Introduction

Research question
How do different usability evaluation methods, focussed on experts and users, contribute to the evaluation of a system during an iterative design process?

The PAT Workbench was used as a case to test usability methods. This is a tool that was developed at the department of Communication and Information Sciences of the University of Groningen, to store, annotate, retrieve and view multimodal instructions (MIs). MIs are instructions that consists of text and pictures. The goal of the PAT Workbench is to create a corpus of annotated MIs for further research. The framework used for this process was the design research cycle (Hevner, 2007).

Methods

Design research cycle

Environment

Design Research Cycle

Knowledge Base

Problems/Issues

Design Processes

Scenarios/Objectives

Tasks/Steps

Evaluation

Expert review

• Functional analysis
• Heuristic inspection
• Cognitive walkthrough w.r.t. PAT’s top tasks (upload MI, search MI and annotate MI

User evaluation

Two tests with real users, students enrolled in the master course on multimodal instructions in Communication and Information Sciences.

• Test 1 at beginning of course, with 9 users
• Test 2 after seven weeks, with 4 users
• Three top tasks while thinking aloud
• A questionnaire after each task
• Two general usability questionnaires
• Interview with the participants.

User evaluation

Two tests with real users, students enrolled in the master course on multimodal instructions in Communication and Information Sciences.

• Test 1 at beginning of course, with 9 users
• Test 2 after seven weeks, with 4 users
• Three top tasks while thinking aloud
• A questionnaire after each task
• Two general usability questionnaires
• Interview with the participants.

Results

Expert review: Heuristic inspection

This inspection was based on the 10 heuristic principles by Nielsen (1995). Results showed that, given the three top tasks, the two most severe usability problems were ‘visibility of system status’ (e.g. lack of feedback to the user, available information and functions were not clear) and ‘consistency and web standards’ (e.g. clickable links styled as text, site behaved not as expected).

Functional analysis

• Provides useful insights in time based efficiency, error count and task completion,
• Displays multiple aspects of the system, which an expert may overlook,
• Helps to overcome obvious issues in a more expensive user evaluation.

Heuristic inspection

• Is cheap,
• Offers a detailed system description,
• Helps to overcome obvious issues in a more expensive user evaluation.
• Displays multiple aspects of the system, which an expert may overlook,
• Provides useful insights in time based efficiency, error count and task completion,
• Benefits from Think Aloud Protocols (although with concurrent TAP participants continuously need reminders to verbalise their thoughts).

In longitudinal studies, added value in iterative tests may be gained from fresh participants in addition to the original ones.

In between testing, it is advised to not let participants use a beta version of the product for their own work.

In both types of evaluation good communication and collaboration between developer and tester are crucial.

Conclusion

An iterative design process allows for the use of various evaluation methods, which contribute to the evaluation of a system in different ways:

Expert review

• Is cheap,
• Offers a detailed system description,
• Helps to overcome obvious issues in a more expensive user evaluation.
• Displays multiple aspects of the system, which an expert may overlook,
• Provides useful insights in time based efficiency, error count and task completion,
• Benefits from Think Aloud Protocols (although with concurrent TAP participants continuously need reminders to verbalise their thoughts).

In longitudinal studies, added value in iterative tests may be gained from fresh participants in addition to the original ones.

In between testing, it is advised to not let participants use a beta version of the product for their own work.

In both types of evaluation good communication and collaboration between developer and tester are crucial.

References
