarios will help address the effect of the system without exposing the user to undue stress or increased injury risk. With adequate, proven safety mechanisms in place to protect the user, a heavy or bulky system may be evaluated during more difficult physical and dynamic tasks. Similarly, systems that require tethers for power or actuation (e.g., pneumatic systems requiring a compressor or a system with off-board actuators) may limit the selection of highly dynamic tasks. Appropriate tasks must be selected on a system-specific basis with considerations for known system limitations and user safety. Evaluations that are too rigorous for the current technical maturity of the system may result in prematurely discarding promising technology. Conversely, evaluations that do not challenge the system will fail to help developers identify weaknesses in the current design, resulting in the continued development of ineffective technologies.

In addition to system-specific tasks, it is also recommended that secondary tasks be included in the evaluation that provide insight on the system's current field readiness or compatibility with other equipment. Regardless of its intended function, a system should not impede a Soldier's mobility, cause discomfort, or interfere with the use of other essential equipment (e.g., rifle or rucksack). Additional tasks that examine range of motion, ability to assume common firing postures, or agility or functional tasks are helpful for determining the physical restrictions imposed by the presence of the system. This information can be used to direct the redesign of bulky or cumbersome components. Subjective surveys and interviews are also valuable tools for identifying sources of pain or discomfort, documenting the user's comments about their experience with the system, and collecting suggestions for future improvements.

5.2.3.2 Environment

Another important consideration in protocol development is the environment in which the evaluation will be conducted. For military applications, most physical augmentation systems will need to function freely and untethered in a field environment. Ideally, systems could be evaluated in a simulated field environment,

but often this is not feasible due to system constraints or safety concerns. System evaluations may be conducted in a laboratory environment, outdoors in a real-world environment, or include a combination of the two. Laboratory evaluations permit highly controlled, high-fidelity data collection with minimal abuse to the system. The disadvantage, however, is that they do not effectively quantify system performance in an operational scenario and there are limitations to the tasks that may be performed. Outdoor assessments are more operationally relevant and indicative of overall system efficacy, but the types of metrics that may be used to quantify performance are more restricted due to measurement equipment portability constraints.

Selecting the appropriate environment for an evaluation will depend largely on the system's robustness as well as health hazards and safety concerns. Robustness refers to resistance to environmental elements and to breakage of system components. For full field use, systems must be resistant to water, dirt, mud, rain, heat, cold, and other environmental factors. Often, resistance to these weather and terrain elements is a secondary concern that is addressed only after the system's functionality has been validated in indoor environments. Mature systems should also be robust to repeated use, impact, minor snag hazards, and cyclic/cumulative stress or damage. Often this is achieved through extensive testing to identify the most appropriate materials for the system and building in strain relief for cables and electronic connections. If systems are at a low level of technical maturity (e.g., exposed electronics, no stress relief for cables/connectors), it is advised to conduct testing in a laboratory environment. The biggest concerns are often wet electronics and breakage. If wet electronics will short the system, understand the risks of sudden system failure prior to outdoor testing. If wet electronics expose participants to a risk of electric shock, it is recommended that testing occur in a controlled lab environment until the electronics can be waterproofed. However, if most of the electronics are contained in a housing or are reasonably protected from the environment, outdoor testing is possible as long as the weather and test site conditions are monitored. During outdoor or field assessments, having extra materials on hand to fix commonly broken elements of the system will ensure that testing can proceed with minimal interruption. If the system includes protruding cables or wires, this may also dictate the environment used for evaluation. These snag hazards are a safety risk to the user but also a likely source of system failure if pulled or disconnected. If these snag hazards can be secured to minimize risk, outdoor testing is still possible with careful monitoring.

Having tethers for power actuation may also affect the selection of the test environment. Systems that are tethered either for power or actuation requirements are likely restricted to indoor testing. A possible exception is if a support vehicle

can be within range of the system at all times. This does, however, restrict the user's agility during any outdoor or dynamic tasks. These tasks must also be executable without the user becoming entangled in the tether. This testing paradigm is not recommended. If a system is intended to afford free ambulation through the environment but still relies on a power or actuation tether, it is likely an indication that the technical maturity of the system is not great enough to warrant outdoor testing. Generally, if the system requires a power or actuation tether, testing should be restricted to laboratory testing or testing near a stationary support vehicle.

Active systems with reliable controls may be tested outdoors in more real-world environments, which introduce new challenges to the user and the system. The duration of this testing will be determined by power supply (e.g., battery) life, and it is recommended that spare power supplies be available during testing. Although power supply life may be predictable in the lab, systems may require additional power supplies (e.g., additional batteries) when operating over uneven terrain or accommodating fluctuations in speed that are typically observed with more realistic testing. Beyond this increase in power consumption, extreme ambient temperatures (hot and cold) in outdoor environments may further reduce power supply capacity, and thus the system's operating duration. This more realistic assessment of human/system performance can reveal important deficiencies in power output, control algorithms, or the human/system interface that can inform design improvements.

ARL and NSRDEC have had success using a combination of laboratory and outdoor tasks in evaluating prototype systems. This has enabled the collection of high-fidelity human biomechanical and physiological data in controlled environments as well as overall system performance in more operationally relevant environments. Therefore, if possible, incorporating both types of environments in the protocol is recommended.

5.2.3.3 System Conditions

Often, physical augmentation systems are developed to aid with a specific function. While the primary goal of system evaluations is to quantify performance with these systems, it is also important to consider how the system's presence will affect the user when the specific function is not being performed. For a system to be fielded, it must demonstrate added capability that provides military overmatch compared to the current military standard. It also must not impede the Soldier's ability to perform other critical tasks. Thus, an important consideration when developing a protocol is which system conditions to compare.

Evaluations are typically conducted by comparing performance with the system actively powered ("ON" condition) to performance without the system (No Device condition, "ND"). These evaluations can determine whether the system provides a significant capability that justifies its additional weight or power requirements. Early in a system's lifecycle, however, it may be inappropriate to compare these "ON" and "ND" states. Instead, early evaluations may elect to compare the actively powered system state to an unpowered state ("OFF" condition; system worn but not powered). These evaluations can help identify mobility/compatibility issues along with performance decrements due to the current system weight and design. Finalized designs may also benefit from evaluating the system in an "OFF" condition. These evaluations can reveal how the Soldier's performance may be affected if an active system is suddenly disabled in the field due to damage or loss of power.

Some augmentation systems are designed to function passively. That is, they provide augmentation through mechanical design rather than powered actuation. In these cases, "ON" and "OFF" as described above may essentially represent the same system condition. For passive systems, "ON" and "ND" are likely the two test conditions of interest permitting the determination of total performance changes with the presence of the system. A possible exception may be if a mechanical linkage can be easily removed to render the system nonfunctional (e.g., a modified "OFF") while still being borne by the user. In this case, "ON" and "OFF" comparisons may be useful to determine whether it would be more beneficial to have the system in its enabled or disabled state when performing non-system-supported tasks.

5.2.4 Metrics

Previous evaluations have produced a list of metrics (Table 1) that have been used to successfully distinguish between system conditions (e.g., ON, OFF, and ND). While it is not an exhaustive list, categories and use cases are provided such that these metrics may be leveraged in the design of future evaluations. Specific metrics are described in detail in Section 6.

Metrics will, of course, be chosen as appropriate to quantify the elements of performance that are of interest during the selected tasks. These metrics may be influenced by the goals of the evaluation, the environment in which the evaluation will occur, tasks selected, or the form factor of the system. If the goal of the evaluation is to conduct a formal evaluation to quantify specific changes in physical performance, biomechanical or physiological metrics may be most appropriate. If, instead, the evaluation is intended to be an assessment of real-world system performance, operational metrics may be most useful. For evaluations whose

primary purpose is to collect user feedback regarding human/system interface, the human factors metrics would be appropriate. Often, studies will employ a variety of metrics to answer questions regarding different elements of performance. While this is certainly a useful approach that yields a rich data set, it is advised that each of the metrics be selected for a particular purpose with specific hypotheses to be tested. Excessive instrumentation can interfere with system performance and negatively impact the user's perception of their experience with the system.

The evaluation environment, the tasks to be performed, and the form factor of the system help to refine the selection of metrics. Certain metrics such as those requiring optical motion capture cannot be collected in a field environment easily. In addition, some portable measurement systems may not be rugged enough to be used in an obstacle course environment due to the risk of equipment damage. Points of attachment may preclude the use of skin-mounted sensors such as motion capture markers or surface electromyography electrodes. Metrics must be selected that quantify elements of performance without interfering with the subject's ability to complete a specific task.

It must be acknowledged that augmentation systems discussed here do not function independently of a user. When a system interfaces with a human operator, the overall system being evaluated becomes the combination of the user and the system. The system may impose specific changes on the behavior of the user that vary between users due to individual differences. For example, consider a system that is intended to aid in load carriage by providing additional power into the gait cycle. One possible effect may be that the user relies on the system to provide that power and their metabolic cost for a specific load carriage task declines. Another possible effect is that the user will continue working as hard as they normally would while also receiving the additional benefits of the system. In this case, metabolic cost for load carriage would not decline with system use, but it is likely that the load carriage task would be completed more quickly. Thus, a single metric may be insufficient to quantify the performance of the system. When researchers consider appropriate methods for evaluation, they should also consider all the possible means by which a system may be eliciting its effects. Utilizing multiple metrics to quantify several aspects of user/system performance will result in the most informative evaluation.

5.2.5 Resources, Training, and Scheduling

5.2.5.1 Resources

As with any research endeavor, the resources available such as research staff, test participants, time, equipment, and funding should be identified as they influence the selection of methods and metrics to be included in the evaluation. The research

staff available to assist with the data collection, processing, and analysis strongly influences the time and funding needed for the evaluation. The cost of conducting an evaluation includes expenses such as paying for the time of the research staff to collect, analyze, and report data, as well as the supplies necessary to conduct the study. Travel costs and compensation for the subjects may also need to be budgeted. Depending on the complexity of the evaluation and the number of systems available, the research staff requirements can be quite large. Staff are needed not only to carry out data collection and analysis but may also be needed to assist test participants to don the system, escort them during training and data collection trials, or operate support vehicles. Personnel requirements to consider also include members of the system design team and research staff. For early evaluations, it is preferred that members of the design team be present to repair breakages and maintain the system. This consideration should become less critical, however, as the system matures. Systems with a high level of technical maturity should be expected to complete a protocol without breakage or adjustment of controls.

Another important resource is the equipment that is available. This includes the number of augmentation systems and the data collection equipment (e.g., units for measuring metabolic cost). The equipment available will influence the schedules for training and data collection. If there are multiple augmentation systems available, then training can be done with multiple participants at the same time (assuming research and developer staff are available as well). Similarly, if multiple units of the data collection equipment are available, then data collections with multiple participants can be done at the same time. Often, however, multiple units of data collection equipment such as force sensing treadmills are not available. In that situation, the data collection sessions have to be staggered so that only one participant at a time is scheduled for the treadmill. If possible, it is recommended that multiple augmentation systems and multiple units of the data collection equipment be available for an evaluation. This will allow the evaluation to be completed quickly and efficiently. This also helps in the event that an augmentation system or piece of data collection equipment malfunctions. Generally, if the number of available augmentation systems and data collection resources are known, use of time can be optimized by determining which elements of the evaluation may be executed in parallel and which must occur in series.

5.2.5.2 Training

Training is a key part of the evaluation of any physical augmentation system. To conduct a fair assessment of the current capabilities of a system, the participants must be adequately trained on its use. The appropriate duration of this training varies by system and its specific application. This may initially be unknown,

particularly in the case of early prototypes and novel technologies. Additionally, training requirements for a particular technology may change over time as TRL increases. Typically, participants in evaluations conducted at ARL and NSRDEC have begun their training by donning the system and briefly practicing the evaluation tasks with the system unpowered. Then they practiced the evaluation tasks with the system powered. Initial criteria for determining when participants were sufficiently trained have been when they could demonstrate familiarity with the operation of the system and when they reported that they were comfortable using the system to perform the evaluation tasks. Historically, during evaluations of load carriage gait-assist systems conducted by ARL and NSRDEC, training was conducted over multiple days to familiarize participants with both the weight and functionality of the system. This paradigm is supported by a recent evaluation of adaptation to a hip exoskeleton, which found that changes in metabolic cost stabilized beyond the second training session (Panizzolo et al. submitted/in review). Formal criteria have not been established for determining when participants have been sufficiently trained to use a physical augmentation system. Research is needed to establish the best training tasks, the appropriate number of training sessions, the length of each training session, and the retention of training.

Time required for an individual to train with the system as well as the time allotted between required training sessions should be factored into the protocol development. If more than one system is available, training and experimental protocols can often be conducted with multiple participants concurrently. This approach makes good use of the time of all involved: participants, research staff, and the system development team. However, if only a single system is available, training for each participant as well as experimental protocols must be conducted serially. It is also worth noting that the additional operation time required for training increases the potential for system breakage, particularly for systems of low technical maturity. Training itself may also result in abrasions, blisters, or other injuries that can affect the participant's readiness for the experimental protocols. In the event that participants sustain minor injuries during training (e.g., abrasions, blisters), it is recommended that time be allotted for treatment and recovery prior to data collection. If a more severe injury is sustained (e.g., muscle strain, ankle sprain), it may be best to replace that participant if possible. If the injury resulted directly from a system malfunction, it may be inappropriate or unethical to continue with the evaluation until the system is modified to prevent future injuries.

5.2.5.3 Scheduling

Time is, of course, a critical resource to account for when preparing the protocol for an evaluation. For the evaluation to run smoothly, good scheduling is essential.

The time required for each step in the evaluation, from donning the system, to training participants to use it, to rest time between training sessions, to each task in the data collection trial, to doffing the system, and completing after-action questionnaires needs to be estimated. In addition, time should be allowed for unexpected malfunctions and repairs for systems with low technical maturity. Time will also vary depending on the type of evaluation needed. Some evaluations with proof-of-concept systems (TRL 3) have been done in 3 to 5 days, but that included only a limited number of participants (one to six) and required training and data collections to occur on the same day. More in-depth evaluations require more time, subjects, and research staff. As an example, recent evaluations with early prototypes (TRL 4) have been done over two-week periods with six to eight Soldiers. These evaluations provided quantitative and qualitative data along with feedback for the program manager and development team to use in assessing the progress of the systems. In these evaluations, the first week was for setup by the system developers and training of the participants. Data collections were held the second week using a research staff of 11. There were two data collection sessions per participant with at least one rest day between sessions. The data collection sessions involved laboratory and outdoor metrics collected during walking on an instrumented treadmill and overground through an outdoor cross-country course. This larger, full evaluation required considerably more research staff, subjects, and time than a simpler proof-of-concept evaluation.

5.2.6 Safety and Risk Mitigation

In all ARL and NSRDEC evaluations, safety is a priority. Prior to collecting data with any research subjects, research staff thoroughly discuss with developers the risk of injury or malfunction with a system. It is important to find out from the developer the results of tests they have done with the system to assure its safety and effectiveness. In particular, it is important to know about tests or inspections to identify injury hazards. It is also important to know how many subjects tested the system and basic information such as their height and weight to know what size individuals the system will fit. In addition, results of any electrical or thermal safety, flammability, biocompatibility, durability, mechanical, and software testing should be shared by the developer. For high TRL systems, the developer should provide the results of formal tests conducted in accordance with standards published by organizations such as ISO and the American National Standards Institute to describe the relative risk associated with system use. Even with information and test results from the developer about safety, time is included in the schedule to permit research staff to don the system and conduct internal pilot testing to identify potential safety hazards. Any safety risks are documented, and the

research protocol only proceeds when there is an appropriate plan in place to mitigate these risks.

Regardless of the specific research facility (academia, industry, government), protocols using human subjects must be reviewed by an Institutional Review Board (IRB). The documentation that is needed for each protocol depends on requirements from the individual IRB but typically includes a thorough description of all data collection equipment, tasks, and surveys; the population to be targeted as participants and inclusion/exclusion criteria for their selection; potential risks associated with participation; plans to mitigate these risks; and the process through which participants will provide informed consent. When evaluating a physical augmentation system, additional steps may be needed such as informing or obtaining approval from an institutional Safety Office. The Safety Office may approve the protocol, or they may require some modifications. If the subjects participating in the evaluation are Soldiers, a Safety Release will be required. For the Army, Safety Releases are issued by the US Army Evaluation Center. A Safety Release is a formal documentation of the system, including a description of its components, a description of how it works, and potential risks to users. Institutions may also require a demonstration of the system prior to the evaluation. Based on the documentation and demonstration, a Safety Release will be issued that includes the warnings, cautions, procedures, and mitigations that must be followed for Soldiers to use the system in the evaluation.

When developing a protocol, consideration should also be given to the state of the system during an unexpected failure. If, for example, unexpected system failure results in the system simply going slack and following the user's motion without hindering agility, this poses only a minimal risk of injury. However, if the system would seize and suddenly restrict motion, serious consideration needs to be given to safety precautions that minimize the risk of injury to the subject. When possible, safety harnesses may be used in a laboratory environment. Systems may also be equipped with quick-release or emergency stop mechanisms to reduce the risk of injury. It is also considered a best practice to have a member of the research team near the participant at all times during the evaluation. This permits a quick response if the participant experiences a problem with the system. If the evaluation involves maneuvers over difficult or remote terrain, a support vehicle is recommended. The support vehicle can provide water and first-aid supplies as well as be used to quickly transport the participant indoors if required due to inclement weather, injury, or system malfunction.

As part of the commitment to safety, participants are generally not asked to do something that the researchers who will be conducting the evaluation have not already done or would not do themselves. As mentioned, pilot testing is conducted

with ARL and NSRDEC staff to address possible safety concerns prior to testing with research subjects. Such pilot testing is strongly encouraged before any physical augmentation system evaluation. This process helps the researchers to understand how to don and doff the system, how the system works, and what it should feel like for the participants going through the evaluation. In addition, this allows the researchers to make adjustments to the evaluation procedures if necessary and identify performance limitations, safety issues, or unexpected failure modes of the system.

5.3 Protocol Development Summary and Recommended Best Practices

The process of developing a protocol involves a number of steps, including identifying the primary research question, selecting the tasks and metrics to be used to quantify performance, and planning for subject training and personnel needs. System technical maturity, available resources (e.g., facilities, data collection equipment, time), and user safety must be carefully considered when designing the protocol. Although each evaluation needs to be somewhat customized to the unique system being evaluated and the goal of the evaluation, the considerations described here should aid in the development of a safe and successful protocol. In addition, the following recommendations are based on the experiences of researchers at ARL and NSRDEC.

- If it does not interfere with the goals of the research staff, and if time and resources will permit, consider addressing additional research questions posed by developers, which can result in a wealth of data that can accelerate system development.
- Recruit and train more than the minimum number of subjects needed for the evaluation to account for possible attrition.
- For the initial evaluation of a system, start with easy versions of the evaluation tasks and gradually increase the difficulty to identify changes in system performance.
- To minimize delays during the evaluation, suggest that developers have extra parts and supplies on hand to repair their system.
- If possible, incorporate laboratory and outdoor tasks in the evaluation to get data in a well-controlled environment and a more operationally relevant environment.

- To avoid interference with the system, which may negatively impact the subject's impression of it, use the minimum instrumentation necessary to answer the research questions posed for the evaluation.
- Particularly for proof-of-concept systems and early prototypes, have developers at the evaluation to maintain and repair the systems.
- Have multiple augmentation systems and multiple units of the data collection equipment available for the evaluation.
- For safety, have a member of the research team near the participant whenever he or she is using the augmentation system. In addition, if the evaluation takes place in terrain that is remote or difficult to maneuver through, have a support vehicle near the participant.
- Have members of the research staff use the system that will be evaluated to pilot test the procedures.

6. Performance Metrics: Biomechanical, Physiological, Operational, and Human Factors Metrics

This section of the report focuses on biomechanical, physiological, operational, and human-factors-based metrics that have been used at ARL and NSRDEC for physical augmentation system assessments. Though cognitive improvements may result from the use of physical augmentation systems, these metrics are focused on evaluating physical performance augmentation. The metrics identified here include both quantitative and qualitative metrics. All or a subset of these metrics may be used based on the TRL of the system, the population of interest that will utilize the system, and the intended physical augmentation system application(s). Not all metrics listed here will be relevant to all evaluations, all types of physical augmentation systems, or all types of use-cases. Typically, the metrics outlined here are used to systematically compare the ND condition (no system being worn or used) to the ON condition (system worn and active). However, in many cases the OFF condition (system worn but not active) may need to be considered as well. When evaluating physical augmentation systems, it is expected that a successful system will demonstrate significantly improved performance over the ND condition. Within each category of metrics (i.e., biomechanical, physiological, operational, and human factors) there are brief descriptions of the current lab and field capabilities for measuring the metrics. These descriptions include the type of equipment needed and, in some cases, additional information relevant to data collections. The current lab and field capabilities refer to the capabilities of Army and DOD laboratories that conduct these types of evaluations. The metrics outlined

in this section can be used in many of the phases of system development including TRL 3-4 Proof of Concept testing, TRL 4-5 Early Prototype testing, and TRL 5-6 Advanced Prototype testing (Table 1).

6.1 Biomechanical Metrics

Biomechanical metrics of performance are physics-based measures that quantify the movement or structure of a biological system of interest. For the purposes of this document, biomechanical metrics will include only those measures that may be helpful in quantifying the behavior and performance of a physical augmentation system user, a physical augmentation system, or the relationship between the user and the system. Biomechanics-based metrics of interest for physical augmentation system testing may include the following.

6.1.1 Postural Stability Metrics

Postural stability assessments are used to determine how particular conditions (e.g., injury, disability, load carriage) or interventions (e.g., balance training, orthotics) influence standing balance or stabilization during stance or after a ground contact event. Metrics of interest often include variability of ground reaction forces (Le Clair and Riach 1996; Karlsson and Frykberg 2000; Sell 2012), center of pressure (COP) deviations, path lengths, and velocities (Yaggie and McGregor 2002; Heller et al. 2009; Rugelj and Sevsek 2011), limits of stability (Holbein and Chaffin 1997; Lee and Lee 2003; Schiffman et al. 2006, 2008), and dynamic postural stability index (Wikstrom et al. 2005). In general, decreases in postural stability are associated with increased risk of slips, trips, falls, and injuries. Therefore, outcomes indicating increased stability are generally favorable. Postural stability metrics may include the following:

- Variability of Ground Reaction Forces – assesses the ability to maintain balance during static states as well as during transitions from a dynamic to a static state by evaluating the standard deviation of the ground reaction forces (Sell et al. 2012). If a reduction in the risk for falls is desired, then outcomes that indicate reduced variability are generally favorable.
- COP Deviations and Path Length – COP describes the point through which the total sum of forces acting on a body exerts its effect. COP position shifts in response to even small postural adjustments. COP length measures the maximum COP distance in the medio-lateral and anterior-posterior directions and the total length traveled by the COP to reach the maximal distance during static standing (Rugelj and Sevsek 2011). Additionally, the center of mass (COM) trajectory in relation to COP can be measured to

assess stability. The "COP-COM" postural stability variable is calculated as the difference between the COP position and the vertical projection of the participant's center of mass at each instant in time during a postural stability trial. This measure may be calculated for either mediolateral or anteroposterior sway directions, and the total COP-COM is expressed as a root mean square of these instantaneous differences over the duration of a trial. Larger COP-COM measures are indicative of poorer stability. This measure has been shown to have high reliability, and is sensitive enough to identify stability differences among those with compromised balance (Corriveau et al. 2000, 2001).

- Limits of Stability – measures maximum COP distance in the medio-lateral and anterior-posterior directions beyond which a body can no longer return to its initial position or state (Schiffman et al. 2008). A reduction in the limits of stability may compromise balance during static and dynamic tasks, thus increasing the risk of falls.
- Dynamic Postural Stability Index (DPSI) – assesses the ability to maintain balance while transitioning from a dynamic state to a static state. DPSI is a composite score of the medio-lateral, anterior-posterior, and vertical stability indices where mean square deviation fluctuations are assessed around a 0 point, rather than standard deviations assessing fluctuations around a group mean (Wikstrom et al. 2005). Outcomes that indicate increased dynamic stability are generally favorable.

Current Lab Capability: Current lab-based force platforms and pressure sensing systems can measure all relevant postural stability metrics.

Current Field Capability: Not all relevant postural stability metrics can currently be measured accurately in a field setting. However, COP can be measured utilizing pressure-sensing insoles integrated into footwear.

6.1.2 Spatiotemporal Gait Metrics

Spatiotemporal metrics quantify the spatial (distance) and temporal (time) components of gait. Common measures include gait speed, stride length and frequency, step width, and double support time (Winter 1987; Hills and Parker 1991; Vaughan et al. 1992; Owings and Grabiner 2004a, 2004b; Hollman et al. 2011). Measurement of spatiotemporal metrics can be used to identify deviations from normal, healthy, or baseline gait patterns. In general, outcomes in which the baseline measures of an individual's preferred gait pattern are not altered by the use of an intervention (such as a physical augmentation system) are desirable; however,

there can be exceptions such as advantageous increases in preferred (self-selected) walking speeds. Spatiotemporal gait metrics may include the following:

- Gait Speed – refers to the distance covered by the body per unit time. Reductions in self-selected speed may indicate hindrance of natural capability or greater physical workload. In addition, individuals may adopt a slower walking speed to increase dynamic stability during walking (Kang and Dingwell 2008) or to reduce ground reaction forces, thereby lessening the risk of injury (Browning and Kram 2007).

- Cadence (also referred to as step time or stride time) – refers to the number of steps or strides per unit time. Deviations in cadence from baseline conditions are typically associated with increases in metabolic cost (Umberger and Martin 2007). Imposing a fixed-cadence can increase ground reaction forces (Gutekunst et al. 2010), while increasing cadence during running can be a technique for reducing joint loading (Lenhart et al. 2014).

- Stride Length – refers to the distance between consecutive points of heel contact of the same foot. Deviations in stride length from baseline conditions are typically associated with increases in metabolic cost (Russell et al. 2010). Reducing stride length during running can reduce impact forces (Thompson et al. 2014), which may actually lower the risk of lower extremity injury.

- Step Length – refers to the distance between consecutive points of heel contact of the opposite foot. Deviations in step length from baseline conditions are typically associated with increases in metabolic cost (Donelan et al. 2002). Reducing step length during running can reduce joint loading (Heiderscheit et al. 2011), which may actually lower the risk of lower extremity injury.

- Step Width – refers to the mediolateral distance between the two feet during gait. Changes in step width from preferred baseline gait patterns are associated with gait instability and can result in increased metabolic cost (Donelan et al. 2004).

- Step Width Variability – refers to the range and standard deviation of the step width. Changes in step width variability from preferred baseline gait patterns are associated with gait instability (Owings and Grabiner 2004a, 2004b).

- Double Support Time – refers to the amount of time both feet are in contact with the ground during the gait cycle. Increased double support time is an

adaptation to increase stability during walking (Winter et al. 1990), but does not indicate improved stability over normal walking.

- Stance/Swing Ratio – refers to the ratio of the stance period to the swing period. For normal walking, this ratio is typically 1.6 (62% stance phase, 38% swing phase; Vaughan et al. 1992). Changes in this ratio from preferred baseline gait patterns are associated with gait instability, or they may be an indication of adaptation to the augmentation system or the environment.

- Toe Clearance – refers to the distance between the toe and the ground when the foot swings forward. The most commonly used measure is Minimum Toe Clearance (MTC). A decrease in this value indicates an increased risk of tripping and falling if there are obstacles in the path or changes in the incline of the path (Begg et al. 2007; Mills et al. 2008).

Current Lab Capability: Current lab-based motion capture systems (optical, video, and IMU-based) can measure all relevant spatiotemporal gait metrics. Force plates and pressure sensor systems can also be used to measure most of the spatiotemporal gait metrics.

Current Field Capability: Limited field capability at this time using optical, IMU, and pressure-sensing systems. Some, but not all, relevant spatiotemporal gait metrics can be captured in a field setting at this time.

6.1.3 Kinematic Metrics

Kinematic metrics describe the motion of the body without concern for the forces associated with those motions. Common kinematic measures include joint angles, joint ranges of motion, and joint angular velocities. Kinematics can be used to identify gait patterns or mechanics that may be indicative of poor joint stability and higher injury risk including sprains and strains, overuse injuries, knee ligament and meniscus damage, and patellar tendon and patellofemoral pain (Jones and Knapik 1999; Allen et al. 2000; Willems et al. 2005; Hewett et al. 2005; Malliaras et al. 2006; Boling et al. 2009). Kinematic metrics may include the following:

- Maximum and Minimum Joint Angles – refers to the angle between two body segments linked by a common joint. In some instances, increases in joint angles as compared to a baseline during specific tasks may be associated with increased injury risk such as hyperextension of the joint. This has been found to be true in the joints of the lower extremity including the ankle and knee (Malinzak et al. 2001; Konradsen et al. 2002; Ford et al. 2003; Hewett et al, 2005; Kristianslund et al. 2011). In other cases,

increased joint angles (e.g., knee flexion angle during stance) are indicative of adaptations to reduce loading rates and injury risk (Blackburn and Padua 2008; Podraza and White 2010).

- Range of Motion (ROM) – refers to the difference between the maximum joint angle and the minimum joint angle (or "range") of a joint (Mitchell et al. 2017). Reach distances are another ROM measure. Differences observed in joint and reach ROM from a baseline condition may indicate non-optimal movement patterns, which may result in a greater metabolic cost or increased risk of musculoskeletal injury (Willems et al. 2005; Malliaras et al. 2006).
- Joint Angular Velocity – refers to the rate of rotation of a joint and is related to the dynamics of muscle activation and force generation during gait. Differences observed in joint angular velocity from a baseline condition may indicate non-optimal movement patterns, which may result in increased risk of musculoskeletal injury (Gehring et al. 2009).

Current Lab Capability: Current lab-based motion capture systems (optical, video, and IMU-based) can measure all relevant kinematic metrics including joint angles, ROMs, and velocities. Static ROM measures can also be captured with inclinometers and goniometers.

Current Field Capability: Limited field capability at this time. Commercial off-the-shelf (COTS) kinematic measurement systems do exist and are designed to collect kinematic data in field-relevant environments. For gross kinematics these systems can perform well. However, the fidelity of measurement needed for biomechanical research is currently not met by these systems.

6.1.4 Kinetic Metrics

Kinetic metrics quantify the forces associated with motion. Typical kinetic measurements include ground reaction forces (GRFs), loading rates, internal joint forces, moments, and powers (Novacheck 1998; Gok et al. 2002; Decker et al. 2003; Hewett et al. 2005; Yeow et al. 2009; Schache et al. 2011; Ali et al. 2012). These measures provide a means to quantify the severity of impact with the ground and the demand on the lower extremity joints. The design of the augmentation system will influence the instrumentation for and interpretation of the specific kinetic metrics listed below. If the system transmits some of the load to the ground or if it transmits some of the load to the user or across one of the user's joints, force and torque sensors (placed between the user and the system) would be needed to determine the kinetic metrics for the user and the system separately. Generally, larger GRFs and higher loading rates, felt by the user, are considered hazardous and

more likely to cause injury (Davis et al. 2004; Milner et al. 2006). Greater internal (i.e., internal to the user) joint forces and moments are associated with shear, strain, and stress fracture injuries (Hewett et al. 2005; Milner et al. 2006). When carrying a load, as Soldiers are often required to do, GRFs increase in proportion to the load (Tillbury-Davis and Hooper 1999) and may increase further during prolonged fixed-cadence military marching (Gutekunst et al. 2010). Kinetic metrics may include the following:

- Peak Ground Reaction Forces (GRF) – The ground reaction force (GRF) is the force exerted by the ground on a body in contact with it. Increases in peak GRFs are risk factors for knee injuries and tibial stress fractures (Davis et al. 2004; Hewett et al. 2005; Milner et al. 2006). Increases in other GRF metrics may be associated with increased injury risk and/or non-optimal movement patterns.

- Loading Rates – The change in a force on a body as a function of time. Increases in loading rate for the vertical ground reaction force are risk factors for knee injuries and tibial stress fractures (Davis et al. 2004; Hewett et al. 2005; Milner et al. 2006).

- Joint Moments – A joint moment is the product of the force acting on a segment and the perpendicular distance from its line of action to the joint's center of rotation. Increases in joint moments are a potential risk factor for injury (Hewett et al. 2005; Milner et al. 2006).

- Joint Reaction Forces (JRF) – The joint reaction force is the net force within the joint, not including muscle forces. Increases in JRFs are a potential risk factor for injury (Hewett et al. 2005; Milner et al. 2006).

- Joint Contact Forces (JCF) – Joint contact forces refers to the net force within the joint including the contribution of muscle forces. Calculation of JCF requires muscle force estimations, which are typically done non-invasively through computer simulations. JCF are a means to compare muscle force contributions or co-contraction patterns stabilizing the joint (Shelburne et al. 2005; Erdemir et al. 2007; Winby et al. 2009). These measures remain difficult to validate empirically due to the difficulty of measuring muscle forces in vivo.

- Joint Powers – Power is the rate at which work is done. Joint powers are calculated as the dot product of the moment about a joint and the angular velocity of the body segment about that joint. The calculation of joint power shows when energy is absorbed or generated at the joint (Eng and Winter 1995) and the flow of energy between segments (Winter 2005).

Current Lab Capability: Utilizing lab-based motion capture systems (optical, video, or IMU-based) in conjunction with instrumented force platforms and simulation software allows for measurement of all relevant kinetic measures on the user alone (i.e., ND condition) or the user-augmentation system combination (i.e., ON condition). Specialized instrumentation (i.e., force and torque sensors) built into the augmentation system and specialized analysis techniques are required to separate the kinetics of the user from the kinetics of the augmentation system. Computer models to accurately separate the kinetics of the user from the kinetics of the augmentation system have not yet been developed and validated.

Current Field Capability: No current, fully validated field capability at this time. Pressure sensing insoles are available as COTS items, but their use to quantify ground reaction forces is limited.

6.2 Physiological Metrics

Physiological performance metrics, for the purposes of this report, are quantitative measures of the physical and chemical phenomena necessary for successful task completion in a biological system of interest. Similar to the biomechanical metrics, physiological metrics stated here will include only those measures that may be helpful in quantifying the behavior and performance of an augmentation system user or the relationship between the user and the system. Physiological-based metrics of interest for physical augmentation system testing may include:

6.2.1 Metabolic Metrics

Metabolic cost is typically quantified by measuring the rate of volume of oxygen consumed ($\dot{V}O_2$) during a given task. Greater oxygen consumption indicates a greater energy expenditure, a greater level of exertion, and increased potential for fatigue. The metabolic cost of walking is strongly related to the mechanical work required for step-to-step transitions (Donelan et al. 2002), and it is thought that individuals select their natural walking pace as that which maximizes metabolic efficiency (Cavagna and Kaneko 1977; Waters and Mulroy 1999). Increases in body mass, load mass, walking speed, grade, and changes in terrain can increase metabolic cost (Knapik et al. 1996; Quesada et al. 2000; Beekley et al. 2007; Fallowfield et al. 2012). Additionally, being fatigued prior to initiating a strenuous task can increase metabolic cost compared to completing the task from a rested state (Ratkevicius et al. 2006). Metabolic metrics may include the following:

- Rate of Oxygen Consumption – refers to the volume of oxygen consumed per minute of activity (typically measured in milliliters of oxygen per minute; $\dot{V}O_2$). Increases in $\dot{V}O_2$ indicate increased physical workload and

may imply heightened risk of fatigue and decreased capacity for task performance.

- Kilocalories Expended – refers to the total amount of calories (or energy) used over the course of a task. Increases may indicate increased physical workload, which would lead to increased nutritional demand and decreased capacity for task performance.
- Cost of Transport (COT) – refers to the amount of energy above resting to move per unit of body mass per unit of distance. An increase in the calculated cost of transport indicates decreased efficiency of movement and increased physical workload.
- Heart Rate (HR) – refers to the number of heart beats per minute of activity. Increases in measured heart rate indicate increased cardiopulmonary demand and physical workload, and may imply heightened risk of fatigue with a decreased capacity for task performance.
- Heart Rate Variability (HRV) – refers to the measure of the variation in the time intervals between successive R-peaks in the QRS complex of the ECG signal. It is also referred to as cycle length variability, or RR variability. Significant alterations in HRV have been found to be related to emotional arousal and are significantly impacted by increased physical activity as well as increased physical and cognitive stress (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology 1996).
- Respiratory Exchange Ratio (RER) – refers to the ratio of carbon dioxide produced to oxygen consumed. It is used to estimate respiratory quotient and the fuel type (fat or carbohydrates) being used during an activity. Carbohydrate usage depends on the intensity of work. RER increases from rest to exercise. At a fixed work load, trained individuals have lower RERs, deplete glycogen more slowly, and can work longer than untrained individuals. A higher RER indicates increased muscular work and can be used as an indicator that a participant is nearing exhaustion during a VO_2 max test.
- VO_2 Max – refers to maximal oxygen consumption, which reflects the aerobic physical fitness of the individual and is an important determinant of their endurance capacity during prolonged, sub-maximal exercise. Higher VO_2 max values indicate higher aerobic physical fitness and are, in general, preferable. VO_2 max can be estimated using a maximal effort 2-mile run time (Mello et al. 1984).

Current Lab Capability: Current cardiopulmonary exercise testing equipment allows for the measurement of all relevant metabolic metrics.

Current Field Capability: Heart rate is easily measured in the field using standard heart rate monitors and associated watches. Portable cardiopulmonary systems can be used in the field, but require Soldiers to wear additional gear that limits their ability to communicate, increases their load, and may limit their movement ability and field of view. VO_2 max testing is conducted in a laboratory. However, a maximal effort 2-mile run time can be used to estimate VO_2 max (Mello et al. 1984).

6.2.2 Muscle Function Metrics – Electromyography (EMG)

Electromyography (EMG) measures the myoelectric signal produced during muscle contraction. Although not directly correlated to muscle force generation, the magnitude of these signals describes the relative level of muscle contraction during a given task. Greater EMG magnitudes indicate more muscle activity and thus a higher level of muscular exertion. EMG signals may also be analyzed for their frequency content. Decreases in median frequency represent a shift to lower contractile frequencies and are used as an indicator of muscle fatigue (De Luca et al. 1983; Hakkinen and Komi 1983; Allison and Fujiwara 2002; Cifrek et al. 2009). Carrying heavy loads can accelerate such changes in these parameters. The effect of backpack load carriage on EMG has been studied in recreational hikers (Simpson et al. 2011) and military personnel (Knapik et al. 1996; Harman et al. 2000; Lindner et al. 2012) and generally demonstrates that muscle activity of the legs, trunk, and shoulders increases with increases in load. Electromyography metrics include the following:

- Peak EMG – refers to the magnitude of the maximum signal produced by an active muscle. Increases in magnitude of muscle activity from baseline conditions may indicate increased muscular work that could result in increased localized muscle fatigue and overall increase in metabolic cost (Hibbs et al. 2011).
- Integrated Muscle Activity (iEMG) – refers to the integral of the rectified EMG-time curve (i.e., the area under the rectified EMG-time curve). iEMG refers to the amount of work a muscle performs during a given activity. Systems that reduce the overall amount of work a muscle needs to perform can be considered preferable.
- Time of Muscle Activity Onset – refers to the timing of the start of muscle activity. Differences in timing of muscle activity onset from baseline

conditions may indicate the adoption of non-optimal movement and motor control patterns, which can increase the risk of injury.

- Frequency Component Analysis – refers to the analysis of the power density spectrum of the measured muscle signal. Shifts to lower mean and median frequencies are indicators of localized muscle fatigue (De Luca et al. 1983; Hakkinen and Komi 1983; Allison and Fujiwara 2002; Cifrek et al. 2009).
- Muscle Synergies and/or Co-Contraction – refers to the analysis of the patterns of muscle activity throughout a task. Changes in neuromuscular control strategies may increase risk of injury and reduce joint stability (Myer et al. 2007).

Current Lab Capability: Current commercially available EMG systems allow for the measurement of all relevant muscle function metrics.

Current Field Capability: The suitability and feasibility of collecting EMG in the field is task dependent. EMG signals are very sensitive to motion artifacts, and electrodes must remain securely attached over the muscle belly for adequate signal quality. High-impact activities or activities in which the EMG electrodes may be compressed under equipment or against obstacles may prohibit the use of EMG data collection systems.

6.2.3 Muscle Function Metrics – Muscle Strength and Endurance

Unlike electromyography-based metrics, measures of muscle strength and endurance directly correlate to muscle force generation. Decreases found in muscle force-producing capability (e.g., after prolonged load carriage) may reduce the ability of individuals to carry out endurance, strength, and skilled tasks and may increase their risk of musculoskeletal injury (Bigland-Ritchie and Woods 1984; Blacker et al. 2010; Clarke et al. 2013). However, the expression of muscular strength is often specific to the test being used to assess it. While isometric peak torque has been the most frequently used metric in investigating muscle strength and endurance of Soldiers, other metrics listed below can also be considered.

- Isometric Peak Torque – refers to the highest torque generated by the muscles about a joint during a maximum contraction when held at a constant joint angle. Decreases in post-task metrics as compared to pre-task metrics indicate muscle or central fatigue and may indicate lower capability of the Soldier to perform their mission. Decreases found in muscle force-producing capability after prolonged load carriage, for example, may reduce the ability of individuals to carry out strength or skilled tasks and increase

their risk of musculoskeletal injury (Blacker et al. 2010, 2013; Grenier et al. 2012).

- Isokinetic Peak Torque – refers to the highest torque generated about a joint by the muscles during a maximum contraction at a constant rate of speed. This type of muscle strength measurement more accurately reflects the contractions of some muscles during locomotion.
- Rate of Torque Development – refers to the derivative of the torque-time curve and reflects the speed at which a muscle can produce its maximal torque. This is often used as a metric of power or "explosiveness" of a muscle/muscle group (Maffiuletti et al. 2016).
- Average Work – refers to the mean work performed by the muscles during a contraction. Decreases in post-task values as compared to pre-task values may indicate muscle fatigue and lower capability of the Soldier to perform their mission (Blacker et al. 2010).
- Vertical Jump Height – is measured as the maximum height reached during a standardized vertical jump task. It is used as a functional assessment of lower extremity power. Decreases in post-task vertical jump height indicate lower extremity muscle fatigue (McGinnis et al. 2016) and may indicate impaired capacity of the Soldier to perform their mission.

Current Lab Capability: Commercially available dynamometers provide the capability to measure joint torques. Jump height can be measured with mechanical systems and optical systems. It can also be estimated using optical systems or a force plate.

Current Field Capability: Limited field capabilities. The vertical jump height test can also be used in the field. However, measurement of isolated muscle groups is limited in field studies. In the field, portable strength measurement systems can be used for more general regional or whole-body isometric strength measurements.

6.2.4 Psychophysiological Metrics

Psychophysiological performance metrics refer to those measures that attempt to quantify the relationship between physiological and psychological phenomena. Typically these measures are subjective and therefore not as robust as other directly measured physiological metrics. However, psychophysiological measures can be useful in understanding user acceptance (or rejection) of an intervention such as a physical augmentation system. They can also be used to determine whether statistically significant differences in biomechanical and physiological measures

are enough to produce an operational outcome that is noticeable by a population of interest such as Soldiers. Psychophysiological metrics may include:

- Rating of Perceived Exertion (RPE) – measures an individual's perceived exertion level. Increases in RPE indicate increased physical workload and may imply heightened risk of fatigue and decreased capacity for task performance (Borg 1970, 1982; Borg and Noble 1974). An example RPE scale is shown in Appendix A.
- Rating of Pain, Soreness, and Discomfort (RPSD) – measures an individual's perceived pain, soreness, and discomfort (at distinct body locations) associated with wearing specific gear and/or while performing a specific task. The RPSD can be used to identify areas where the compatibility between the user and the system needs to be improved (Corlett and Bishop 1976). The original diagram of body regions that Corlett and Bishop used has been recently modified into the RPSD Questionnaire shown in Appendix B. This questionnaire includes space for subjects to record more details about the source of their pain, soreness, or discomfort.
- Comfort – refers to the degree of freedom from pain or constriction within the physical augmentation system. A system should be usable for the duration of a mission without causing discomfort, hot spots, chafing, and so on, based on defined mission length. Comfort level should be maintained whether the system is operating or idle. The system should be as transparent as possible to the user. Users should not feel as though they are fighting the system while it is either operating or idle. In addition to questionnaires and interviews, the Comfort Affected Labeled Magnitude (CALM) Scale (Cardello 2003) is frequently used to assess comfort or discomfort with systems. The CALM Scale is shown in Appendix C.

Current Lab Capability: Current questionnaires, forms, and other means allow for the collection of relevant psychophysiological metrics.

Current Field Capability: Current questionnaires, forms, and other means allow for the collection of relevant psychophysiological metrics.

6.3 Operational Metrics

Operational metrics are field-based measures that quantify performance of Soldier-relevant tasks in the general categories of strength, movement, and marksmanship. Typically, Soldier tasks are simulated and performance on each task is defined based on how these tasks would be successfully completed within an actual combat

scenario. Simulated operational tasks and their associated performance metrics for physical augmentation system testing may include the following:

6.3.1 Strength Metric

This metric is intended to provide an assessment of whether a Soldier has the strength to do a task that simulates the kind of lifting tasks done in a combat environment. This task is different from the Muscle Function Metrics – Muscle Strength and Endurance of Section 6.2. This metric assesses large groups of muscles working together to accomplish a task rather than examining isolated muscles or muscle groups. Also, this metric can be measured in a lab or field environment whereas most of the metrics in Section 6.2 are measured in the lab.

- Box Lift – A box lift task consists of performing repeated lifts of a weighted box to a pre-determined height, then lowering the box back to the ground, and repeating the task to exhaustion, usually in time with a metronome. Sometimes the weighted box is carried a specified distance before it is placed back on the ground or lifted to a pre-determined height. Box lift tasks have been shown to be good predictors of performance of military tasks such as a rucksack lift and place, artillery gunner loading simulation, and bridge building simulations (Carstairs et al. 2016). Performance metrics include overall time to complete a certain number of lifts (or lift-and-carries) or number of repetitions in a given time. Higher numbers of repetitions or longer times before falling off pace are indicators of higher strength and endurance. Decreased numbers of repetitions or shorter times before falling off pace indicate decreased performance and earlier onset of fatigue.

Current Lab Capability: Using commonly available materials, this strength metric can easily be measured in laboratory settings.

Current Field Capability: Using commonly available materials, this strength metric can also be measured in field settings.

6.3.2 Movement Metrics

These metrics are used to quantify performance on a variety of tasks that simulate the types of movements Soldiers might make in a combat scenario. These metrics are adapted from tests that assess Soldier mobility with respect to the clothing worn and equipment carried and tests that assess equipment portability. A physical augmentation system can be considered another piece of Soldier clothing or equipment. The movement metrics are measures of time to complete a task or distance traveled, so they are relatively easy to assess.

- Preliminary Mobility – These metrics are general checks on the system's basic performance related to operationally-based mobility. They are most important in the initial stages of development, prior to assessment of the remaining operational metrics. If performance under these metrics is poor, changes to system design and/or performance are likely prior to assessment of the remaining operational metrics. These metrics are assessed by the evaluator through direct observation of participant performance (as pass/fail) and/or through participant feedback (i.e., questionnaires, interviews, or other means), typically as a rating of interference or restriction. Examples of these types of tasks include the following:
    - Sitting: No components should restrict the ability to sit down, nor should any components cause discomfort while seated with or without back support.
    - Kneeling/Squatting/Crawling: A system should not impede the ability to squat on one's heels, to kneel, or crawl, nor should any of the components cause discomfort when doing so.
    - Assuming a Weapon Firing Position: User should be able to assume firing positions. A system should not impede the user from engaging targets, achieving a shoulder/cheek weld, and firing in a standing, kneeling, or (if appropriate to concept of use) prone position.
    - Running: Short sprints should be possible without system interference or restriction.
    - Walking uphill/downhill: The system should allow these tasks without restriction.
    - Obstacle clearance: Users should be able to maintain their ability to climb over obstacles (e.g., logs, holes, low walls), as well as climb stairs, and ladders.
    - Balance maintained: A system should have no negative impact on a user's balance.
    - Modification to movement or gait: The system should not modify movement or gait in a way that increases the potential for injury.
- Obstacle Course – Obstacle courses are designed to simulate relevant military tasks and can be used as assessment tools to determine differences in Soldier performance. Overall time to complete a course and time to complete individual sections of a course are the primary outcome metrics used to quantify performance and agility. Courses are often used to assess Soldier fitness as well as Soldier mobility as a function of the clothing and equipment worn (Tack et al. 2012; Bossi et al. 2014; Brewster 2014; Dutton

and Stryker 2015). They can also be used to assess the portability of equipment carried by Soldiers (Mitchell et al. 2016 and Batty et al. 2016).

- Agility Run – Agility runs are similar to obstacle courses in that they can be used to assess Soldier fitness or Soldier mobility as a function of the clothing and equipment worn. Total time to complete an agility run is the primary outcome metric used to quantify performance and agility. Recent developments in wearable sensors such as IMUs have also allowed for the measurement of the following metrics: turn speed, turn radius, heading angles, tangential acceleration range, angular velocities of segments, straightaway body speed, straightaway distance, cord length at turn, number of footfalls during a turn/straightaway, and cumulative ground contact duration during turns and straightaways (McGinnis et al. 2016, 2017). Significant deviations in these measures from a baseline condition may help to explain significant differences in task performance.

- Balance Beam – Balance beam tasks can be used as a functional assessment of Soldier balance and mobility. Time to traverse the beam and number of falls from the beam are the primary outcome metrics used to quantify performance. Recent developments in wearable sensors such as IMUs have also allowed for the measurement of the following metrics: foot used to push onto beam, foot used to land off beam, number of steps on beam, double support time on beam, step frequency, stride duration, medial-lateral and anterior-posterior acceleration and angular velocity of sacrum, torso medial-lateral range of motion, torso and rifle angular velocities, average time to complete turns of the beam, torso lean (Cain et al. 2016). Significant deviations in these measures from a baseline condition may help to explain significant differences in overall task performance.

- Window Traversal – Window traversal tasks (i.e., climbing through a window-size opening) can be used as an assessment of Soldier mobility as well as the function of clothing and equipment worn. They can also be used to assess the portability of equipment carried by Soldiers. Time to complete the traversal is the primary outcome metric used to quantify performance. Recent developments in wearable sensors such as IMUs have also allowed for the measurement of the following metrics: peak approach speed, peak departure speed, velocity ratio, vertical takeoff speed, takeoff work, average takeoff power, duration of takeoff, vertical landing speed, landing work, average landing power, duration of landing, max pelvis heading angle, range of pelvis heading angle, and range of body twist angle (Cain et al. 2015). Significant deviations in these measures from a baseline condition may help to explain differences in overall task performance.

- Road March – Road marches are one of the most common dismounted infantry tasks and involve walking with required military gear for a specified time or distance. Performance metrics include the total time to complete the march and the rate of march (walking speed). Longer times to complete a march and slower walking speeds indicate reduced Soldier performance and increased fatigue (Knapik et al. 1993).

- Traverse Natural Terrain – Dismounted Soldiers must often walk or run through natural terrain. If a particular physical augmentation system is intended to be used in natural terrain, then it is important to have an area with natural terrain for subjects to traverse. Depending on the technical maturity of the system, the path through the natural terrain can range from a worn trail clear of obstacles (for systems of low technical maturity) to an unmarked path with obstacles that subjects must navigate using a map and compass and/or GPS (for systems of higher technical maturity). Performance traversing natural terrain is typically measured by course completion time, or distance traveled in a specified time, and metabolic metrics (Section 6.2.1). Significant deviations in these measures from a baseline condition may help to explain differences in overall task performance.

- Specific Mission-Based Operational Tasks – This category encompasses any specialized tasks that a user might be expected to conduct with a physical augmentation system. The particular tasks will depend on the role a Soldier is playing, the military occupational specialty (MOS) of the user, and other factors. These tasks also involve actions or activities beyond those specifically addressed by other movement metrics detailed in this section. Assessment of these specific tasks is accomplished through creation of a mock scenario or mission incorporating them. By running a mock mission, system capabilities and shortcomings in a dynamic environment can be determined and documented. Specific operational tasks will be most useful somewhat later in system development, once other movement metrics have been used to assess and optimize performance. These specific tasks/scenarios are assessed through a variety of means. These metrics are assessed by the evaluator through direct observation of participant performance (as pass/fail), timed performance versus published mission performance standards or other means appropriate to the task/scenario, and/or through participant feedback (questionnaires, interviews, or other means), typically as a rating of interference or restriction.

Current Lab Capability: Current lab capabilities can accommodate the measurement of most of these movement metrics; however, the operational setting itself is of interest and therefore a non-lab setting for many of these tasks would be preferred.

Current Field Capability: Current field capabilities allow for the measurement of all relevant movement metrics.

6.3.3 Marksmanship Metrics

Marksmanship is an important skill for Soldiers to possess. The performance measures are relatively easy to collect in a marksmanship simulator or at a live-fire range. It must be noted, however, that performance in a simulator and live-fire performance are not the same. Therefore, simulator performance should only be compared to simulator performance, and live-fire performance should only be compared to live-fire performance.

- Simulated Marksmanship – Quantifying marksmanship ability is critical when determining the effects of an intervention on Soldier operational performance. Simulated marksmanship allows measures of accuracy, precision, and motor control without the use of live ammunition or the need for live-fire shooting facilities. Performance metrics include precision, accuracy, probability of hit, probability of lethal hit, aiming time, stability (horizontal, vertical, and overall), trigger control, time between shots, mean target acquisition time, total target acquisition time, engagement time, and total scenario completion time. Clothing and individual equipment have been shown to impact timing and stability metrics (Choi et al. 2016; Brown et al. 2017a, 2017b). Significant decreases in any of these performance metrics may indicate impaired marksmanship ability and reduced lethality.
- Live-Fire Marksmanship – Marksmanship is one of the most critical tasks a Soldier performs. The ability to measure marksmanship is critical when determining the effects of an intervention on Soldier operational performance. Live-fire marksmanship permits the same accuracy, precision, and motor control assessment as simulated marksmanship, but it presents a more realistic engagement scenario to the Soldier. Live-fire also permits the quantification of recoil forces and their effects on accuracy, precision, fatigue, shot timing, and target acquisition. Further, live-fire can be used to assess the compatibility between military equipment, weapons, optics, muzzle devices, or other barrel-borne accessories. Similar to simulated marksmanship, the following metrics can be used to assess performance: precision, accuracy, probability of hit, probability of lethal hit, aiming time,

stability (horizontal, vertical, and overall), trigger control, time between shots, mean target acquisition time, total target acquisition time, engagement time, and total scenario completion time. Significant decreases in any of these performance metrics may indicate impaired marksmanship and reduced lethality.

Current Lab Capability: Current lab capabilities allow for the measurement of marksmanship using simulators only.

Current Field Capability: Portable marksmanship simulators can be used in field settings to measure simulated marksmanship, and designated ranges are needed for measuring live-fire marksmanship.

6.4 Human Factors Metrics

Human factors metrics are measures used to understand the interactions among humans and a system including its environment in order to optimize functionality, human comfort, safety, and overall system performance. These measures can be either qualitative (both user opinions and tester observations) or quantitative (e.g., time to complete a task or noise level measurements). Both types of measures can be useful in understanding user acceptance and success (or rejection) of an intervention such as a physical augmentation system. This includes understanding user opinions, what the driving needs of the user are, how to improve the form and design of the system for easy use by all stakeholders (including those beyond the primary users such as logisticians, maintenance personnel, etc.). Guidelines and recommendations for human factors engineering of military devices can be found in MIL-STD-1472 (MIL-STD-1472G 2012). Human factors metrics of interest for physical augmentation system testing may include the following.

6.4.1 Fitting and Sizing Metrics

These metrics are closely linked, and they are used to assess a system's means of attachment to and position on the body as well as the range of body dimensions a system should accommodate. System fitting should be easy to understand and accomplish, with a minimum of pieces and fasteners. Early prototypes may require assistance for donning the system, but mature systems should permit self-donning (if within the needs and goals of the program) with only minor adjustments made by a partner if necessary. Systems (including all subcomponents/pieces, fasteners, straps, etc.) should fit a broad range of wearers (via adjustability or alternative sizes), consistent with dimensions in Anthropometric Survey (ANSUR) II database (Gordon et al. 2014).

Fitting and sizing metrics include the following:

- Size and Body Dimensions – refers to the size(s) of the system relative to the specific body areas the system must accommodate. Metrics include the sizing dimensions the system is designed to accommodate and the number of sizes necessary for accommodation. A system should fit a wide range (typically a minimum of 90% [threshold] up to 98% [objective]) of the Soldier population, both males and females, consistent with dimensions in the ANSUR II database (Gordon et al. 2014). This accommodation range is critical whether a system is designed to fit a broad range of sizes (i.e., via adjustments and/or other changes) or is custom manufactured. If the system is custom manufactured for a specific Soldier, the developer must have the ability to produce systems in the sizes needed to accommodate the Soldier population (typically 90% to 98% of male and female Soldiers). Assessment of sizing is conducted by measurement and fitting trials and observation by a researcher trained in measurement techniques (Gordon et al. 2014).
- Donning and Doffing – refers to the ability of the user to put on or remove the system. Metrics include: Self-don/Self-doff Capability (able to be donned by the wearer with no assistance or a minimum of assistance in a reasonable amount of time based on requirements), Self-Adjustment Capability (no or minimal assistance should be needed to make adjustments once system is donned), and Quick Doff Capability (system should allow for quick doffing for emergency purposes without tools). These metrics are assessed by the researcher through heuristic and participant trial observation. Feedback from evaluation participants (via questionnaires, interviews, etc.) is also recorded.

6.4.2 Usability Metrics

These metrics are used to determine whether a system is field capable. These metrics will be dependent on the operational context in which a system is expected to perform. Note that many usability metrics are related to system design. System attributes and features should be chosen at the outset with the knowledge that they are likely to be interrelated. Therefore, system designs should be undertaken so that features and performance choices do not preclude modification later. Modifications may be desired or required for a host of reasons, to include improving performance, user acceptance, usability, maintenance, reparability, or other system attributes at a later time. Common usability metrics include ease of use, safety, and signature.

Usability metrics include the following:

- Ease of Use – refers to the ease with which the system is operated and maintained. Ease of use can be assessed through determination of accessibility, system feedback, storage, and repairs. For accessibility, Soldiers should have easy access to all main and emergency power switches, to ensure that the user can control power on/off. All parts/connections/cables and so forth should be easy to reach, connect, and disconnect with no or minimal assistance. For system feedback, the system should have a method to inform Soldiers what mode it is currently in, if applicable. For storage, the system should be easy to disassemble (if necessary) and easy to stow. For repairs, simple repairs by a Soldier in the field should be possible and field-expedient. All of these metrics are assessed by the researcher both heuristically and via participant trials. All observations and participant feedback are recorded. Participants can offer feedback through questionnaires, interviews, or other appropriate means.

- Safety – refers to the safety of the system, which can be assessed by evaluating the following: snag hazards, quick release capability, quick-doff capability, emergency shut-off capability, fail-safe capability (if system failure occurs, the operator should be able to continue the mission without having to doff the system and without interference), thermal burden, and the potential for either physical or thermal injury (burns). Details regarding military safety standards can be found in MIL-STD-882 (MIL-STD-882E 2012). These metrics are assessed by the researcher/subject matter expert through heuristic and user/participant observation. All observations are noted and recorded. Participant information can be obtained through questionnaires, interviews, or other feedback methods.

- Signature – refers to the distinctive pattern, product, or characteristic by which a system can be identified. Military systems that are not easily identified by enemy combatants are preferable. The system signature can be evaluated by measuring the noise level, light level, electronic signature, heat signature, and visual signature to determine if they exceed the levels to which the system was designed or limits such as those in MIL-STD-1472 (MIL-STD-1472G 2012).

- Reliability – refers to the system's ability to perform its intended function under defined conditions. Reliability can be measured in a number of ways. Typically, the mean time between failures (MTBF) is calculated to give a measure or reliability. Other calculations involving MTBF, failure rates,

and probability of surviving for a specified operating time are also used to quantify reliability (MIL-HDBK-781A 1996).

6.4.3 Equipment Compatibility Metrics

These metrics are used to determine the degree to which a system integrates with other, standard Soldier equipment systems and with the Soldier system as a whole. Equipment compatibility metrics may include (but are not limited to) the following:

- Weight and Weight Distribution – refers to the load that the Soldier carries or wears and how it is distributed on the body. Given the current state of Soldier overburden, lightweight systems that do not add significantly to the load the Soldier bears are preferable. System weight should be offset as much as possible by the system benefit. Weight should be distributed on the Soldier so that it has minimal effect on the ability to move and perform required tasks, and so that it maximizes comfort. These metrics can be assessed by direct measurement of the weight of the system, and via questionnaires or other means. Participants can be asked to note body areas where the weight they are bearing causes any pain, soreness, or discomfort. In addition, weight and weight distribution assessment results should be reviewed in the context of the biomechanical and mobility assessment results.

- Mobility – refers here to the ability to move or be moved freely and easily in conjunction with the clothing and equipment with which the system is expected to function. The metrics are the same as those previously described under range of motion (Section 6.1 in the Kinematic Metrics) and movement metrics (Section 6.3). They are assessed as "pass-fail" by evaluator observation or through restriction and/or interference subjective ratings through use of Likert scales (Likert 1932). They include torso rotation, bending at the waist, joint rotations (at the hips, knees, and ankles), sitting, kneeling/squatting/crawling, weapon firing positions, running, walking up/downhill, obstacle clearance (climbing over logs/holes, climbing stairs/ladders), maintaining balance, unmodified gait.

- Compatibility – refers to the degree to which a system interferes with the normal functioning of other clothing and equipment worn by the Soldier. Compatibility can be assessed with worn or carried clothing and equipment, and with other military platforms. For worn or carried clothing and equipment, the system should be compatible with body armor, rucksacks, duty/combat uniforms, boots and other items used by Soldiers. Any required modifications to standard-issued clothing and individual equipment should

be very limited or unnecessary. For other military platforms (i.e., vehicles and aircraft), a system should not interfere with egress/ingress, vehicle operation, transport/sitting, or maintenance procedures as appropriate to the system's concept of use. All of these metrics are assessed by the researcher by direct observation of participants using the system with the relevant items. Observations are noted and recorded. Typically these metrics are assessed as "pass/fail" or "go/no-go." In addition, subjective ratings of restriction and/or interference from participants can be elicited via Likert scales through various means (questionnaires, interviews, etc.).

6.4.4 Training Metrics

Refers to the degree and types of instruction required for successful operation of a system. Questionnaires (with 5 point Likert scales) can be used to probe users for understanding and instructional issues. These metrics are human factors best practices, and may not have a standard method for evaluation. Many can also be assessed relative to a previous iteration of a system to examine progress. Training metrics may include (but are not limited to) the following:

- Habituation Time – The amount of time required for a Soldier to become comfortable using a system. Soldiers should become comfortable with a system in a reasonably short time without causing pain or excessive muscle fatigue. The specific amount of habituation time needed will vary with the complexity of the device and each individual, but should improve with system use. Time in the device should be tracked along with any pain, discomfort, or muscle fatigue issues. The elapsed time in the system should be recorded as these issues occur or at regular intervals. If the issues do not improve, adjustments to the system (e.g., fit adjustment or level of assistance) may be necessary to avoid injury to the user. This time can be compared across successive system iterations to assess whether increases in habituation time have occurred.

- Instruction Time – The amount of time that is required for instruction of proper system usage and familiarization. Orientation and instruction should not take longer than new equipment training for similar items. System design should be mindful of the instruction time required and work toward ease of user understanding. Soldiers trained on the system should be asked whether the instruction time was reasonable for the amount of information needed to convey system operation to them. One method of assessment is to ask for feedback at the end of instruction.

- Familiarization – Users should be able to complete the task or demonstrate an understanding of the required task that is to be augmented prior to completing the task with the system. Users should be queried to ensure that they are familiar with the system; this is a safety concern as well as a training issue.
- Understandability – Users should be able to understand system use and operation without overly extensive, specific instructional training. At the completion of training, instructors may query users on the main points of system use and operation. In that case, trainees should demonstrate understanding of the system via answers to a brief list of orally administered questions.
- Physical Fitness Requirements – The physical fitness requirements should not be greater than the fitness levels of the types of Soldiers who will use the augmentation system. Each MOS in the Army has defined physical demands (Department of the Army 2007). These physical demands are often defined in terms of weight and frequency of lifts required for that MOS. MIL-STD-1472 (MIL-STD-1472G 2012) also provides guidance on limits for carried weight, lifts, and other physical strength attributes. A system's physical fitness requirements should be compared to those of the Soldiers in the MOS expected to use it and the limits in the MIL-STD-1472.

Current Lab Capability: Most of the human factors metrics are assessed via observations, questionnaires, and interviews, which can be done in the lab. Measuring the signatures of an augmentation system requires instruments that are also available in the laboratories.

Current Field Capability: The measurement of all relevant human factors metrics can also be done in the field through observations, questionnaires, interviews, and portable instruments for measuring signatures.

7. Conclusion

This report documents physical augmentation technology terminology, protocol development considerations, metrics, and other related best practices. It covers quantitative and qualitative biomechanical, physiological, operational, and human factors metrics. The selection of the appropriate metrics to use during an evaluation requires judgment, and key selection considerations include the evaluation goals, intended use, intended environment, technical maturity of the candidate system, available resources, safety, and training. Table 1 provides high-level suggestions to orient the reader based on technical system maturity.

The practices detailed throughout this report capture lessons learned, which were originally documented for internal NSRDEC and ARL-HRED use; however, additional details have been added to clarify Army applications and interest, technology and military terminology, evaluation processes, and desirable metric outcomes for a broader audience. In effect, this report supports discussion of standard military physical augmentation technology methods and metrics, which may be further considered by the broader community encompassing numerous non-military stakeholders. These methods and metrics themselves are not standards; however, similar to standards, they may enhance future activities intended to foster physical augmentation technological innovation while improving technology quality, safety, competition, and user confidence.

Due to the rapidly evolving nature of the physical augmentation technology space, it is expected that new technology designs and applications will emerge. Similarly, it is expected that new measurement metrics and methods, as well as enabling human and operational performance measurement technologies will emerge and mature. This report is not an all-encompassing blueprint, but rather is intended to provide a starting point from which to tailor your organization's unique physical augmentation technology evaluations. As such, depending on your organization's needs and the system's technical maturity to accomplish its intended use, there is considerable flexibility in which subsets of the methods and metrics outlined in this report may be adopted.

8. References

Ali N, Robertson D, Rouhi G. Sagittal plane body kinematics and kinetics during single-leg landing from increasing vertical heights and horizontal distances: implications for risk of non-contact ACL injury. Knee. 2012;21(1):38–46.

Allen C, Wong E, Livesay G, Sakane M, Fu F, Woo S. Importance of the medial meniscus in the anterior cruciate ligament-deficient knee. J Orthop Res. 2000;18(1):109–115.

Allison G, Fujiwara T. The relationship between EMG median frequency and low frequency band amplitude changes at different levels of muscle capacity. Clin Biomech (Bristol, Avon). 2002;17(6):464–469.

Bachkosky J, Andrews M, Douglass R, Feigley J, Felton L, Fernandez F, Rumpf R. Lightening the load. No. NRAC-07-02. Arlington (VA): Naval Research Advisory Committee; 2007.

Batty J, Coyne M, DeSimone L, Mitchell K, Bensel C. Evaluation of weight effects on a soldier physical readiness test course. Proceedings of the 2016 American Biomechanics Society; 2017; Raleigh, NC.

Beekley M, Alt J, Buckley C, Duffey M, Crowder T. Effects of heavy load carriage during constant-speed, simulated, road marching. Mil Med. 2007;172(6):592–595.

Begg R, Best R, Dell'Oro L, Taylor S. Minimum foot clearance during walking: strategies for the minimization of trip-related falls. Gait Posture. 2007;25(2):191–198.

Bennell K, Goldie P. The differential effects of external ankle support on postural control. J Orthop Sports Phys Ther. 1994;20(6):287–295.

Bigland-Ritchie B, Woods J. Changes in muscle contractile properties and neural control during human muscular fatigue. Muscle Nerve. 1984;7:697–699.

Blackburn J, Padua D. Landing, influence of trunk flexion on hip and knee joint kinematics during a controlled drop. Clin Biomech (Bristol, Avon). 2008;23(3):313–319.

Blacker S, Fallowfield J, Bilzon J, Willems M. Neuromuscular function following prolonged load carriage on level and downhill gradients. Aviat Space Envir Md. 2010;81(8):745–753.

Blacker S, Fallowfield J, Bilzon J, Willems M. Neuromuscular impairment following backpack load carriage. J Hum Kinet. 2013;37:91–98.

Boling M, Padua D, Marshall S, Guskiewicz K, Pyne S, Beutler A. A prospective investigation of biomechanical risk factors for patellofemoral pain syndrome. Am J Sports Med. 2009;37(11):2108–2116.

Borg G. Perceived exertion as an indicator of somatic stress. Scand J Rehabil Med. 1970;2:92–98.

Borg G. Psychophysical bases of perceived exertion. Med Sci Sport Exer. 1982;14(5):377–81.

Borg G, Noble B. Perceived exertion. Exerc Sport Sci Rev. 1974;2:131–53.

Bossi L, Kelly A, Wojtarowicz D, Jones M, DuCharme M. Load Effects Assessment Program (LEAP): Sensitivity of LEAP to operationally-relevant clothing and equipment conditions. 3rd International Congress on Soldier's Physical Performance; 2014; Boston, MA.

Brewster F. Observation report for the Load Effects Assessment Program – Army (LEAP-A) pilot evaluation. Columbus (GA): Maneuver Battle Lab, Maneuver Center of Excellence; 2014.

Brown C, Mynark R. Balance deficits in recreational athletes with chronic ankle instability. J Athl Training. 2007;42(3):367–373.

Brown S, McNamara J, Mitchell K. Dynamic marksmanship: a novel methodology to evaluate the effects of clothing and individual equipment on mission performance. Proceedings of the Human Factors and Ergonomics Society 61th Annual Meeting; 2017; Austin, TX.

Brown S, McNamara J, Mitchell K. Shooting stability: a critical component of marksmanship performance as measured through aim path and trigger control. Proceedings of the Human Factors and Ergonomics Society 61st Annual Meeting; 2017; Austin, TX.

Browning R, Kram R. Effects of obesity on the biomechanics of walking at different speeds. Med Sci Sports Exerc. 2007;39(9):1632–1641.

Cain S, McGinnis R, Davidson S, Vitali R, McLean S, Perkins N. Quantifying performance and effects of load carriage during completion of a window obstacle using an array of wireless inertial sensors. Proceedings of the American Society of Biomechanics; 2015; Columbus, OH.

Cain S, McGinnis R, Davidson S, Vitali R, Perkins N, McLean S. Quantifying performance and effects of load carriage during a challenging balancing task using an array of wireless inertial sensors. Gait Posture. 2016;43:65–69.

Cardello A, Winterhalter C, Schutz H. Predicting the handle and comfort of military clothing fabrics from sensory and instrumental data: development and application of new psychophysical methods. Text Res J. 2001; 73(3):221–237.

Carstairs G, Ham D, Savage R, Best S, Beck B, Doyle T. A box lift and place assessment is related to performance of several military manual handling tasks. Mil Med. 2016;181(3):258–64.

Cavagna G, Kaneko M. Mechanical work and efficiency in level walking and running. J Physiol. 1977;268(2):467–481.

Choi H, Mitchell K, Garlie T, McNamara J, Hennessy E, Carson J. Effects of body armor fit on marksmanship performance. Advances in Physical Ergonomics and Human Factors. 2016;489:341–354.

Cifrek M, Medved V, Tonković S, Ostojić S. Surface EMG based muscle fatigue evaluation in biomechanics. Clin Biomech (Bristol, Avon). 2009;24(4):327–340.

Clarke H, Shay C, Mathews D. Strength decrements from carrying various Army packs on military marches. Res Q Exercise Sport. 2013;26(3):253–265.

Corlett E, Bishop R. A technique for assessing postural discomfort. Ergonomics. 1976;19(2):175–182.

Corriveau H, Prince F, Hebert R, Raiche M, Tessier D, Maheux P, Ardilouze J. Evaluation of postural stability in elderly with diabetic neuropathy. Diabetes Care. 2000;23:1187–1191.

Corriveau H, Hebert R, Prince F, Raiche M. Postural control in the elderly: an analysis of the test-retest and interrater reliability of the COP-COM variable. Arch Phys Med Rehabil. 2001;82:80–85.

Crowell H, Milner C, Hamill J, Davis I. Reducing impact loading during running with the use of real-time visual feedback. J Orthop Sports Phys Ther. 2010;40(4):206–213.

Davis I, Milner C, Hamill J. Does increased loading during running lead to tibial stress fractures? A prospective study. Med Sci Sports Exerc. 2004;36(5):S58.

De Luca C, Adam A, Wotiz R, Gilmore L, Nawab S. Decomposition of surface EMG signals. J Neurophysiol. 2006;96(3):1646–1657.

Decker M, Tokish J, Ellis H, Torry M, Hawkins R. Subscapularis muscle activity during selected rehabilitation exercises. Am J Sports Med. 2003;31(1):126–134.

Decker M, Torry M, Wyland D, Sterett W, Steadman J. Gender differences in lower extremity kinematics, kinetics and energy absorption during landing. Clin Biomech. 2003;18:662–669.

Department of the Army. Pamphlet 611-21: military occupational classification and structure. Washington (DC): Department of the Army; 2007.

Dickstein R, Abulaffio N. Postural sway of the affected and nonaffected pelvis and leg in stance of hemiparetic patients. Arch Phys Med Rehabil. 2000;81(3):364–367.

Donelan J, Kram R, Kuo A. Mechanical and metabolic determinants of the preferred step width in human walking. Proc Biol Sci. 2001;268(1480):198–1992.

Donelan J, Kram R, Kuo A. Mechanical work for step-to-step transitions is a major determinant of the metabolic cost of human walking. J Exp Biol. 2002;205(23):3717–3727.

Donelan J, Shipman D, Kram R, Kuo A. Mechanical and metabolic requirements for active lateral stabilization in human walking. J Biomech. 2004;37(6):827–835.

Dutton B, Stryker T. Soldier load benchmark evaluation. Columbus (GA): Maneuver Battle Lab, Maneuver Center of Excellence; 2015.

Eng J, Winter D. Kinetic analysis of the lower limbs during walking: What information can be gained from a three-dimensional model? J Biomech. 1995;28(6):753–758.

Erdemir A, McLean S, Herzog W, van den Bogert A. Model-based estimation of muscle forces exerted during movements. Clin Biomech. 2007;22:131–154.

Fallowfield J, Blacker S, Willems M, Davey T, Layden J. Neuromuscular and cardiovascular responses of Royal Marine recruits to load carriage in the field. Appl Ergon. 2012;43(6):1131–1137.

Ford K, Myer G, Hewett T. Valgus knee motion during landing in high school female and male basketball players. Med Sci Sports Exerc. 2003;35(10):1745–1750.

Gehring D, Melnyk M, Gollhofer A. Gender and fatigue have influence on knee joint control strategies during landing. Clin Biomech. 2009;24:82–87.

Gök H, Ergin S, Yavuzer G. Kinetic and kinematic characteristics of gait in patients with medial knee arthrosis. Acta Orthop Scand. 2002;73(6):647–652.

Gordon C, Blackwell C, Bradtmiller B, Parham J, Barrientos P, Paquette S, Corner B, Carson J, Venezia J, Rockwell B, Mucher M. 2012 anthropometric survey of US Army personnel: methods and summary statistics. 2014.

Grenier J, Millet G, Peyrot N, Samozino P, Oullion R, Messonnier L, Morin J. Effects of extreme-duration heavy load carriage on neuromuscular function and locomotion: A military-based study. PLoS One. 2012;7(8):e43586.

Guskiewicz K, Perrin D. Research and clinical applications of assessing balance. J Sport Rehabil. 1996;5(1):45–63.

Guskiewicz K, Ross S, Marshall S. Postural stability and neuropsychological deficits after concussion in collegiate athletes. J Athl Training. 2001;36(3):263–73.

Gutekunst D, Frykman P, Seay J. Over-striding during "fixed-cadence" load carriage leads to increased ground reaction forces. Annual Meeting of the American Society of Biomechanics; 2010; Providence, RI.

Häkkinen K, Komi P. Electromyographic changes during strength training and detraining. Med Sci Sports Exerc. 1983;15(6):455–460.

Harman E, Hoon K, Frykman P, Pandorf C. The effects of backpack weight on the biomechanics of load carriage. Military Performance Division. Natick (MA): Army Research Institute of Environmental Medicine (US); 2000.

Heiderscheit B, Chumanov E, Michalski M, Wille C, Ryan M. Effects of step rate manipulation on joint mechanics during running. Med Sci Sports Exerc. 2011;43(2):296–302.

Heller M, Challis J, Sharkey N. Changes in postural sway as a consequence of wearing a military backpack. Gait Posture. 2009;30(1):115–117.

Hewett T, Myer G, Ford K, Colosimo A, LcLean S, van den Bogert A, Paterno M, Succop P. Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: a prospective study. Am J Sports Med. 2005;33(4):492–501.

Hibbs A, Thompson K, French D, Hodgson D, Spears I. Peak and average rectified EMG measures: Which method of data reduction should be used for assessing core training exercises? J Electromyogr Kinesiol. 2011;21(1):102–111.

Hills A, Parker A. Gait characteristics of obese children. Arch Phys Med Rehabil. 1991;72:403–407.

Holbein M, Chaffin D. Stability limits in extreme postures: Effects of load positioning, foot placement, and strength. Hum Factors. 1997;39:456–468.

Hollman J, McDade E, Petersen R. Normative spatiotemporal gait parameters in older adults. Gait Posture. 2011;34:111–118.

Jones B, Knapik J. Physical training and exercise-related injuries. Surveillance, research and injury prevention in military populations. Sports Med. 1999;27(2):111–125.

Jones B, Bovee M, Harris J, Cowan D. Intrinsic risk factors for exercise-related injuries among male and female Army trainees. Am J Sports Med. 1993;21(5):705–710.

Kang H, Dingwell J. Effects of walking speed, strength and range of motion on gait stability in healthy older adults. J Biomech. 2008;41:2899–2905.

Karlsson A, Frykberg G. Correlations between force plate measures for assessment of balance. Clin Biomech (Bristol, Avon). 2000;15(5):365–369.

Knapik J, Ang P, Reynolds K, Jones B. Physical fitness, age, and injury incidence in infantry soldiers. J Occup Med. 1993;35(6):598–603.

Knapik J, Harman E, Reynolds K. Load carriage using packs: A review of physiological, biomechanical and medical aspects. Appl Ergon. 1996;27(3):207-216.

Knapik J, Reynolds K, Harman E. Soldier load carriage: historical, physiological, biomechanical, and medical aspects. Mil Med. 2004;169(1):45–56.

Knapik J, Jones S, Darakjy S, Hauret K, Bullock S, Sharp M, Jones B. Injury rates and injury risk factors among US Army wheel vehicle mechanics. Mil Med. 2007;172(9):988–996.

Konradsen L. Sensori-motor control of the uninjured and injured human ankle. J Electromyogr Kinesiol. 2002;12(3):199–203.

Kristianslund E, Bahr R, Krosshaug T. Kinematics and kinetics of an accidental lateral ankle sprain. J Biomech. 2011;44(14):2576–2578.

Le Clair K, Riach C. Postural stability measures: what to measure and for how long. Clin Biomech (Bristol, Avon). 1996;11(3):176–178.

Ledin T, Odkvist L. Effects of increased inertial load in dynamic and randomized perturbed posturography. Acta Otolaryngol. 1993;113:249–252.

Lee T, Lee Y. An investigation of stability limits while holding a load. Ergonomics. 2003;46:446–454.

Lenhart R, Thelen D, Wille C, Chumanov E, Heiderscheit B. Increasing running step rate reduces patellofemoral joint forces. Med Sci Sports Exerc. 2014;46(3):557–564.

Likert R. A technique for the measurement of attitudes. Arch Psychol. 1932;140:1–55.

Lindner T, Schulze C, Woitge S, Finze S, Mittelmeier W, Bader R. The effect of the weight of equipment on muscle activity of the lower extremity in Soldiers. Sci World J. 2012;2012:976513.

Liu K, Glutting J, Wikstrom E, Gustavsen G, Royer T, Kaminski T. Examining the diagnostic accuracy of dynamic postural stability measures in differentiating among ankle instability status. Clin Biomech. 2013;28(2):211–217.

Maffiuletti N, Aagaard P, Blazevich A, Folland J, Tillin N, Duchateau J. Rate of force development: Physiological and methodological considerations. Eur J Appl Physiol. 2016;116(6):1091–1116.

Malinzak R, Colby S, Kirkendall D, Yu B, Garrett W. A comparison of knee joint motion patterns between men and women in selected athletic tasks. Clin Biomech (Bristol, Avon). 2001;16(5):438–445.

Malliaras P, Cook J, Kent P. Reduced ankle dorsiflexion range may increase the risk of patellar tendon injury among volleyball players. J Sci Med Sport. 2006;9(4):304–309.

McGinnis R, Cain S, Davidson S, Vitali R, McLean S, Perkins N. Inertial sensor and cluster analysis for discriminating agility run technique and quantifying changes across load. Biomed Signal Proces. 2017;32:150–156.

McGinnis R, Cain S, Davidson S, Vitali R, Perkins N, McLean S. Quantifying the effects of load carriage and fatigue under load on sacral kinematics during countermovement vertical jump with IMU-based method. Sports Eng. 2016;19(1):21–34.

McNair P, Marshall R, Matheson J. Important features associated with acute anterior cruciate ligament injury. N Z Med J. 1990;103(901):537–539.

McNamara J, Brown S, Choi H, Mitchell K. Evaluating the effects of clothing and individual equipment on marksmanship performance using a novel five target methodology. Proceedings of the Human Factors and Ergonomics Society 60th Annual Meeting. 2016;60(1):2043–2047.

Mello R, Murphy M, Vogel J. Relationship between the Army two mile run test and maximal oxygen uptake. Natick (MA): Army Research Institute of Environmental Medicine (US); 1984.

Meyer A, D'Lima D, Besier T, Lloyd D, Colwell C, Fregly B. Are external knee load and EMG measures accurate indicators of internal knee contact forces during gait? J Orthop Res. 2013;31(6):921–929.

MIL-HDBK-781A. Handbook for reliability test methods, plans, and environments for engineering, development qualification, and production. Department of Defense. Arlington (VA): Space and Naval Warfare Systems Command; 1996.

Mills P, Barrett R, Morrison S. Toe clearance variability during walking in young and elderly man. Gait Posture. 2008;28(1):101–107.

Milner C, Ferber R, Pollard C, Hamill J, Davis I. Biomechanical factors associated with tibial stress fracture in female runners. Med Sci Sports Exerc. 2006;38(2):323–8.

MIL-STD-882E. System safety. Department of Defense. Wright-Patterson Air Force Base (OH): Headquarters Air Force Materiel Command (US); 2012.

MIL-STD-1472G. Design criteria standard, human engineering. Department of Defense. Redstone Arsenal (AL): Army Aviation and Missile Command (US); 2012.

Mitchell K. Standard methodology for assessment of range of motion while wearing body armor. Natick (MA): Natick Soldier Research, Development, and Engineering Center (NSRDEC); 2013.

Mitchell K, Batty J, Coyne M, DeSimone L, Choi H, Gregorczyk K, Bensel C. Impact of weight on military mission oriented obstacle course performance. Poster presentation at the 7th International Conference on Applied Human Factors and Ergonomics; 2016; Orlando, FL.

Mitchell K, Choi H, Garlie T. Anthropometry and Range of Motion (ROM) of the encumbered soldier. Natick (MA): Natick Soldier Research, Development, and Engineering Center (NSRDEC); 2017.

Mudie KL, Boynton AC, Karakolis T, O'Donovan MP, Kanagaki GB, Crowell HP, Begg RK, LaFiandra ME, Billing DC. Consensus paper on testing and evaluation of military exoskeletons for the dismounted combatant. J Sci Med Sport. [accessed on 2018]. https://doi.org/10.1016/j.jsams.2018.05.016.

Myer G, Ford K, Brent J, Hewett T. Differential neuromuscular training effects on ACL injury risk factors in "high-risk" versus "low-risk" athletes. BMC Musculoskel Dis. 2007;8(39).

Novacheck T. The biomechanics of running. Gait Posture. 1998;7(1):77–95.

Owings T, Grabiner M. Variability of step kinematics in young and older adults. Gait Posture. 2004;20(1):26–29.

Owings T, Grabiner M. Step width variability, but not step length variability or step time variability, discriminates gait of healthy young and older adults during treadmill locomotion. J Biomech. 2004;37:935–938.

Paasuke M, Ereline J, Gapeyeva H. Knee extension strength and vertical jumping performance in Nordic combined athletes. J Sports Med Phys Fitness. 2001;41:354–61.

Panizzolo F, Freisinger G, Karavas N, Eckert-Erdheim A, Siviy C, Long A, Zifchock R, LaFiandra M, Walsh C. Metabolic cost adaptations during training with a soft exosuit assisting the hip joint. (submitted, in review).

Podraza J, White S. Effect of knee flexion angle on ground reaction forces, knee moments and muscle co-contraction during an impact-like deceleration landing: Implications for the non-contact mechanism of ACL injury. Knee. 2010;17:291–295.

Quesada P, Mengelkoch L, Hale R, Simon S. Biomechanical and metabolic effects of varying backpack loading on simulated marching. Ergonomics. 2000;43(3):293–309.

Ratkevicius A, Stasiulis A, Dubininkaite L, Skurvydas A. Muscle fatigue increases metabolic costs of ergometer cycling without changing VO_2 slow component. J Sports Sci Med. 2006;5(3):440–448.

Rozzi S, Lephart S, Sterner R, Kuligowski L. Balance training for persons with functionally unstable ankles. J Orthop Sports Phys Ther. 1999;29(8):478–86.

Rugelj D, Sevsek F. The effect of load mass and its placement on postural sway. Appl Ergon. 2011;42(6):860–866.

Russell E, Braun B, Hamill J. Does stride length influence metabolic cost and biomechanical risk factors for knee osteoarthritis in obese women? Clin Biomech. 2010;25:438–443.

Schache A, Blanch P, Dorn T, Brown N, Rosemond D, Pandy M. Effect of running speed on lower limb joint kinetics. Med Sci Sports Exerc. 2011;43(7):1260–1271.

Schiffman J, Bensel C, Hasselquist L, Gregorczyk K, Piscitelle L. Effects of carried weight on random motion and traditional measures of postural sway. Appl Ergon. 2006;37:607–614.

Schiffman J, Gregorczyk K, Bensel C, Hasselquist L, Obusek J. The effects of a lower body exoskeleton load carriage assistive device on limits of stability and postural sway. Ergonomics. 2008;51(10):1515–1529.

Sell T. An examination, correlation, and comparison of static and dynamic measures of postural stability in healthy, physically active adults. Phys Ther Sport. 2012;13:80–86.

Shelburne K, Torry M, Pandy M. Muscle, ligament, and joint-contact forces at the knee during walking. Med Sci Sports Exerc. 2005;37(11):1948–1956.

Simpson K, Munro B, Steele J. Backpack load affects lower limb muscle activity patterns of female hikers during prolonged load carriage. J Electromyogr Kinesiol. 2011;21(5):782–788.

Tack D, Kelly A, Richter M, Bray-Miners J. Preliminary results of MC-LEAP testing of US Marine combat load order configurations. Guelph (Ontario): Humansystems Inc.; 2012.

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: Standards of measurement, physiological interpretation and clinical use. Eur Heart J. 1996;17(3):354–381.

Thompson M, Gutmann A, Seegmiller J, McGowan C. The effect of stride length on the dynamics of barefoot and shod running. J Biomech. 2014;47:2745–2750.

Tilbury-Davis D, Hooper R. The kinetic and kinematic effects of increasing load carriage upon the lower limb. Hum Movement Sci. 1999;18(5):693–700.

Umberger B. Short communication Effects of suppressing arm swing on kinematics, kinetics, and energetics of human walking. J Biomech. 2008;41(11):2575–2580.

Umberger B, Martin P. Mechanical power and efficiency of level walking with different stride rates. J Exp Biol. 2007;210(18):3255–3265.

US Army Training and Doctrine Command. The US Army robotic and autonomous systems strategy. Fort Eustis (VA): Army Training and Doctrine Command (US); 2017.

Vaughan C, Davis B, O'Connor J. Dynamics of human gait. 2nd ed. Cape Town (South Africa): Kiboho Publishers; 1992.

Wang Y, Watanabe K. Limb dominance related to the variability and symmetry of the vertical ground reaction force and center of pressure. J Appl Biomech. 2012;28(4):473–478.

Warren R, Marshall J. Injuries of the anterior cruciate and medial collateral ligaments of the knee. A retrospective analysis of clinical records – part I. Clin Orthop Relat Res. 1978;136:191–197.

Waters R, Mulroy S. The energy expenditure of normal and pathologic gait. Gait Posture. 1999;9(3):207-231.

Wikstrom E, Tillman M, Smith A, Borsa P. A new force-plate technology measure of dynamic postural stability: the dynamic postural stability index. J Athl Training. 2005;40(4):305.

Willems T, Witvrouw E, Delbaere K, Mahieu N, De Bourdeaudhuij L, De Clercq D. Intrinsic risk factors for inversion ankle sprains in male subjects: a prospective study. Am J Sports Med. 2005;33(3):415–423.

Winter D. Moments of force and mechanical power in jogging. J Biomech. 1983;16(1):91–97.

Winter D, Patla A, Frank J, Walt S. Biomechanical walking pattern changes in the fit and healthy elderly. Phys Ther. 1990;70(6):340–347.

Winter D. Biomechanics and motor control of human movement. Hoboken (NJ): John Wiley & Sons, Inc.; 2005.

Wisloff U, Castagna C, Helgerud J. Strong correlation of maximal squat strength with sprint performance and vertical jump height in elite soccer players. Brit J Sport Med. 2004;38:285–288.

Yaggie J, McGregor S. Effects of isokinetic ankle fatigue on the maintenance of balance and postural limits. Arch Phys Med Rehab. 2002;83(2):224–228.

Yeow C, Lee P, Goh J. Effect of landing height on frontal plane kinematics, kinetics and energy dissipation at lower extremity joints. J Biomech. 2009;42(12):1967–1973.

Appendix A. Rating of Perceived Exertion Scale

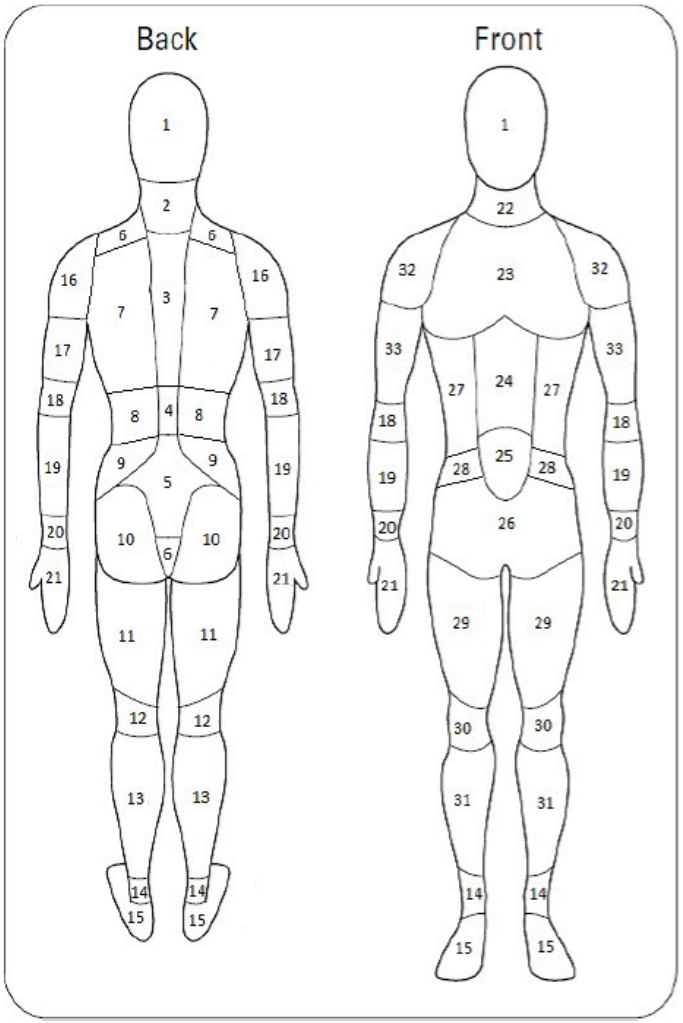
This appendix appears in its original form, without editorial change.

6	No exertion
7	
8	
9	
10	
11	Light
12	
13	Somewhat hard
14	
15	Hard (heavy)
16	
17	Very hard
18	
19	
20	Maximal exertion

Appendix B. Rating of Pain, Soreness, and Discomfort Scale

This appendix appears in its original form, without editorial change.

3. *Pain, Soreness, and Discomfort.* Using the diagram below, please highlight any areas that you felt pain, soreness or discomfort during the condition you have just completed. In the chart on the right hand side, please list the numbers of the highlighted areas and then mark the box indicating the level of pain, soreness or discomfort you experienced in each of those areas. On the next page, please indicate the source of pain, soreness and discomfort for the highlighted areas.

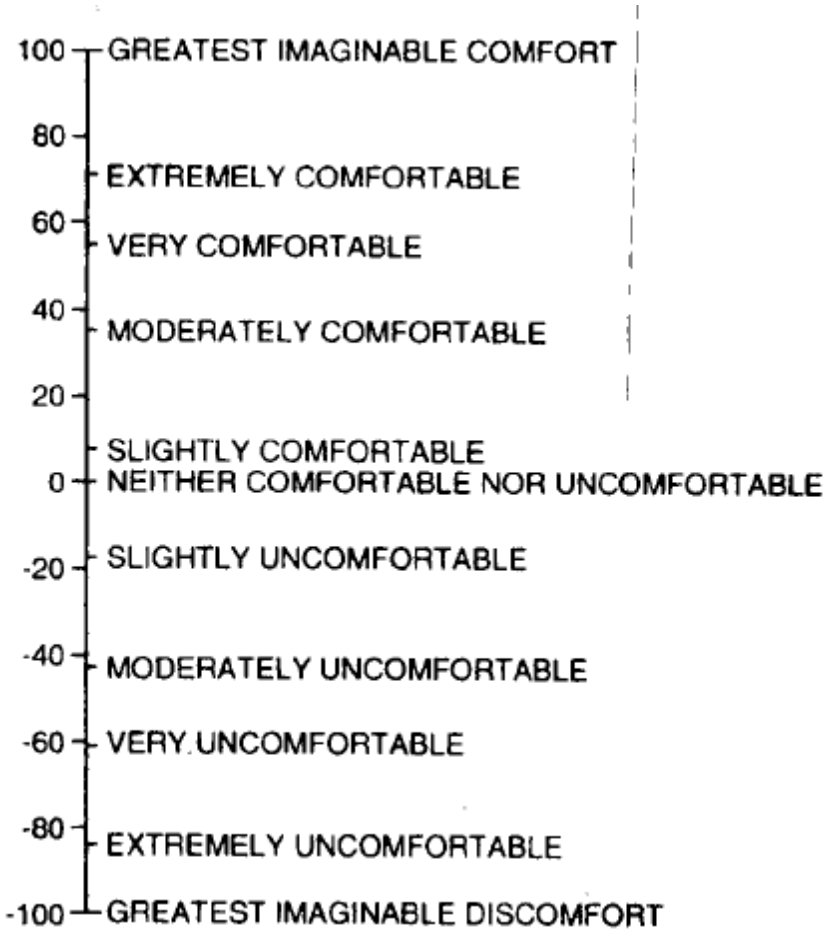


Area of the body	Slight	Moderate	Severe	Extreme
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④
	①	②	③	④

Area of the body _____Source of pain, soreness, or discomfort (e.g., rucksack straps, sensor, mechanical component, blister)

Appendix C. Comfort Affected Labeled Magnitude Scale

This appendix appears in its original form, without editorial change.


100
GREATEST IMAGINABLE COMFORT
80
EXTREMELY COMFORTABLE
60
VERY COMFORTABLE
40
MODERATELY COMFORTABLE
20
SLIGHTLY COMFORTABLE
0
NEITHER COMFORTABLE NOR UNCOMFORTABLE
-20
SLIGHTLY UNCOMFORTABLE
-40
MODERATELY UNCOMFORTABLE
-60
VERY UNCOMFORTABLE
-80
EXTREMELY UNCOMFORTABLE
-100
GREATEST IMAGINABLE DISCOMFORT

List of Symbols, Abbreviations, and Acronyms

ANSUR	dimensions in Anthropometric Survey
ARL	US Army Research Laboratory
CALM	Comfort Affected Labeled Magnitude
CBRNE	chemical, biological, radiological, nuclear, and explosive
COM	center of mass
COP	center of pressure
DOD	Department of Defense
DPSI	Dynamic Postural Stability Index
EMG	electromyography
GRF	ground reaction force
HR	heart rate
HRED	Human Research and Engineering Directorate
HRV	heart rate variability
iEMG	integrated muscle activity
IRB	Institutional Review Board
ISO	International Organization for Standardization
JRF	joint reaction forces
MOS	military occupational specialty
MTBF	mean time between failures
MTC	Minimum Toe Clearance
NSRDEC	Natick Soldier Research, Development & Engineering Center
RER	Respiratory Exchange Ratio
ROM	Range of Motion
RPE	Rating of Perceived Exertion
RPSD	Rating of Pain, Soreness, and Discomfort

TPA	Technology Program Agreement
TRL	technology readiness level
VO_2	volume of oxygen

1 (PDF) DEFENSE TECHNICAL INFORMATION CTR DTIC OCA

2 (PDF) DIR ARL
IMAL HRA
RECORDS MGMT
RDRL DCL
TECH LIB

1 (PDF) GOVT PRINTG OFC
A MALHOTRA

1 (PDF) ARL
RDRL HRB B
T DAVIS
BLDG 5400 RM C242
REDSTONE ARSENAL AL
35898-7290

1 (PDF) ARL
RDRL HRB A
J EVERETTE
BLDG E2929
DESERT STORM DR
FORT BRAGG NC
28310-0001

7 (PDF) ARL
SFC PAUL RAY SMITH CTR
RDRL HRA M CLARKE
RDRL HRA I MARTINEZ
RDRL HRR R SOTTILARE
RDRL HRA C A RODRIGUEZ
RDRL HRA B J HART
RDRL HRA A C METEVIER
RDRL HRA D B PETTIT
12423 RESEARCH PARKWAY
ORLANDO FL 32826

1 (PDF) USA ARMY G1
DAPE HSI M SAMS
300 ARMY PENTAGON
RM 2C489
WASHINGTON DC 20310-0300

1 (PDF) USAF 711 HPW
711 HPW/RH K GEISS
2698 G ST BLDG 190
WRIGHT PATTERSON AFB OH
45433-7604

1 (PDF) USN ONR
ONR CODE 341 J TANGNEY
875 N RANDOLPH STREET
BLDG 87
ARLINGTON VA 22203-1986

1 (PDF) USA NSRDEC
RDNS D D TAMILIO
10 GENERAL GREENE AVE
NATICK MA 01760-2642

1 (PDF) OSD OUSD ATL
HPT&B B PETRO
4800 MARK CENTER DRIVE
SUITE 17E08
ALEXANDRIA VA 22350

ABERDEEN PROVING GROUND

19 (PDF) ARL
RDRL HR
J LANE
J CHEN
P FRANASZCZUK
K MCDOWELL
K OIE
RDRL HRB
J LOCKETT
RDRL HRB C
J GRYNOVICKI
RDRL HRB D
D HEADLEY
C PAULILLO
RDRL HRF A
A DECOSTANZA
RDRL HRF B
A EVANS
A BOYNTON
H CROWELL
C HAYNES
J NEUGEBAUER
J PARK
A TWEEDELL
RDRL HRF C
J GASTON
RDRL HRF D
A MARATHE

1 (PDF) NSRDEC TECH LIB
RDNS COS IMT
P BREMNER

1 (PDF) NSRDEC
RDNS SEO MA
G KANAGAKI

1 NSRDEC
(PDF) H GIROLAMO

1 NSRDEC
(PDF) RDNS SEW PB
K GREGORCZYK

1 NSRDEC
(PDF) RDNS SEO PH
E HENNESSY

1 NSRDEC
(PDF) RDNS SEO PH
K MITCHELL

1 NSRDEC
(PDF) RDNS SEO PB
M O'DONOVAN

1 NIST
(PDF) W BILLOTTE

1 NIST
(PDF) R BOSTELMAN

1 NORTHERN ILLINOIS UNIV
(PDF) D PETERSON

1 ATEC
(PDF) TEDT AT
J CHRISTOPHER