Usability Evaluation

or, “I can’t figure this out...do I still get the donuts?”
Two types of evaluation

- user-based evaluation: involves users other than the designers
- expert evaluation: checking the design on paper, groups of designers or UI experts
- both important
  - user-based: reveals unexpected problems
  - expert: keeps design on track
Purposes of Evaluation

- which design is better?
- are there any problems with the design as it now stands?
- does the design meet usability targets?
When do you evaluate?

- **formative evaluation**
  - helps with form of the solution
  - deciding between competing designs
  - at intermediate design stages—for iteration

- **summative evaluation**
  - a summary of usability
  - often compares new design with existing or alternative solutions
  - when design is complete
What do you evaluate?

- prototypes
  - during design iterations
  - can be very low-fidelity
  - “wizard of Oz” technique

- working systems
  - system can be prototype
  - can also evaluate at end of design cycle
Types of Evaluation

- range of formality
- informal user studies
- usability studies
- formal experiments
User studies

- early stages of design
- usually only a few users
- non-structured tasks
- collect comments, observations, suggestions, ...

- ex: VR manipulation
Usability studies

- usually more complete prototype
- structured tasks
- collect timings, errors, verbal protocol, ...

issues:
  - finding users
  - testing environment
  - compensation

ex: Virtual Habitat
Formal experiments

- working system or piece of system
- # users determined by desired difference
- tightly controlled tasks
- precise measurements
- statistical analysis of hypotheses

ex: VR manipulation
Components of a study

- Informed consent
- User familiarization
- User questionnaire
- Background testing
- Pre-testing
- System exploration
- Specific tasks
- Post-testing
- Post-questionnaire
- Debriefing
Informed consent

- Form advising users of their rights
- Tell them you are studying the system, not them!
- Tell them if there are any known risks
- Tell them they can stop at any time
- Get signature, give them a copy
User familiarization

- Make the user comfortable
- Tell them what you’re doing without giving everything away
- Show them the facilities and equipment
- Give them written instructions
- Tell them approximately how much time the evaluation will take
- Ask if they have any questions
User questionnaire

- Obtain demographic information
  - Age
  - Gender
  - Occupation or major
  - Computer experience
  - Domain experience
  - Eyesight
  - Handedness
Background testing

- You may want to correlate your findings with some standardized test score

- Examples
  - Spatial ability tests
  - Memorization tests
  - Domain knowledge tests

- Example result: “Users with high spatial ability preferred interface X, while others preferred interface Y”
Pre- and post-testing

- Some systems have a specific learning goal (i.e. education or training)
- Giving users the same test or type of test both before and after usage of the system is a way to measure learning
- Only used with very well-developed prototypes
System exploration

- User gets “free play” time with the system
- You can observe what they seem to understand easily and what is troublesome
- See if they “find” all the features or parts of the system
- Good with verbal protocol
Specific tasks

- Give the user a specific goal
  Ex: Find out what school Dr. Bowman went to for undergraduate studies using these web pages
- Observe problems
- Record performance, errors
Post-questionnaire

- Get user’s reaction to the system
- Subjective levels of satisfaction, perceived ease of use, usefulness, etc.
Debriefing

- Talk to the users about the session
- Assure them they did well
- Give more details about what you’re doing if they are interested
- Thank them
- Give them donuts 😊
Other evaluation issues

- Recruiting users
- Testing facilities
- Measurements and observations
- Compensation
Recruiting users

- Attempt to match the proposed user population
- Perhaps divide into several user groups
- Techniques
  - Posted advertisements
  - Internet/email/newsgroup
  - Colleagues/friends/classmates
  - People already using existing system
- Make sure your users are not overly knowledgeable!
Testing facilities

- For informal studies, a simple setup
- In all cases, privacy is important
- Replicate the usage environment?
- Evaluator present or not?
- More formal studies may use a special usability lab (e.g. McBryde 102)
Usability lab

- One-way glass
- Omni-directional mics
- Remote controlled cameras
- Audio/video recording equipment
- Configurable computing equipment
Measurements & observations

- User comments
- General observations
- Specific “critical incidents”
- Task timing
- User errors
Session management

- May need multiple evaluators
- Use checksheets or pre-printed tables for filling in results
- Video/audio as backup if something is missed
- Don’t ask the user to stop while you catch up!
Compensation

- Most studies give the user something for their time and effort
- Doesn’t have to be monetary – can also be:
  - Food
  - Extra credit in a class
  - Special discounts on the company’s products
  - Tour of your facility
  - ...