Chapter 6
Challenges to Effective Evaluation of Assistive Technology

Richard Simpson

Abstract Evaluation in the context of assistive technology can take several forms. Engineers evaluate the devices and technologies they develop. Clinicians perform clinical evaluations to decide which device(s) are most appropriate for their client. Finally, researchers hope to evaluate the long-term outcomes of assistive technology interventions. Evaluating a new assistive technology in the lab can be complicated by small user populations and the lack of universally accepted performance measures. The obstacles to effective clinical evaluations include the wide variety of devices that can meet each client’s needs and the limited time and resources available to clinicians and their clients. Evaluating assistive technology outcomes is a relatively new pursuit, with many open questions.

6.1 Introduction

Evaluation is a critical aspect of the research and development process. The scientific method rests on the ability to measure the results of actions or interventions. Similarly, the development process is almost always depicted as a repeating cycle of implementation and evaluation. There are at least three different kinds of evaluation in regards to assistive technology (AT). Scientists and engineers evaluate the devices and technologies they develop as part of the research and development process. Clinicians perform clinical evaluations of a client’s needs and abilities to decide which device(s) are most appropriate for their client. Finally, there is the idea (though rarely put into practice) of evaluating the long-term outcomes once a client has received AT.
6.2 Evaluating Technology in the Lab

Evaluating a new AT in the lab can be a challenge for even the most rigorous scientist with the best of intentions. Depending on the target user population, identifying a sufficient number of potential users who are able to travel to the lab to trial devices can be difficult. An alternative is to utilize single-subject or “small N” experimental designs, but these are difficult to generalize.

As an example, consider the DriveSafe System (DSS) [15]. The target user population for the DSS includes individuals who have both mobility impairments and visual impairments. This population is small to begin with and, almost by definition, has difficulty traveling to a lab. We were only able to successfully recruit two individuals from this population to test out the DSS. One approach we took to augment our testing with potential users was to use blindfolded able-bodied individuals as subjects.

Unfortunately, able-bodied subjects lack the orientation and mobility (O&M) skills of individuals who are actually visually impaired. A second approach was to blindfold actual O&M specialists, which was more realistic, but not even O&M specialists have the navigation skills of someone who is truly blind. A third approach we pursued was to use ambulatory blind individuals. These participants had the appropriate navigation skills, but they were unable to give us much insight into the needs of visually-impaired wheelchair users. In all three cases, we were able to recruit enough participants to perform group statistical analyses, at the expense of a realistic appraisal of the system’s performance with the target user group.

A significant obstacle to accurately evaluating some technologies is the need for lots of training. Some technologies are inherently difficult to master, especially in cases where using a different technology outside of the lab in between training sessions can interfere with retention of skills taught during training. Even with regular use, some technologies can take months to master. For example, operating an augmentative and alternative communication (AAC) device with a sophisticated language encoding scheme like Semantic Compaction [1] is akin to learning a new language.

In the case of the DSS, some participants were given several hours of training prior to initiating data collection. Participants needed to learn how to operate a powered wheelchair that occasionally refused to travel in the indicated direction because of perceived obstacles. Then, participants needed to learn how to do this blindfolded.

A final obstacle to evaluating AT in the lab is the decision of what to measure. Investigators often emphasize speed at the expense of other valid measures like accuracy, comfort, or workload. When evaluating the DSS, for example, we knew the DSS was likely to cause participants to take longer to complete navigation tasks because the DSS slows down the wheelchair in the presence of obstacles. A participant completing a navigation task without the DSS could drive straight from the start point to the goal, if he or she chose, completing the navigation task in the shortest time possible at the expense of