

## Development and testing of new upper-limb prosthetic devices: Research designs for usability testing

---

**Linda Resnik, PT, PhD, OCS**

*Providence Department of Veterans Affairs Medical Center, Providence, RI; Department of Community Health, Brown University, Providence, RI*

**Abstract**—The purposes of this article are to describe usability testing and introduce designs and methods of usability testing research as it relates to upper-limb prosthetics. This article defines usability, describes usability research, discusses research approaches to and designs for usability testing, and highlights a variety of methodological considerations, including sampling, sample size requirements, and usability metrics. Usability testing is compared with other types of study designs used in prosthetic research.

**Key words:** amputation, case studies, consumer, medical devices, methodology, prosthetics, rehabilitation, research design, upper limb, usability.

### INTRODUCTION

Participants at the 2006 State-of-the-Science (SOS) Meeting in Prosthetics and Orthotics (sponsored by the National Institute of Disability and Rehabilitation Research and held at the Rehabilitation Engineering Research Center, Northwestern University, Feinberg School of Medicine; [http://www.nupoc.northwestern.edu/news-publications/papers/sos\\_reports/SOS\\_2006report.pdf](http://www.nupoc.northwestern.edu/news-publications/papers/sos_reports/SOS_2006report.pdf)) identified a wide range of research priority areas that spanned the continuum from product development to identifying the predictors of successful prosthetic wear to improving education for prosthetists fitting complex prostheses [1]. The majority of SOS recommendations related to upper-limb prosthetics focused on expanding the capabilities and control inputs for prosthetic technologies; in other words, the development and

evaluation of new prosthetic devices. Clearly, product development research was identified as a high priority for the field.

Design of highly functional upper-limb prosthetic devices that are acceptable to users has proven to be particularly challenging, as evidenced by high rates of device abandonment reported in the literature [2]. In their 2007 review of more than 200 articles on prosthesis use and abandonment, Biddiss and Chau noted that technological factors relating to discomfort and limited function, particularly for those with higher levels of limb loss, were among the major reasons for prosthetic dissatisfaction and abandonment [2]. Additional reasons for dissatisfaction and abandonment included problems related to durability and mechanical failure, discomfort, control, and cosmesis [2–6]. Although design priorities identified by

---

**Abbreviations:** FDA = Food and Drug Administration; ICF = International Classification of Functioning, Disability, and Health; ISO = International Organization for Standardization; MCID = minimal clinically important difference; MDC = minimum detectable change; RCT = randomized clinical trial; SOS = State-of-the-Science; ULPOM = Work Group on Upper Limb Prosthetic Outcome Measures; VA = Department of Veterans Affairs.

**Address all correspondence to Linda Resnik, PT, PhD, OCS; Research Health Scientist, Providence VA Medical Center, 830 Chalkstone Avenue, Providence, RI 02908; 401-273-7100, ext 2368; fax: 401-863-3489.**

**Email:** [Linda.Resnik@va.gov](mailto:Linda.Resnik@va.gov)

DOI:10.1682/JRRD.2010.03.0050

consumers varied by type of device and level of amputation, identified priorities generally included such key areas as reduced weight, lower cost, improved comfort, increased movement, and greater dexterity [2].

Successful design of prosthetic devices hinges upon a research and development process that intimately combines end-users with device developers [2]. This process focuses on the usability of the product [7], in what is sometimes called a “user-centered” approach to design [7].

The concept of usability and the design and conduct of usability testing studies may be unfamiliar to many in prosthetics and rehabilitation because of the dearth of studies published in these areas. The field of usability engineering originates from the disciplines of human factors science and ergonomics [8], with origins in the aerospace and automotive industries [9]. The usability concept has been embraced by the fields of information technology, assistive technology [8], and medical device development [10]. However, little has been written about usability of prosthetic devices or the best design for conducting usability tests of prosthetic devices. Thus, the overall purposes of this article are to describe the concept of usability testing and introduce designs and methods of usability testing research. The perspective on usability testing presented in this article draws from my experience in usability testing and mixed methods (qualitative and quantitative) research. This synthesis is a direct result of my experience designing and implementing the Department of Veterans Affairs (VA) “Study to Optimize the DEKA Arm,” a multisite usability study funded by the VA Rehabilitation Research and Development Service.

## WHAT IS USABILITY?

The concept of device usability, considered an aspect of “usefulness,” is a qualitative attribute that assesses the ease of use of device-user interfaces. A variety of definitions for usability have been proposed by standardization bodies such as the International Organization for Standardization (ISO) and human factors/usability experts [7,10–12]. The ISO definition of product usability that may be most applicable to medical devices is “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” [13].

Usability is associated with the functionality or utility of the device; in other words, the usefulness of the

device and its perceived value for the purpose for which it is intended. Usability, according to Krug, means “making sure that something works well: that a person of average (or even below average) ability and experience can use the thing—whether it’s a website, a fighter jet, or a revolving door—for its intended purpose without getting hopelessly frustrated” [12].

In his seminal text on usability engineering, Nielson describes a generic framework for product usability delineating the following elements: learnability, efficiency, memorability, errors, and satisfaction [7]. Learning refers to the ease of accomplishing basic tasks with the device the first time users encounter the design. After the basic controls of the device are learned, efficiency relates to the speed at which users can perform tasks with the device, as well as the work and time involved in device use. Memorability relates to the ease of reestablishing proficiency with the device after a period away from it. Errors refer to the frequency and severity of errors in using the controls. Satisfaction refers to the user’s satisfaction with the design and its usefulness.

Usability can have a variety of other attributes related to user comfort, aesthetics, durability, perceived complexity, or other features; specific usability attributes may vary depending on the product and its functions [10]. For users of prosthetic devices, I believe that important attributes of usability include, but are not limited to, functionality achieved while using the device, ease of initial setup of the device, use of good posture and body mechanics while operating the device, daily maintenance, and repairs.

## WHAT IS USABILITY RESEARCH?

Studies that focus on the development and assessment of usability of new products or technology can be classified as usability research or usability engineering [7]. Usability engineering refers to research and design methods for improving ease-of-use during the design process of new products and devices. Usability research can be conducted to inform optimization efforts to refine pre-commercially available devices and to evaluate existing commercially available technology. Usability research has two basic approaches: (1) usability inspection by non-users (such as designers and other experts) [14], and (2) usability testing by device users themselves [15]. Usability testing is the focus of this article.

Information on usability is needed at each stage of prosthesis development to provide feedback to device developers and manufacturers [16]. In the early research and development phase, technical data, i.e., speed and torque of joint movements and angular motions and durability of components, are essential. Although achievement of product engineering goals could be considered an aspect of usability assessment, as development progresses, user testing is needed to assess the use of the device in real-life settings. Just because a device meets the manufacturer's development goals does not necessarily mean that a device is usable. For example, a prosthetic manufacturer may have specific goals regarding torque or speed of grip. User testing is needed to assess whether or not the grip speed and torque are satisfactory for the user's purposes. Thus, user testing can provide information on the performance of devices and the perceived usability that differs from assessment of manufacturing goals.

Usability research can result in recommendations for material changes and other improvements to the device itself, as well as improvements to the human-device interface. As refinements are implemented, additional research is required to assess the changes. Thus, usability research for product development is an iterative process of user engagement and product refinement. Ideally, the incorporation of usability data derived from end-user feedback in an iterative design process enables repeated refinements of product design such that an acceptable level of usability is achieved.

Usability research can also be conducted to evaluate a product after the design is complete and/or after the product is commercially available, in what is sometimes called summative usability research [11]. Summative usability testing can be conducted to compare the usability of several products.

Usability inspection by nonusers may employ strategies such as heuristic evaluation [17], cognitive walkthroughs (to assess ease of learning to use the device), and design reviews [14], formally assessing elements of task performance using the device.

Dozens of methods can be used in usability testing, and the choice of methods depends upon the usability attributes that need to be assessed. Because device usability cannot be directly measured, it is often assessed by means of indirect measures or attributes, such as observation of the user interacting with the device, user's impressions of ease-of-use of the device, and user reports of satisfaction with the device. Many of the qualitative

methods employed in usability testing are drawn from ethnography and involve participant observation, observation of live or video-recorded use of the device, and observations made by investigators of responses (both verbal and visual) made by participants during usability testing. Qualitative methods also include administration of open-ended surveys and questionnaires, interviews, and focus groups. Usability testing methods typically involve gathering and analyzing various quantitative metrics as well. These may be derived from various testing methods, such as quantitative surveys, time-based tests, measures of task performance and completion, and tracking errors and usage. Some usability testing methods ask subjects to choose between several hypothetical engineering trade-offs in device design. This method of usability assessment may have limitations because users may have differing short- and long-term preferences. Additionally, in the assessment of hypothetical trade-offs, users do not get to actually compare the different versions of the device with and without the various components, so their stated preferences are not actually based on user testing.

## RESEARCH APPROACHES TO USABILITY TESTING

Researchers take a variety of approaches to usability testing, choosing qualitative, quantitative, or mixed methodologies, depending upon the researcher, the project, and the objectives. Quantitative and qualitative research paradigms have distinctly different sets of assumptions, terminologies, and methods. Qualitative research takes an integrative, naturalistic approach, which may provide researchers with greater understanding of a complex phenomenon [18]. Although qualitative research is sometimes categorized under the general term descriptive research, its purpose often extends beyond description to include interpretation, prediction, and explanation. Thus, qualitative methodology is an appropriate choice for addressing research questions about complex relationships, clinical situations, or new areas of inquiry. In contrast, a quantitative approach to research is "based upon testing a theory, composed of variables, measured with numbers, and analyzed with statistical procedures" [18].

The mixed methodology approach combines both qualitative and quantitative methods. Mixed method studies are particularly useful when seeking triangulation (i.e.,

convergence) or complementarity (i.e., fuller explanation) of results; for examining overlapping and different facets of a phenomenon; and for discovering paradoxes, contradictions, or fresh perspectives [19]. A mixed methods approach is generally seen as adding breadth and scope to a project [19]. The diversity of methods and data sources can help in understanding all types of usability concerns. Triangulating findings from both qualitative and quantitative data analyses allows a more comprehensive understanding of usability testing.

## STUDY DESIGN FOR USABILITY TESTING

Usability testing typically involves case study or case series designs [20–23] but can also be conducted as within-subjects (crossover) designs and between-subjects quasi-experimental studies. Case studies are typically conducted as uncontrolled observational studies. The single case study involves the study of only one case or individual. A case series, used in many health technology assessments [24], is the study of more than one individual with similar characteristics. Evidence derived from a case series is considered more robust than that derived from a single case study alone. A multiple case series involves the conduct of more than two case series.

Case studies are considered the weakest type of study design for effectiveness research because they do not have a control group that matches the experimental group on all factors except receipt of the intervention (i.e., the prosthesis). Primary threats to the validity of case study designs relate to several elements: history, maturation, and instrumentation. A historical threat to validity occurs when an observed change over time is attributable to events that occur between measurement points and not to the intervention (i.e., the prosthesis). The threat of maturation is that subjects may change because they grow or age, and observed changes may be attributable to maturation, not the intervention. The threat of instrumentation relates to the effects of test taking or changes in the way that measurements are taken. Without a control or comparison group, it is difficult to guard against these threats to internal validity. Thus, case studies and case series designs cannot be used to conduct conclusive evaluations of effectiveness of devices. Despite these limitations, case studies, case series, and multiple case series are the most appropriate design choices for studies of device

usability for prosthetics, where per subject costs are high and the sample is hard to find.

Various design strategies can be used to address the threats to internal validity of simple observational case studies and case series [25]. Quasi-experimental designs for use in case studies and case series employ repeated measures, at least one prior to intervention (baseline) and one after intervention. Use of repeated measures at baseline provides information on the stability of tests and measures and helps minimize threats to instrumentation.

Single subject experimental designs (also called “*n*-of-1” clinical trials) are most appropriate for early studies on treatment efficacy [26]. These designs employ techniques used in larger randomized controlled trials, for example, random assignment to treatment groups and blinding of subject and assessor to treatment condition. Thus, *n*-of-1 trial design would not be appropriate if the goal of the study is to obtain user feedback on device usability.

Study designs that could be used to compare effectiveness of one device with another include randomized clinical trials (RCTs), quasi-experimental designs involving two comparison groups, and crossover designs. Larger RCTs and quasi-experimental designs may not be feasible to conduct because of reasons discussed previously (i.e., heterogeneity of population). Subjects in a crossover study are followed over time and receive a sequence of different treatments (in this case, prostheses) in a within-subjects design [27]. Thus, each subject serves as his or her own control/comparison. The advantages of a crossover study over a quasi-experimental design that compares two separate groups of subjects is that the influence of confounding covariates is reduced because subjects serve as their own comparison or control. The strongest crossover study designs also use methods from clinical trial design, including random assignment to experimental conditions. This reduces threats to internal validity much the way an RCT does.

It is well known that the order in which interventions are administered may affect study outcomes. An example might be that hours of training with any type of prosthesis may carryover from one device to another, improving ease of use with a subsequent device. Thus, the strong crossover study designs both randomize the order in which interventions are received and provide a washout period between treatments to minimize carryover.

## METHODOLOGICAL CONSIDERATIONS IN USABILITY TESTING OF PROSTHETIC DEVICES

### Subjects

A distinctive feature of prosthetic device development is the need to evaluate the device, not only for the amputee's use but also in terms of how easily and proficiently clinical staff (e.g., prosthetists, therapists) use the device, i.e., staff can fit and setup the device and train amputees in its optimal use. If clinicians find it difficult to prescribe, fit, or configure the prosthetic device or refine the setup and train a patient with it, amputee users may experience greater problems with the device. Thus, it is important to consider clinicians as well as amputees as subjects in usability research studies.

Involvement of device users in all stages of product development increases the likelihood of producing devices that are safe, usable, effective, and acceptable to users. Thus, failure to engage users in device development may adversely affect the quality of outcomes and, ultimately, the acceptability of the device [28]. Some of the best examples describing user involvement in device development come from the VA [29–30]. The VA model involved device users at all stages of medical device development, starting with the process of identifying clinical need and conceptualizing a new product and continuing patient and caregiver involvement through an iterative process of device design, development, and testing. Other models of usability research have also arisen from the VA system [31].

### Sample Size

The question of how many subjects is “enough” for usability research is an ongoing issue in usability testing and user interface discussions. In all research, the cost of development and testing must be weighed against the potential benefits of the knowledge to be gained. The expense of manufacturing research and development versions of prototype devices for study and the complexity of using the new devices is a major factor limiting study size. In addition, available research funding for new devices often limits study size. In prosthetics, the amount of time it takes to fit subjects with prosthetic sockets (if necessary) and train them in the use of a new device adds to the cost of conducting such studies. There is a need to balance the costs of producing new devices for testing and/or experimental prosthetic control procedures with having a large enough sample to test for usability.

Whereas drug clinical trials require increasingly larger samples of human subjects to establish safety, effectiveness, side effects, and dosage, usability research typically requires a small number of subjects to identify the improvements, limits, and characteristics of the product. Some research has found that 80 percent of usability problems are detected by the first four or five subjects who use the device and the most severe usability problems are likely to have been detected by the first few subjects [32].

However, other factors must be considered when estimating sample size for usability in upper-limb prosthetics. Sampling in these types of studies is often considered “purposeful” [33]. The objective of purposeful sampling is to gather data that will maximize opportunities to discover variations among concepts and to increase understanding of the phenomena under study. As discussed previously, the upper-limb amputee population is heterogeneous. Thus, usability of a new prosthetic device or technology may need to be assessed separately for subsamples, for example by level of limb loss (if the device is intended for use at different levels); for bilateral or unilateral users; or for those with comorbid conditions, cognitive deficits, or sensory loss. It may also be important to compare and contrast usability of a prosthetic terminal device for new versus experienced users or for males versus females.

The low prevalence of upper-limb amputation, particularly at the higher levels of limb loss, may make accrual of even a small sample size for purposeful sampling quite a challenge. There may be few subjects with the necessary characteristics residing in a given area available to participate in research studies. Thus, research studies may need to use multiple sites to accrue sufficient numbers and include those amputees who are available (convenience sampling).

Purposeful sampling has advantages in helping to maximize the understanding of wide-ranging usability concerns. That said, this sampling strategy, like convenience sampling, also creates an important potential limitation. The use of purposeful or convenience sampling, rather than random selection of subjects from a representative group, can introduce bias by selecting subjects who may not be representative of the broader population. As a hypothetical example, if subjects who agreed to participate in these types of studies were more likely to be unemployed and less happy with their existing devices compared with the broader population of upper-limb amputees, then this could create selection bias.

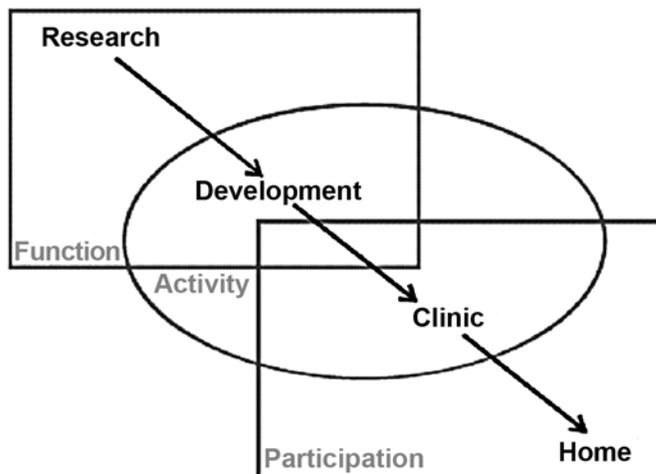
## USABILITY METRICS

A variety of metrics can be used in usability testing. Usability performance measures may include quantitative assessments, such as number of errors made while operating the device, time to complete specified tasks, accuracy of task performance, and time to become proficient in device use.

A particular challenge in the conduct of studies of usability of upper-limb prosthetic devices is that until recently, little consensus has existed on the important domains or constructs that should be assessed related to prosthetic performance. Thus, researchers have needed to develop their own conceptual framework for assessment that is tailored to the device itself and its intended uses and capabilities. Recently, the Work Group on Upper Limb Prosthetic Outcome Measures (ULPOM) recommended the use of the World Health Organization's International Classification of Functioning, Disability, and Health (ICF) as an ideal framework for organizing the selection of outcome measures [16]. The group emphasized the importance of selecting assessment tools that captured each of the important elements of the ICF—including body structures and functions (performance of the prosthesis), activity (carrying out tasks), and participation (use of the prosthesis in real-life situations). It is generally accepted that multiple outcome measures are needed to cover the range of important elements in upper-limb prosthetic research because no single tool captures all aspects [16]. Thus, a tool kit consisting of multiple types of measures may be needed.

The ULPOM also acknowledged that selection of measures should be matched to the type of research being conducted and targeted to the stage of development of the prosthesis (**Figure**). Thus, the most appropriate assessments to use at the earliest stage of prosthesis development include technical measures of grip force, speed, etc.

At the later stages of device development, outcome measures need to address broader areas, such as performance of specific tasks and use in everyday activities. Thus, objective measurements, such as timed dexterity tests, often used as outcomes in studies of upper-limb amputees, need to be supplemented by assessments of the amputee's experience to understand which changes "make a real difference in the lives of patients" [34]. Of particular importance in assessment of prosthetic usability



**Figure.**

Development cycle of prostheses, from research through daily use (black). International Classification of Functioning, Disability, and Health components of assessment for prostheses and their location within development cycle (gray). Adapted from Hill W, Stavadahl O, Hermansson LN, Kyberd P, Swanson S, Hubbard S. Functional outcomes in the WHO-ICF model: Establishment of the Upper Limb Prosthetic Outcome Measures Group. *J Prosthet Orthot.* 2009;21(2): 115–19. DOI: [10.1097/JPO.0b013e3181a1d2dc](https://doi.org/10.1097/JPO.0b013e3181a1d2dc)

ity is the patient's perspective on the usefulness of the device for performing everyday functions, as well as the comfort and fit of their prosthetic limb, their health-related quality of life, and their mobility.

## CHOOSING USABILITY METRICS

A challenge for selection of quantitative outcome measures for usability testing is that few measures have been developed and validated for use with adult upper-limb amputees [35]. Nor has prior research been conducted on the responsiveness and sensitivity to change that would aid researchers in interpreting measurement change scores of upper-limb prosthetic measures [16,35]. Additional research is needed before the field has all necessary data to select and interpret a complete battery of measures [16,35].

In large clinical trials, the effectiveness of interventions is assessed statistically by comparing mean change in outcomes scores between groups of patients. These same comparisons are impossible in clinical practice, individual case studies, or trials using very small patient samples. Common measures of responsiveness of a measure, such

as effect size, standardized response means, or the responsiveness statistic, summarize test responsiveness and are useful for making relative comparisons between measures but do not contribute to the interpretation of test results in individual subjects or patients.

The best instruments for use in case studies, case series, and other types of studies using small samples would have superior measurement properties. In particular, measurement instruments designed to measure change for an individual person must have greater reliability than measures designed for group use. Additionally, research must be done to know how to interpret changes in scores from one point in time to the next. For interpretability, one should be able to answer questions such as, “Does a change in score of 10 points in a certain measure denote an important change for these patients?” and, “Is a 5-point change in score the same as a 10-point change in score?”

Data on two important constructs, minimum detectable change (MDC) [36–38] and minimal clinically important difference (MCID) [39] can assist in assessing the interpretability of scores for individuals and small samples. MDC is a statistical measure of meaningful change, defined as the minimum amount of change that exceeds measurement error [38]. In contrast, MCID defines the threshold at which an individual has experienced an important change [39]. The MDC of measures can be calculated using data from studies of the reliability of the measures.

Unfortunately, at this time, few outcome measures have been validated for adults with upper-limb prosthetics and information is lacking on the MDC and within-subject MCID of outcomes instruments that might be used to assess patients using upper-limb prosthetics. Until further data are available, researchers must use their best judgment, informed by a sound theoretical rationale, to guide their instrument choices and be cautious in interpreting the meaningfulness of change scores using these measures.

### WHY IS USABILITY TESTING A NEW TERM IN UPPER-LIMB PROSTHETICS RESEARCH?

While research has been conducted to assess consumer priorities and needs [2,6], a necessary first step in user-centered design, little research has been conducted that directly involves users of upper-limb prostheses in testing of products during the development stage. Lack of usability research in upper-limb prosthetics may be

attributed to several factors. First, limited funding has been available for this type of research in the United States. Until recently (with the Defense Advanced Research Projects Agency’s Revolutionizing Prosthetics Program), few or no major U.S. federally funded research initiatives in upper-limb prosthetic development existed. Studies of upper-limb prosthetics are inherently challenging because of the lack of homogeneity and relative rarity of upper-limb amputation. Major upper-limb amputees constitute only 3 percent of the U.S. amputee population, an estimated 41,000 persons in 2005. Each year in the United States, an estimated 1,908 upper-limb amputations are performed compared with 56,912 lower-limb amputations [40]. Furthermore, this relatively small population is diverse in terms of the level of etiology, level of limb loss, and number of limbs lost.

The paucity of usability research in upper-limb prosthetics may also be explained, in part, by the fact that studies of new devices are not typically driven by regulatory requirements. Although prostheses are considered “Medical Devices,” subject to regulation by the U.S. Food and Drug Administration (FDA), they are generally considered “Physical Medicine Devices,” which are Class I or low-risk devices. As such, manufacturers are not required to submit a “Premarket Approval Application” to the FDA. Thus, manufacturers are rarely required to conduct research on their new products.

### OTHER TYPES OF PROSTHETICS RESEARCH

Usability studies are one type of scientific research needed to advance the field of prosthetics and to support the design and development of highly functional prosthetic devices that are acceptable to users. Usability research complements other types of prosthetic research, including descriptive epidemiology, neuroscience/engineering research, effectiveness, measurement studies, and cost-effectiveness. These types of studies call for designs and methods that specifically address their goals and thus may employ varied research strategies that may or may not overlap. While a full description of each of the study types and designs is beyond the scope of this article, representative goals of each of these types of studies are shown in the **Table**. User studies to understand the needs and preferences of prosthetic users or predictors of prosthetic use and abandonment are typically observational studies that involve surveys of users [41–47] and providers [48], medical record abstraction [47,49], focus groups [10], or

**Table.**

Types of studies in prosthetics and goals for each type.

Study Type	Example of Goals
Descriptive Epidemiology	Understanding prosthetic users and their needs.
Neuroscience and Engineering	Designing and testing prosthetic components, control mechanisms, and control interfaces.
Usability	Developing/refining prosthetics to meet consumer needs.
Effectiveness and Comparative Effectiveness	Evidence-based prosthetic prescription.
Development of Outcome Measures	Developing reliable and valid measures for clinical and research use.
Cost-Effectiveness	Developing evidence to support cost-effective prosthetic care.

in-person interviews [50]. User studies are typically recommended as the first step in product development research [10]. Research in the fields of neuroscience and engineering is needed to develop better prosthetic components and controls and advance the interface between the amputee user and prosthetic controls. Usability research complements neuroscience and engineering research to ensure that new inventions and products meet user needs. Studies to evaluate effectiveness of devices and compare effectiveness of more than one device type require experimental or quasi-experimental designs [51]. These designs may also be used to compare the usability of several devices or device prototypes. Studies to develop and validate outcome measures for prosthetics involve a variety of development designs and psychometric testing to ensure that measures have adequate reliability, validity, and responsiveness [52–53]. Such studies are needed to guide selection of metrics for studies of effectiveness and usability. Studies to evaluate cost-effectiveness of prosthetic care would analyze data on costs and benefits of treatment, providing data to assist in payment and prescription.

## CONCLUSIONS

Upper-limb prostheses must work well to meet user's needs, or they are likely to be abandoned. The development and evaluation of new upper-limb prosthetic devices that better meet the needs of users was identified as a high priority by participants at the 2006 SOS Meeting in Prosthetics and Orthotics.

Usability research helps to ensure that products are designed, developed, and optimized to meet user needs. User testing studies provide feedback on device performance and perceived usability—providing the user's perspective. User testing studies typically employ mixed

methods and case study designs. Usability testing is often conducted with purposefully selected samples that best represent typical users or subgroups of users. Choice of testing methods depend on the usability attributes that need to be assessed but often include observations of the user interacting with the device, the user's reports of satisfaction and assessment of usability of specific device features, and collection of standardized usability metrics. Measures used in usability studies should be highly reliable and valid. Although usability testing may be new to many in prosthetics research, usability studies complement other types of studies necessary to advance research and practice in upper-limb prosthetics.

## ACKNOWLEDGMENTS

**Financial Disclosures:** The author has declared that no competing interests exist.

**Funding/Support:** This material was based on work supported by VA Rehabilitation Research and Development (grants A6780 and A6780I).

## REFERENCES

1. Rehabilitation Engineering Research Center in Prosthetics and Orthotics. Research in P&O: Are we addressing clinically-relevant problems? Evanston (IL): Northwestern University; 2006.
2. Biddiss E, Chau T. Upper-limb prosthetics: Critical factors in device abandonment. *Am J Phys Med Rehabil.* 2007; 86(12):977–87. [PMID: 18090439] DOI:10.1097/PHM.0b013e3181587f6c
3. Biddiss EA, Chau TT. Upper limb prosthesis use and abandonment: A survey of the last 25 years. *Prosthet Orthot Int.* 2007;31(3):236–57. [PMID: 17979010] DOI:10.1080/03093640600994581

4. Glynn MK, Galway HR, Hunter G, Sauter WF. Management of the upper-limb-deficient child with a powered prosthetic device. *Clin Orthop Relat Res.* 1986;(209):202–5. [\[PMID: 3731596\]](#)
5. Bhaskaranand K, Bhat AK, Acharya KN. Prosthetic rehabilitation in traumatic upper limb amputees (an Indian perspective). *Arch Orthop Trauma Surg.* 2003;123(7):363–66. [\[PMID: 12827395\]](#)  
[DOI:10.1007/s00402-003-0546-4](#)
6. Atkins DJ, Heard DC, Donovan WH. Epidemiologic overview of individuals with upper-limb loss and their reported research priorities. *J Prosthet Orthot.* 1996;8(1):2–11. [\[DOI:10.1097/00008526-199600810-00003\]](#)
7. Nielson J. Usability engineering. San Francisco (CA): Morgan Kaufmann Publishers; 1994.
8. Arthanat S, Bauer SM, Lenker JA, Nochajski SM, Wu YW. Conceptualization and measurement of assistive technology usability. *Disabil Rehabil Assist Technol.* 2007;2(4):235–48. [\[PMID: 19263540\]](#)  
[DOI:10.1080/17483100701343665](#)
9. Meister D. The history of human factors and ergonomics. Mahwah (NJ): Lawrence Erlbaum Associates; 1999.
10. Wiklund M. Medical device and equipment design: Usability engineering and ergonomics. Buffalo Grove (IL): Interpharm Press; 1995.
11. Tullis T, Albert B. Measuring the user experience: Collecting, analyzing and presenting usability metrics. Boston (MA): Elsevier/Morgan Kaufmann; 2008.
12. Krug S. Don't make me think!: A common sense approach to Web usability. Indianapolis (IN): New Riders Press; 2000.
13. ISO 9241-11:1998. Ergonomic requirements for office work with visual display terminals (VDTs)—Part 11: Guidance on usability. Geneva (Switzerland): International Organization for Standardization; 1998.
14. Nielson J, Mack RL. Usability inspection methods. New York (NY): Wiley; 1994.
15. Barnum CM. Usability testing and research. New York (NY): Longman; 2002.
16. Hill W, Kyberd P, Hermansson LN, Hubbard S, Stavadahl Ø, Swanson S. Upper limb prosthetic outcome measures (UPLOM): A working group and their findings. *J Prosthet Orthot.* 2009;21(9):69–82. [\[DOI:10.1097/JPO.0b013e3181ae970b\]](#)
17. Lathan CE, Sebrechts MM, Newman DJ, Doarn CR. Heuristic evaluation of a web-based interface for internet telemedicine. *Telemed J.* 1999;5(2):177–85. [\[PMID: 10908430\]](#)  
[DOI:10.1089/107830299312140](#)
18. Creswell JW. Research design: Qualitative, quantitative, and mixed methods approaches. Thousand Oaks (CA): Sage; 2009.
19. Tashakkori A, Teddlie C. Mixed methodology: Combining qualitative and quantitative approaches. Thousand Oaks (CA): Sage; 1998.
20. LeBlanc M. Current evaluation of hydraulics to replace the cable force transmission system for body-powered upper-limb prostheses. *Assist Technol.* 1990;2(3):101–7. [\[PMID: 10149042\]](#)  
[DOI:10.1080/10400435.1990.10132159](#)
21. Andrade NA, Borges GA, De O Nascimento FA, Romariz AR, Da Rocha AF. A new biomechanical hand prosthesis controlled by surface electromyographic signals. *Conf Proc IEEE Eng Med Biol Soc.* 2007;2007:6142–45. [\[PMID: 18003417\]](#)
22. Cupo ME, Sheredos SJ. Clinical evaluation of a new, above-elbow, body-powered prosthetic arm: A final report. *J Rehabil Res Dev.* 1998;35(4):431–46. [\[PMID: 10220222\]](#)
23. Cupo ME, Sheredos SJ. Clinical evaluation of the Modular Electromechanical Lock Actuator (MELA) for above-elbow prostheses: A final report. *J Rehabil Res Dev.* 1996;33(1):56–67. [\[PMID: 8868418\]](#)
24. Dalziel K, Round A, Stein K, Garside R, Castelnuovo E, Payne L. Do the findings of case series studies vary significantly according to methodological characteristics? *Health Technol Assess.* 2005;9(2):iii–iv,1–146. [\[PMID: 15588556\]](#)
25. Backman CL, Harris SR. Case studies, single-subject research, and N of 1 randomized trials: Comparisons and contrasts. *Am J Phys Med Rehabil.* 1999;78(2):170–76. [\[PMID: 10088595\]](#)  
[DOI:10.1097/00002060-199903000-00022](#)
26. Guyatt GH, Keller JL, Jaeschke R, Rosenbloom D, Adachi JD, Newhouse MT. The n-of-1 randomized controlled trial: Clinical usefulness. Our three-year experience. *Ann Intern Med.* 1990;112(4):293–99. [\[PMID: 2297206\]](#)
27. Chinchilli VM, Esinhart JD. Design and analysis of intra-subject variability in cross-over experiments. *Stat Med.* 1996;15(15):1619–34. [\[PMID: 8858786\]](#)  
[DOI:10.1002/\(SICD\)1097-0258\(19960815\)15:15<1619::AID-SIM326>3.0.CO;2-N](#)
28. Ram MB, Browne N, Grocott P, Weir H. Methods to capture user perspectives in the medical device technology life cycle: A review of the literature in health care, social science, and engineering & ergonomics. London (UK): Multidisciplinary Assessment of Technology Centre for Health; 2005.
29. Sheredos SJ, Cupo ME. The Department of Veterans Affairs Rehabilitation Research and Development Service's technology transfer process. *Technol Disabil.* 1997;7(1):25–29. [\[DOI:10.1016/S1055-4181\(96\)00017-9\]](#)
30. Ross CA. Interim report on VA clinical evaluation of externally powered upper-limb prostheses (June 1971–June 1972). *Bull Prosthet Res.* 1973;10(19):104–23. [\[PMID: 4588223\]](#)
31. Malassigné P, Nelson AL, Cors MW, Jensen RP, Amato M, Schnurr ES, Amerson TL. Iterative design and evaluation of new prone carts for individuals with SCDs: A technical note. *J Rehabil Res Dev.* 2002;39(1):127–39. [\[PMID: 11930905\]](#)

32. Virzi R. Refining the test phase of usability evaluation: How many subjects is enough? *Human Factors*. 1992;34(4):9.
33. Strauss AL, Corbin JM. *Grounded theory in practice*. Thousand Oaks (CA): Sage; 1997.
34. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. *Phys Ther*. 2006;86(5):735–43. [PMID: 16649896]
35. Wright V. Prosthetic outcome measures for use with upper limb amputees: A systematic review of the peer-reviewed literature, 1970 to 2009. *J Prosthet Orthot*. 2009;21(9):3–63. DOI:10.1097/JPO.0b013e3181ae9637
36. Fritz JM, Irrgang JJ. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Phys Ther*. 2001;81(2):776–88. [PMID: 11175676]
37. Stratford PW, Binkley J, Solomon P, Finch E, Gill C, Moreland J. Defining the minimum level of detectable change for the Roland-Morris questionnaire. *Phys Ther*. 1996;76(4):359–68. [PMID: 8606899]
38. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol*. 1999;52(9):861–73. [PMID: 10529027] DOI:10.1016/S0895-4356(99)00071-2
39. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10(4):407–15. [PMID: 2691207] DOI:10.1016/0197-2456(89)90005-6
40. Dillingham TR, Pezzin LE, MacKenzie EJ. Limb amputation and limb deficiency: Epidemiology and recent trends in the United States. *South Med J*. 2002;95(8):875–83. [PMID: 12190225]
41. Biddiss EA, Chau TT. Multivariate prediction of upper limb prosthesis acceptance or rejection. *Disabil Rehabil Assist Technol*. 2008;3(4):181–92. [PMID: 19238719] DOI:10.1080/17483100701869826
42. Datta D, Selvarajah K, Davey N. Functional outcome of patients with proximal upper limb deficiency—Acquired and congenital. *Clin Rehabil*. 2004;18(2):172–77. [PMID: 15053126] DOI:10.1191/0269215504cr716oa
43. Davidson J. A survey of the satisfaction of upper limb amputees with their prostheses, their lifestyles, and their abilities. *J Hand Ther*. 2002;15(1):62–70. [PMID: 11866354] DOI:10.1053/hanthe.2002.v15.01562
44. Saradjian A, Thompson AR, Datta D. The experience of men using an upper limb prosthesis following amputation: Positive coping and minimizing feeling different. *Disabil Rehabil*. 2008;30(11):871–83. [PMID: 17852212] DOI:10.1080/09638280701427386
45. Wright TW, Hagen AD, Wood MB. Prosthetic usage in major upper extremity amputations. *J Hand Surg Am*. 1995;20(4):619–22. [PMID: 7594289] DOI:10.1016/S0363-5023(05)80278-3
46. Pezzin LE, Dillingham TR, Mackenzie EJ, Ephraim P, Rossbach P. Use and satisfaction with prosthetic limb devices and related services. *Arch Phys Med Rehabil*. 2004;85(5):723–29. [PMID: 15129395] DOI:10.1016/j.apmr.2003.06.002
47. Dudkiewicz I, Gabrielov R, Seiv-Ner I, Zelig G, Heim M. Evaluation of prosthetic usage in upper limb amputees. *Disabil Rehabil*. 2004;26(1):60–63. [PMID: 14660200] DOI:10.1080/09638280410001645094
48. Schultz AE, Baade SP, Kuiken TA. Expert opinions on success factors for upper-limb prostheses. *J Rehabil Res Dev*. 2007;44(4):483–89. [PMID: 18247245] DOI:10.1682/JRRD.2006.08.0087
49. Gaine WJ, Smart C, Bransby-Zachary M. Upper limb traumatic amputees. Review of prosthetic use. *J Hand Surg Br*. 1997;22(1):73–76. [PMID: 9061532] DOI:10.1016/S0266-7681(97)80023-X
50. Kejlaa GH. Consumer concerns and the functional value of prostheses to upper limb amputees. *Prosthet Orthot Int*. 1993;17(3):157–63. [PMID: 8134275]
51. Stein RB, Walley M. Functional comparison of upper extremity amputees using myoelectric and conventional prostheses. *Arch Phys Med Rehabil*. 1983;64(6):243–48. [PMID: 6860093]
52. Lindner HY, Linacre JM, Norling Hermansson LM. Assessment of capacity for myoelectric control: Evaluation of construct and rating scale. *J Rehabil Med*. 2009;41(6):467–74. [PMID: 19479160] DOI:10.2340/16501977-0361
53. Wright FV, Hubbard S, Naumann S, Jutai J. Evaluation of the validity of the prosthetic upper extremity functional index for children. *Arch Phys Med Rehabil*. 2003;84(4):518–27. [PMID: 12690590]

Submitted for publication March 30, 2010. Accepted in revised form September 21, 2010.

This article and any supplementary material should be cited as follows:

Resnik L. Development and testing of new upper-limb prosthetic devices: Research designs for usability testing. *J Rehabil Res Dev*. 2011;48(6):697–706.

DOI:10.1682/JRRD.2010.03.0050

