Chapter XLVII
Using Wizard of Oz to Evaluate Mobile Applications

Janet C. Read
University of Central Lancashire, UK

ABSTRACT

This chapter describes Wizard of Oz studies and gives a historical view that includes a summary of the literature in which several studies that used the method with mobile technology are discussed. The use of Wizard of Oz for mobile applications is explored by referencing the literature and by examination of a case study. A taxonomy for Wizard of Oz studies is presented that has been derived from a study of the literature in this area. A set of guidelines is presented that outlines the essential considerations in planning a Wizard of Oz study for mobile applications. The chapter concludes with some thoughts for future Wizard of Oz studies.

INTRODUCTION

In most technology applications, developers are required to evaluate systems for usability and user acceptance. These evaluations require careful planning and there are often pressures on resources and time that constrain the evaluation team. Where technology is particularly novel and where the cost of producing the technology is high, the evaluations that need to be carried out before any investment is provided pose a particular problem.

One method for evaluating novel or incomplete products is to use simulation. Simulation assumes that the evaluators engage with a product that ‘looks like,’ ‘behaves like,’ or ‘feels like’ the eventual product, but is, in one way or another, incomplete. The benefits of using simulations are that certain features of products can be evaluated before a fully functional product is available.

In mobile applications, evaluators are often concerned with the ease of interaction with a product as well as with the effect of the product on
Using Wizard of Oz to Evaluate Mobile Applications

user behaviour. For many mobile applications, the development of the communication technology and the development of the product interface require considerable investment. However, for the purpose of evaluation of concept and evaluation of ease of use, these two components can be simulated and therefore tested early in the development stage keeping the cost of redesign low.

In HCI (Human Computer Interaction), evaluation using simulation has become synonymous with the term ‘Wizard of Oz’. A Wizard of Oz evaluation is one in which some, or all, of the interactivity that would normally be controlled by computer technology is ‘mimicked’ or ‘wizarded’. The Wizard of Oz method is considered mainstream in HCI and, as user groups have diversified and as the technologies under investigation have changed, the method has become more and more popular in evaluation studies.

Wizard of Oz methods have long been associated with novel and emerging interfaces. Originally presented as a method for simulating speech input at a time when speech recognition was flaky, the method is particularly well suited to mobile technology which is often novel and where, as indicated by Cohen and Oviatt (1994), “portable computing and communications devices will soon be too small to allow for the use of a keyboard, implying that the input modalities for such machines will most likely be digitising pens and voice.”

This chapter presents an overview of Wizard of Oz as a method and then focuses down on its use for evaluations and specifically for evaluations of mobile applications. The first section introduces WOz (Wizard of Oz) methods by giving a historical overview which is augmented with references from early studies that used the method. This section concludes with a consideration of how WOz is used in the Product Development Lifecycle. The second section presents the literature on the use of WOz in mobile applications before concluding with a taxonomy that describes the variability of WOz studies. The third section focuses on the process of carrying out a WOz study, beginning with a case study that demonstrates how a WOz study was used in an evaluation of a mobile system in a museum and concluding with a discussion of when and how to use WOz. This section includes some tips and guidelines for the use of WOz with mobile technology. The chapter concludes with some reflections on the use of Wizard of Oz and presents some research ideas for future studies in this exciting area.

THE WIZARD OF OZ METHOD

A traditional Wizard of Oz study (shown in Figure 1) has three components; a human wizard, an interface, and a subject. During the study, the human wizard manipulates the interface in such a way that the subject is unaware (to varying extents) of the existence and the impact of the wizard.

There is some debate about the origins of the method as the concept was being used well ahead of the adoption of the name. As a method, it can be traced back to an IBM technical report by Thomas (1976), but it is more often attributed to Gould, Conti, and Hovanyecz (1983), whose study of a listening typewriter (an early simulation of a speech recognition system) is well cited and whose diagram of a WOz study is often reproduced. In the Gould et al. (1983) study, a skilled typist was employed to enter what the participants said to the computer and to therefore act as a wizard by mimicking the potential behaviour of a speech recognition engine.

The phrase Wizard of Oz comes from the book ‘The Wonderful Wizard of Oz’ by Baum (1900). Central to this book is the character known as the Wizard of Oz who is believed, by Dorothy Gale and her friends, to be the only one able to solve their problems. Unseen for most of the book, Oz is reluctant to meet the travellers and constantly appears in disguise, once as a giant head, once as a beautiful fairy, once as ball of fire, and once as a horrible monster. As the story progresses, it becomes apparent that Oz is actually none of these things, but is in fact just an ordinary American man who has been using a lot of elaborate magic tricks and props to make himself seem ‘great and powerful.’ Incidentally, his name, Oz comes from the first letters of the first two words in his name (Oscar Zoroaster).
Related Content

Structural Effects of Platform Certification on a Complementary Product Market: The Case of Mobile Applications
[www.irma-international.org/chapter/structural-effects-platform-certification-complementary/26620/](www.irma-international.org/chapter/structural-effects-platform-certification-complementary/26620/)

Mobile Gaming
[www.irma-international.org/chapter/mobile-gaming/17124/](www.irma-international.org/chapter/mobile-gaming/17124/)

A CASE Tool for Java Mobile Computing Applications
[www.irma-international.org/article/case-tool-java-mobile-computing/43892/](www.irma-international.org/article/case-tool-java-mobile-computing/43892/)

Prioritization Schemes in Queuing Handoff and New Calls to Reduce Call Drops in Cellular Systems
[www.irma-international.org/article/prioritization-schemes-queuing-handoff-new/55084/](www.irma-international.org/article/prioritization-schemes-queuing-handoff-new/55084/)

Threshold-Based Location-Aware Access Control
[www.irma-international.org/article/threshold-based-location-aware-access/55889/](www.irma-international.org/article/threshold-based-location-aware-access/55889/)