

## **1 Introduction, p. 1**

User testing is used in this paper as another word for usability testing. It can be used for prototypes and for “late beta versions”. The paper describes “Think Aloud Testing” by Lewis 1982

## **2 Overview, p. 2**

Generally a TAT has some practical tasks given to the test subject who is asked to perform them while thinking aloud about what she wonders about and tries to do and so on. Afterwards there is a “debriefing”

## **3 Developing the test**

**3.1** The tasks may be hypothetical, but they need to be realistic. It is good if they are actual problems. They are developed in collaboration between the experimenter, a software designer and a “user representative”. The tasks are given to the test subject in written form to avoid hidden cues.

**3.2** The tasks need to be short. It is favorable if the tasks are tailored to the specific test subject, so that they feel relevant. A problem may be that the specification of the task gives information that the program can actually solve the problem, which is not natural. A way to solve this can be to provide a task in the beginning which can not be solved and intervening after a short while.

**3.3** The test subjects need to be the right people, they have to be realistic possible users and not over- or under-qualified, the sample should be representative. Asking specific functional questions, that test the user in relevant knowledge can be used. The user should be compensated but not too much.

**3.4** The user should be informed about the test: Time and location, what will be tested, how it will be conducted, whether the test will be videotaped, legal matters without being written in legalese and documentation for the software can also be provided.

**3.5** A pilot test can be conducted which tests the test itself; for example the test subject might not understand the question, the tasks are bad

## **4 Conducting the test**

**4.1** The subject should feel comfortable. The experimenter can welcome the user instead of a secretary. It is very good to state that the user is not being tested but the technology. Instruction before the test can contain clues, a way to fix that is to use a prerecorded instruction.

**4.2** Different ways of making the user comfortable: Allowing the user not to be interrupted and be the one who initiates contact to the team, this is natural but might but thinking aloud alone might feel weird. Another option is to smalltalk but again, this can lead to cues. 60-90 minutes with breaks

**4.3** Key role of experimenter is to make test user feel comfortable. Be friendly. The test user can be a software developer with a crash-course, but also a usability professional. Both can be good or bad at the job. The experimenter should have tried being a test user.

**4.4** When debriefing the experimenter should be open and not be defensive. Focus on the difficulties, the test user can state the least and most representative tasks, if it makes sense the developers can be allowed to talk to the user, if they are not merely defending their design.

## **5 Collecting and analysing data**

Data can be collected by taking notes or recording or something else. Data should only be recorded, if it is already known how the data will be analysed. Using notes can make the user uncomfortable because they worry about what has triggered writing the notes. Using video camera can also make the user uncomfortable. Video allows others to watch the test afterwards.

## **6 Communicating test results**

Communication of test results is done in a short report, so that developers will read it. You can also invite people to a short meeting. You have to communicate both good and bad things. Allow developers and others to contribute to the final report.

**6.1** A more expansive test report is written which can be more than 20 pages, where each problem is described in detail. It also has standard content like description of experiment and test users, conclusion and so on.

**6.2** Classification of problems should be done according to the severity and the time needed to fix them.

**7 Ethical considerations** How to make the user comfortable: Make clear that the data is confidential, that the user is not being tested, take a break, make realistic boring tasks

**9** A variant of the test is the “constructive interaction test” where two users are together, making it more natural to think aloud, but makes it harder to assess the way in which a single user conducts a task.