User Evaluation

HCI Research Group
Department of Computer Science

Types of Evaluation

- We have already discussed how there are expert evaluations and user evaluations – each serve their purposes
- There is another dimension that we can thinking about when doing evaluations
  - Formative evaluations
  - Summative evaluations

Formative Evaluations

- *Formative evaluations*
  - Quick evaluations intended to pick up on issues in design early in the process
  - Conducted in service of design – with the intention of improving the design going forward through its lifecycle
  - Expert evaluations can be used as one type of formative evaluation
  - Where possible also useful to have user evaluations at different levels of prototypes

Summative Evaluations

- Summative evaluations are usually final stage prototypes (final products) where we are conducting “disaster checks”
- We can be picking up on some problems – but usually if there are major problems at this point, we will be in trouble in terms of making changes
- Solely the realm of user evaluations – experts should have seen the prototypes well before summative testing beings
- Primarily geared at collecting data about real use of the product

User Evaluations

- User evaluations are our gold standard of evaluation
- Users behave in interesting and often unexpected ways – not everything can be anticipated by good design or through expert review
- When we have a sufficiently concrete prototype, we need to bring in users to undertake tasks

User Evaluations on Low Fidelity Prototypes

- Requirements for a user evaluation
  - A sufficiently detailed prototype, preferably with some depth for a particular user journey
  - Define a set of tasks, taken from your scenarios for people to undertake on your prototype
  - Recruit participants from the target population – using personas as a guide
- Have participants perform tasks on the prototype using a *concurrent verbal protocol* (think-aloud protocol)
- As users take actions in the prototype (or declare they want to take actions) the evaluator plays the part of the computer moving the user along

User Evaluations on Low Fidelity Prototypes (2)

- User undertakes tasks in the interface
  - In low fidelity prototypes this may include simply stating what they would enter into text fields
- User speaks aloud what they are thinking as they work through the interface
  - This can include describing where they are trying to make decisions, what they interpret different interface cues to be or other forms of feedback
- When users encounter problems, they stop and report the problem to the evaluator, giving it a rating similar to the Nielsen ratings from heuristic evaluations

User Evaluations on Low Fidelity Prototypes (3)

- Conducting an evaluation
  - Ensure that you have the tasks and paths through the prototype worked out clearly in advance
  - Obtain *informed consent* from the users to participant – includes what they will do, understanding any risks, how long the evaluation will take, where data will be used and how it will be used
  - Provide a quiet, organised space where you can have the user sit comfortably with the prototype
  - Have users undertake tasks using a think-aloud protocol
  - Record notes of particular areas of problems, and after tasks are done, return to those areas and ask users about the issues they were having
  - At the end of the session provide a debriefing session regarding the evaluation and what will happen next with the results

User Evaluations on High Fidelity Prototypes

- Largely the same as with low-fidelity, but you have a few more things you can do
- User journeys, look and feel and overall functionality will be more complete – but you should still know the routes for completing various tasks
- Concurrent verbal protocols are still possible for collecting information about problems
- Alternatively, can collect quantitative information about specific usability criteria
  - Time to completion, time per page, number of errors, log files (careful with these!), proportion of tasks completed, time to 95% error free performance, proportion of tasks completed on re-engaging after being away from application, personal perceptions/preferences of system

Demonstration

Readings

- Rogers, Preece, Sharp, 3rd ed. Chapter 12 and Chapter 14 cover Evaluation of interactive systems
- Cooper et al Chapter 7 provides a good overview of the entire lifecycle of interactive design, with the last section being about evaluation.