Usability Evaluation

Usability

Q: What is usability?
A: usability: ease with which ***people (users)*** can use a particular tool or object to achieve a specific ***goal (task)***

Q: Which aspects of usability are important (named at least three points)?
A: • aspects of usability:
– learnability: how easy to accomplish tasks the first time?
– efficiency: once learned, how quickly to complete tasks?
– memorability: how easy to reestablish proficiency after not having used a design for a period of time?
– errors: how many, how serious, how easy to recover?
– satisfaction: how pleasant to use the design?

Heuristic Evaluation

use of design principles/heuristics to inspect an interface for usability problems.
evaluation only by developers or experts.

General approach:
take the interface and check for the interface guidelines/heuristics.

Number of evaluations:
– single inspector
– multiple inspectors

Heuristic Evaluation: Individuals vs. Teams

• individual inspectors who look at an interface alone are recommended

• reasons:
– evaluation not influenced by others
– independent and unbiased
– greater variability in the kinds of errors found
– no overhead required to organize group meetings

• problem: some interfaces require groups, then use several independent groups

Self-Guided vs. Scenario Exploration

• self-guided exploration
– open-ended exploration
– not necessarily task-directed
– good for exploring diverse aspects of the interface

• scenario exploration
– step through interface using a number of representative end user tasks (remember task-centered design)
– ensures problems found in relevant parts of the interface

– ensures that specific features of interest are evaluated
– limits scope of evaluation – problems may be missed

Qualitative and Quantitative Evaluation

evaluation with ‘real’ people (users).

Ethics

• Don’t waste the user’s time
– use pilot tests to debug experiments, questionnaires etc
– have everything ready before the user shows up

• Make users feel comfortable
– emphasize that it is the system that is being tested, not the user
– acknowledge that the software may have problems
– let users know they can stop at any time

• Maintain privacy
– tell user that individual test results will be completely confidential
• Inform the user
– explain any monitoring that is being used
– answer all user’s questions (but avoid bias)

• Only use volunteers
– user must sign an informed consent form

• Inform the user
– answer particular questions about the experiment that could
have biased the results before

Qualitative vs. Quantitative Evaluation

• can test numeric or non-numeric criteria

• non-numeric (qualitative) criteria
– e.g., what techniques do people employ to reach a goal?
– e.g., what do people like or dislike about an interface?
– subjective opinions of people about interfaces

• numeric (quantitative) criteria
– e.g., how fast can someone achieve a goal?
– e.g., how precise are people in their interactions?
– e.g., how many mistakes do people make?
– quantitative comparison of techniques
– statistical evaluation: significance

Qualitative Evaluation Techniques

Qualitative Evaluation Approaches

• naturalistic approaches:
observe in realistic settings
– pro: real-life situations
– con: hard to arrange and do, time-consuming, may not generalize easily

• usability engineering approaches:
observe in simulated settings
– pro: situation can be better controlled, easier and cheaper to arrange and do, less time-consuming
– con: changed context for interface use

methods:
– user performance data collection
– controlled experiments

Collecting Performance Data

• people using a system (often lots of data)
• targeted data collection
– look for specific information, but may miss something
– e.g., frequency & type of request for online assistance
– e.g., frequency of use of different parts of the system
– e.g., number of errors and where they occurred
– e.g., time it takes to complete some operation
– all these tell you something about the usability

Controlled Experiments

Participants and Apparatus.

• What is controlled experiments?
– This is when a hypothesis is scientifically tested. In a controlled experiment, an independent variable (the cause) is systematically manipulated and the dependent variable (the effect) is measured.

• striving for
– removal of experimenter bias
– clear and testable hypothesis
– control of variables and conditions
– quantitative measurement
– replicability of experiment
– measurement of confidence in obtained results (statistics)

Experimental design - Participants

• Three experimental design:
– Between-subjects: different participants perform in different conditions.
  • No ordering or training effects.
  • large numbers of participants are needed.

– Within-subjects: all participants perform in all conditions
  • lessen the impact of individual differences and see how performance varies across conditions for each participant.
  • Need to ensure the order in which participants perform tasks for this setup does not bias the results (counterbalancing)

– Pair-wise: participants are matched in pairs based on certain user characteristics such as expertise and gender. Each pair is randomly allocated to each experimental condition.

Experimental design – Collected data

• The data collected to measure user performance on the tasks set in an experiment usually includes:
– Times to complete a task (efficiency)
– Number of errors per task (accuracy)

Clear and Testable Hypothesis

• A hypothesis involves examining a relationship between two things, called variables.
• Variables can be independent or dependent.

• independent variables:
variables that are to be altered, independent of the participants' behavior

• dependent variables:
variables that will be measured, depending on the participants' reactions to the independent variables in the experiment, included in hypothesis.

Qualitative Evaluation Techniques

• direct observations: observe people while they are using a system
• interviews: investigate specific issues
• continuous evaluations: monitor system in use

Direct Observations

1, Think Aloud Method
Q: Briefly describe the "think aloud" technique.
A:
person is asked to speak out their thoughts while doing a task, e.g.:
– what they are trying to do, why they took an action
– how they interpret what the system did

Q: Name an advantage and a disadvantage.
A:
advantages:
gives insight into what the person is thinking.

disvantages:
– may alter how people perform tasks
– unnatural (awkward & uncomfortable)
– hard to talk while concentrating

2, Constructive Interaction
• observe two people working together
– monitor their normal conversations
– removes awkwardness of think-aloud

• different version: co-discovery learning
– use semi-knowledgeable ‘coach’ and ‘novice’
– only novice uses the interface
  • novice to ask questions
  • coach responds

Interviews and Questionnaires

plan a set of central questions
– a few good questions to get things started
– avoid leading/suggestive questions (bias)
– focus on the interview
– could be based on results of user observations

semi-structured interview
– let user responses lead to follow-up questions
– follow interesting discussions

Continuous Evaluation

• monitor systems in actual use
– usually later stage of development
– fix problems in next release

• user feedback
– provide means of reporting feedback: help desks, bulletin boards, e-mail, etc.
– combine with trouble-shooting facilities

• case/field studies
– careful study of “system usage” on-site
– seeing system in “real-life” use
– monitor through external observers and site visits

Qualitative Evaluation Techniques

qualitative techniques to evaluate non-numeric aspects of interfaces
a number of different techniques for evaluation
often combinations of techniques employed:
– e.g., questionnaire with open-ended interviewing
– e.g., single-person or team direct observation and think-aloud with co-discovery learning

results often not backed with statistics and may be biased (reduce as much as possible), but yield important insights into what people think
more general issues discovered