

Chapter 1.7

Evaluating Mobile Human–Computer Interaction

Chris Baber

The University of Birmingham, UK

ABSTRACT

In this chapter the evaluation of human computer interaction (HCI) with mobile technologies is considered. The ISO 9241 notion of ‘context of use’ helps to define evaluation in terms of the ‘fitness-for-purpose’ of a given device to perform given tasks by given users in given environments. It is suggested that conventional notions of usability can be useful for considering some aspects of the design of displays and interaction devices, but that additional approaches are needed to fully understand the use of mobile technologies. These additional approaches involve dual-task studies in which the device is used whilst performing some other activity, and subjective evaluation on the impact of the technology on the person.

INTRODUCTION

This chapter assumes that ‘usability’ is not a feature of a product, that is, it does not make sense to call a product itself ‘usable’. Rather, usability

is the consequence of a given user employing a given product to perform a given activity in a given environment. Holcomb and Tharp (1991) proposed a ‘model’ of interface usability, which is illustrated by Table 1. The definitions presented in Table 1 arose from consideration of the user interface of desk-based computers. However, it ought to be apparent that the majority of the components are defined in terms of an individual’s perceptions of features of the user interface.

The International Standards Organization has a number of standards relevant to human-computer interaction (Bevan, 2001). Current standards for mobile devices tend to focus on product attributes, for example, *ISO 18021: Information Technology—User Interface for Mobiles* (2001) provides interface specifications for Personal Digital Assistants. Other Standards have recognized the multifaceted nature of usability and have sought to encourage an approach that is similar to Quality Assessment (Earthey et al., 2001). Demonstrating compliance with the standards requires analysts to document their evaluation, demonstrating how it meets the objectives of the standard. The

Table 1. Holcomb and Tharp's (1991) "model" of interface usability

Component	Term
Functional	Able to accomplish tasks for which software is intended Perform tasks reliably and without errors
Consistent	Consistent key definitions Show similar information at same place on screens Uniform command syntax
Natural and Intuitive	Learnable through natural conceptual model Familiar terms and natural language
Minimal memorization	Provide status information Don't require information entered once to be re-entered Provide lists of choices and allow picking from the lists Provide default values for input fields
Feedback	Prompt before destructive operations like DELETE Show icons and other visual indicators Immediate problem and error notification Messages that provide specific instructions for action
User help	Online help system available Informative, written documentation
User control	Ability to undo results of prior commands Ability to re-order or cancel tasks Allow operating system actions to be performed within the interface

definition of usability offered by the International Standards Organization, that is, in ISO9241, part 11, 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.*" (ISO9241-11, 1998). The implications are that, first, usability is the consequence of a given user employing a given product to perform a given activity in a given environment (as stated) and, second, that it is possible to measure aspects of this relationship in terms of effectiveness, efficiency, and user satisfaction. It is important to note that these three aspects are inter-connected and that any evaluation activity ought to try to measure some aspect of each (Frøkjær et al., 2000).

Defining Evaluation Targets

If one is able to speak of measures, then it makes sense to be able to determine some criteria that indicate good or poor performance on these

measures. Good et al. (1986) proposed that it is important to define both evaluation targets and metrics that relate to these targets. For example, in a study of conferencing systems, Whiteside et al. (1988) identified 10 attributes that they felt reflected the use of the conferencing system, for example, ranging from a fear of feeling foolish to a number of errors made during task performance. For each attribute, Whiteside et al. (1988) defined a method for collecting data about that attribute, for example, questionnaires, observation, and so forth, and then set performance limits relating to best, worst, and planned levels. A study of a wearable computer for paramedics (Baber et al., 1999) used this concept to produce Table 2. In Table 2, three measures of performance were undertaken, that is, predictive modeling (using critical path analysis), user trials, and performance improvement arising from practice. In addition, three subjective evaluation methods were used. Table 2 shows how the system met (or exceeded) some of the target criteria but fell below the target

13 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/evaluating-mobile-human-computer-interaction/22241

Related Content

Users' Acceptance of Artificial Intelligence-Based Chatbots: An Empirical Study

Mahendar Goli, Anoop Kumar Sahu, Surajit Bagand Pavitra Dhamija (2023). *International Journal of Technology and Human Interaction* (pp. 1-18).

www.irma-international.org/article/users-acceptance-of-artificial-intelligence-based-chatbots/318481

Logistic Clustering Dynamics Determinants: A Bibliometric and Prospective Approach

Naoual Bouhtati, Mourad Aarabe, Lhoussaine Allaand Youssef El Yaacoubi (2025). *Utilizing Technology to Manage Territories* (pp. 139-174).

www.irma-international.org/chapter/logistic-clustering-dynamics-determinants/360491

Exploring the Behavioral Dimension of Client/Server Technology Implementation: An Empirical Investigation

Eitel J.M. Lauría (2006). *International Journal of Technology and Human Interaction* (pp. 63-81).

www.irma-international.org/article/exploring-behavioral-dimension-client-server/2887

Cyberbullying in the World of Teenagers and Social Media: A Literature Review

Sophia Alim (2017). *Gaming and Technology Addiction: Breakthroughs in Research and Practice* (pp. 520-552).

www.irma-international.org/chapter/cyberbullying-in-the-world-of-teenagers-and-social-media/162534

Evaluating the Relevance of Contextual Hyper-Advertising on Social Media: An Empirical Study

Dhote Triptiand Zahoor Danish (2016). *International Journal of Social and Organizational Dynamics in IT* (pp. 39-47).

www.irma-international.org/article/evaluating-the-relevance-of-contextual-hyper-advertising-on-social-media/158055