Understanding and Improving Usability Inspection Methods

Alan Woolrych
University of Sunderland, UK

Mark Hindmarch
University of Sunderland, UK

INTRODUCTION

Usability inspection method (UIM) is the term used for a variety of analytical methods designed to "find" usability problems in an interface design. The basic principle involves analysts inspecting the interface against a set of pre-determined rules, standards or requirements. Analysts inspect the interface and predict potential usability problems based on breaches of these rules. None of the UIMs currently in use are capable of detecting all of the problems associated with an interface. After describing some of the UIMs in use, this article will look at the authors' work on improving these methods by focusing on the resources analysts bring to an inspection.

BACKGROUND

In order to better explain the work we have done on improving UIMs, three of the more commonly used UIMs will be described. These are by no means the only usability inspection methods; other examples being claims analysis (Carroll & Rosson, 1991) and pluralistic walkthroughs (Bias, 1994).

Heuristic Evaluation

This method was developed by Nielsen (1992). The basis of the method is the comparison of an interface with a set of usability guidelines, known as the heuristics. Originally nine, there are currently 10 guidelines dealing with areas such as visibility of system status, user control and freedom and error prevention.

A heuristic evaluation is carried out by a number of different evaluators; five is recommended as the optimal number (Nielsen & Landauer, 1993), and the problems identified by the individual evaluators are then merged into a master problem set.

This technique has been used at numerous stages in the development process from paper prototypes to full software packages (Nielsen, 1990). The advantages of the technique, and the reason the method is so popular, are that it can be used by novices as well as experts, although novices find fewer problems than experts (Nielsen, 1992) and the technique is comparatively quick and inexpensive to employ. The disadvantage is that it tends to only uncover more superficial problems with an interface; problems that require complex interaction on the part of the user are more likely to be missed by heuristic evaluation.

Cognitive Walkthrough

This technique is based on the CE+ [This is a combination of Cognitive Complexity Theory (**C**CT) (Kieras & Polson, 1985) and Explanation-based Learning (**E**XPL) (Polson & Lewis, 1990)] theory of exploratory learning. This theory states that users exploring a new interface are guided by general task goals, and they search for interface elements that promise to move them closer to these goals. Cognitive Walkthrough is a practical technique for applying CE+ in an evaluation and was fully outlined in Wharton, Rieman, Lewis, and Polson (1994). In contrast to Heuristic Evaluation, Cognitive Walkthrough can only be performed by experts.

The technique focuses on how well an interface can support a novice user without formal training. A Cognitive Walkthrough is usually performed by the interface designer with a small group of colleagues. It requires that certain information be available to

the evaluators to be successful, including a description of the users and their knowledge resources, a description of the tasks to be performed and the correct sequence of actions necessary to carry out the tasks. In performing the walkthrough, for each step in a task, the evaluation team asks a series of questions including:

- Is the correct action obvious to the user?
- Will the user match the system's response with the chosen action?

Cognitive Walkthrough has been criticized for being too time consuming and requiring large amounts of paperwork to be completed, although attempts have been made to streamline the method (Rowley & Rhoades, 1992; Spencer, 2000).

Heuristic Walkthrough

Sears (1997) proposed a method which combines aspects of cognitive walkthroughs and heuristic evaluation to address the weaknesses of both.

Heuristic walkthrough is a two phase technique. The first phase has similarities with cognitive walkthrough, in that evaluators have a set of questions to guide their exploration of the interface as well as a set of common user tasks; this is designed to expose the evaluators to the core functionality of the interface. During the second phase, the evaluators use usability heuristics to assess problems with the interface. However, unlike a straightforward heuristic evaluation which is relatively unstructured, the use of the heuristics in a heuristic walkthrough is focused on those areas of the interface identified as important in the first phase of the evaluation.

It is claimed that the major advantage of heuristic walkthrough is its ability to identify severe usability problems compared to heuristic evaluation while avoiding the narrow focus commonly associated with cognitive walkthroughs.

Despite variations in the strengths and weaknesses of the various UIMs, the unreliability of the assessment of all such methods is well documented (e.g., Gray & Salzman, 1998). In practice, these methods fail to predict *all* of the usability problems in a design; not all of the analysts' predictions are true predictions. Such false predictions are commonly known as false positives. For a variety of reasons, analysts will make predictions about usability problems that in reality cause no problems to the users (Cockton & Woolrych, 2001).

For example, UIMs such as Heuristic Evaluation (Nielsen, 1992) are simply not good enough in their current state. The negative outcome of the use of such inspection methods is two-fold. First of all, because such methods are not thorough (they fail to find all of the usability problems), designs subjected to them can result in poor usability, especially if the nature of their flaws is not fully understood. For example, is there a type of usability problem that the method is typically good at finding? Or more importantly, is there a type of problem the method is particularly bad at finding, and are these problems likely to be severe ones? Second, if the false positives are addressed as real usability problems, time and money is wasted in the redesign of usable features.

Although the assessment of UIMs has been very poor (Gray & Saltzman, 1998), this has improved recently (e.g., Cockton, Woolrych, Hall, & Hindmarch, 2003). Research must, therefore, focus on the reliable assessment of inspection methods before work on inspection method improvement can begin.

Despite these problems with inspection methods, there is still a place for reliable UIMs given the original rationale for their development—saving valuable resources such as time and costs. The challenge is to improve UIM quality without increasing costs.

IMPROVING USABILITY INSPECTION METHODS

Thorough assessment of UIMs is reliant on accurate coding of analyst (non)-predictions. UIMs are commonly assessed by their validity, thoroughness, and effectiveness (Sears, 1997), even though percentages fail to comprehensively assess UIMs (Woolrych & Cockton, 2000).

Validity drops as the number of problems found with a UIM exceeds the real problems found. Analysts make false predictions (false positives) as well as successful ones. Fewer false positives mean a more valid UIM.