CMPE 235: User Evaluation of Technology

Preparing for User Evaluation

Course contact: srikur@soe.ucsc.edu

The goals of user evaluation

► To improve aspects of the product with which users come into contact
► To make user interfaces logical, intuitive and clear to people who use them
► Includes the entire user experience
  ▪ physical design (dimensions, layout, packaging)
  ▪ look and feel (displays, controls)
  ▪ procedures (the steps to complete a task)
  ▪ environment (physical, geographic)
  ▪ availability/usefulness of support (online help, training, documentation, help line support)

User evaluation

► Different types, can be complementary
  ▪ Testing with users
  ▪ Heuristic / expert evaluation
  ▪ Online survey
  ▪ Click-path monitoring
  ▪ Concept evaluations

► Timing
  ▪ Formative: early in the development cycle, usually qualitative and open-ended
  ▪ Summative: almost at the end of the development cycle, usually quantitative and had very specific measures.

Planning

► What are the goals of user evaluation?
  ▪ Is the product about to be deployed and is there only time for “tweaking”
  ▪ What if there are show-stoppers?

► Who are the users?
  ▪ Are there multiple user groups?

► How many users should be tested?
  ▪ Do we need to test users from all groups?
  ▪ What type of test is this? (i.e., do we need quantitative, random sampling based data? Do we need to turn around the results in two days?)

► What or who determines if the product is usable?

Planning and Preparation

► What should be evaluated?
  ▪ Do we go with typical tasks?
  ▪ Are all functions created equal?
  ▪ Do anticipated environmental situations affect usability?

► What happens to the information that will be generated?
  ▪ Will changes be made to the product based on the data?

► Materials needed for testing and test plan
  ▪ Screener to “qualify” participants, test plan
  ▪ Consent form and other paper forms (pre/post-test surveys? Sign-up sheet for remuneration?)
  ▪ Tasks for participants (enough to fill the time, and a little more)
Preparing for User Evaluation

1. Developing a test plan ("blueprint" for the test)
   - General intention and individual checks
     ▶ MUST BE PRECISE
     ▶ MUST BE CLEAR
     ▶ MUST BE MEASURABLE/OBSERVABLE
   - Specifies the resources needed to carry out test
   - Governs conclusions that can be made

2. Developing test materials
3. Recruiting participants
4. Carrying out the test
5. Debriefing participants
6. Findings and recommendations

Test plan

1. Purpose of test
2. Problem statements/test objectives
3. Participant profile (inclusion/exclusion criteria)
4. Method/technique to be used
5. List of tasks to be used
6. Test environment (field vs. lab) and material (HW/SW, resources → recorder and batteries, report forms, questionnaires)
7. Experimenter’s role (monitor, coach etc.)
8. Evaluation measures to be taken (qualitative vs. quantitative, subjective vs. objective)
9. Contents of report to be produced and how the report is going to be presented → focus group, informal meeting, big bosses are going to be there

A Good Test Plan Contains...

▶ Objectives of the study broken up into specific questions that the study has been designed to answer.
▶ Description of the stimuli, including version or model numbers where appropriate.
▶ Description of the participants with an indication of strict requirements and allowable variances.

A Good Test Plan Contains...

▶ Description of the methodology, including:
  - The procedure (e.g., use of a think-aloud protocol and within- or between-groups study design)
  - Measures used (e.g., time on task, ratings)

A Good Test Plan Contains...

▶ Requirements for equipment needed for testing (e.g., size of TV or computer screen, computer operating system, and types of cameras – face shot, desktop shot etc.).
▶ Pictures/sketches or description of the lab setup denoting how the stimuli should be arranged, where the moderator and participant should sit, and where the camera(s) should be set up.
▶ Esp. beneficial if complex setup with multiple elements.

A Good Test Plan Contains...

▶ The project logistics, including:
  - Test dates and times
  - Test location(s)
  - Deliverable deadlines for the stakeholders and the research team
  - Etc.
Participant Recruitment

► Probably the hardest part of user evaluation
  ▪ The arts of selling the study
  ▪ Success rate can be low, especially for special population
  ▪ Know when to give up
► Key contacts are key!
► Alternative recruitment venues (barber, bus stop, Laundromat, church)
► Active vs. passive recruitment
  ▪ Passive: flyer, mass mails, media ads
  ▪ Active: phone, bring/introduce a friend, outreach/presentation

Preparation: Screener

► Based on a description of user group(s)
  ▪ Determine which characteristics are important for choosing participants
  ▪ Determine what should be controlled or balanced
► Use the screener to verify participant is willing to be video taped or photographed, if appropriate
► Screening call = scheduling call, too
  ▪ Participants who are "no-shows" are costly so plan on getting contact detail for reminders
► Have a script and checklist before dialing

What can help recruitment

► Know your participants
  ▪ How to get to appointments
  ▪ Frequency and timing of appointments
  ▪ Total length of study
► Be flexible if possible
  ▪ home visits!
  ▪ phone calls and follow up
  ▪ evening and weekend appointments
► Show appreciation
  ▪ Doesn't have to be money (food, movie ticket, altruistic)
  ▪ A draw of something big is better than small change → give the odds of winning
  ▪ Completion bonus

Preparation

► Consent form
  ▪ Is it legally necessary to receive written permission from a participant to video or audio tape them?
  ▪ You need the participant’s agreement to use their image
  ▪ Some clients need written non-disclosure statements from participants
► Pilot test → with unfamiliar participants
  ▪ To assure that the task wording is clear
  ▪ To determine how long it takes to accomplish the tasks
  ▪ To determine if the tasks are "do-able"
  ▪ To find if there are alternate "correct" paths

Preparation: Test Plan

► A script for the administrator
  ▪ Gives place for checklist of what to say
  ▪ Can give precise wording if study needs to adhere to more precision
  ▪ Gives the anticipated “correct” path for ease of the administrator to follow
  ▪ So that everybody does it exactly like the script said
  ▪ Easier when the order of tasks are different for each participant
► Because it is easy to forget things
  ▪ Starting the tape or recording software
  ▪ Debrief, thank, compensate cycle

Report

► User evaluation is usually documented
  ▪ To have the appropriate effect
  ▪ To allow the powers-that-be to make appropriate decisions
  ▪ To serve the in the historical context of a project
► Basic information about participants and evaluation
► Results and analysis → prioritize problem severity
  ▪ High--cannot complete the task or causes data loss
  ▪ Medium--likely to cost user time or difficulty
  ▪ Low--not substantial but worth considering
► Recommendations, including further testing if problems are severe or complicated
Sample results from a usability test
► Overall, 4 of 5 users liked the product
► 3 of 5 users cannot complete Task 3 within the allotted time.
► The average time to finish Task 4 was 3.4s (S.D. = 0.34s).
► The three users that had difficulty selected the blue button instead of the red button because the names are ambiguous and confusing. It is recommended that one or both of these names change to reflect their opposing effects.
► 5 of 5 users completed Task 6. Users commented that they liked this feature and found it easy to complete the task. Average rating was 8.7 on a 9 point scale.

The need for IRB approval
► National Research Act of 1974
  ▪ Required establishment of IRBs at institutions receiving the Department of Health and Human Services funding for human subjects research
  ▪ Respect for Persons: Researchers must treat individuals as autonomous human beings, capable of making their own decision/choices, and not use people as a means to an end.
  ▪ Beneficence: Researchers should minimize the risks of harm and maximize the potential benefits of their work.
  ▪ Justice: Researchers must treat people fairly and design research so that its burdens and benefits are shared equitably.

Respect for Person
► Protect those with limited autonomy in terms of:
  ▪ Mental capacity (the ability to understand and process information)
  ▪ Voluntariness (freedom from the control or influence of others) → can withdraw from research without coercion or undue influence from others.
► Implementations:
  ▪ The req. to obtain and document informed consent.
  ▪ The req. to respect the privacy interests of research subjects.
  ▪ Much more strict procedure for obtaining IRB when conducting research on individuals with limited autonomy.

Beneficence
► Potential benefits should justify/outweigh the risks of harm, and definitely not when death/disab. can occur
► Requires evaluating magnitude and likelihood of harm
  ▪ Harms include physical, psychological, legal, social, and economic harms.
  ▪ "benefit" = positive value to individual subjects or to others, such as a community, or humanity as a whole.
► Implementations
  ▪ The req. to use procedures that present the least risk to subjects consistent with answering the scientific question
  ▪ The req. to gather data from procedures or activities that are already being performed for non-research reasons.
  ▪ The req. to maintain promises of confidentiality and monitor data for the safety of individuals

Justice
► Those who benefit from the research should share in the burden of being subjects in the research.
► Those who serve as subjects in the research should share in the potential benefits from the research.
  ▪ Compensation section in IRB, potential benefit to subjects.
► Individuals or groups should not be selected for research participation solely because they are available, cannot say no or do not know that they can say no.
  ▪ Subject selection should be based on scientific justification
  ▪ That's why IRB requires inclusion criteria, recruitment ads, and places where subjects are recruited from.

Some basic definitions:
► Research: a systematic investigation designed to discover or contribute to a body of generalizable knowledge.
► Research participant: a living individual about whom a researcher obtains either:
  1. data through intervention or interaction with the person
  2. identifiable private information
► Research requiring IRB:
  ▪ Any study intended to result in publication or public presentation, including classroom projects.
  ▪ Any activity resulting in publication or public presentation, even though it involves only review of existing data that was collected with no intent to publish
  ▪ Any use of an investigational drug or device.
### Vulnerable populations

By Federal definition, the populations that require special protection include:

- Children
- Persons with diminished capacity to consent
- Prisoners
- Fetuses and pregnant women
- Terminally ill persons
- Students or employees
- Comatose patients

### The Institutional Review Board

- Mandated for all institutions conducting human research.

- Composition of Board, functions of board, reporting requirements for board are all mandated by Office of Human Research Protection (OHRP).

- According to Common Rule, the IRB must
  - Be representative of community
  - Must have at least five members
  - Include one scientist
  - Include one nonscientist
  - Include more than one profession
  - Include one person not affiliated with the institution.

### Role and Responsibility of IRB

- Review research plan to be sure it meets criteria in Federal regulations (45CFR 46.111)

- Confirm there are no unreasonable risks

- Conduct continuing review

- Assess suspected or alleged protocol violations.

- The IRB has authority to:
  - Approve, disapprove, or terminate all research.
  - Require modifications to protocols.
  - Require that information the IRB deems necessary is provided to participants.
  - Require documentation of informed consent, or allow waiver of consent.
  - Check that all experimenters (or advisor in the case of student experimenter) completed ethics training.

### Types of IRB reviews

1. **Quorum/full board review**
   - Review of a protocol by a quorum of IRB members attending the monthly IRB meeting.
   - Necessary for research involving risk of physical or psychological harm greater than that encountered in daily life, particularly research involving deception, stress, or manipulation

2. **Expedited review (1-2 members)**
   - Collection of data through noninvasive procedures.
   - Research involving materials for nonresearch purposes or continuing research of quorum review.
   - Voice, video, digital, or image records made for research purposes.

3. **Administrative review**
   - Exempt research does not require expedited or quorum review by the IRB, but it does require "exemption approval" at the institution level.
   - Research involving the collection or study of existing data, including documents, records, not collected in a prospective fashion (e.g. the data must exist before the research is initiated)
   - Research conducted in accepted educational settings
   - Research involving only the use of educational (diagnostic, aptitude, or achievement) tests.
   - Research involving only observation of public behavior
   - Research involving only taste and food quality evaluations.
   - Research involving only surveys or interviews
   - Above: only for non-vulnerable populations

### Protocol

- Protocol statement
  - What the study is about: investigators, dates, reasons, review category, location, funding source, payment, risks
  - Items of special concern: filming, deception of subjects, use of placebo, biohazard

- All forms:
  - Consent and assent (for research w. 7-17 years old) forms
  - All other assessment forms
  - Recruitment ads (has to be approved before publishing)
  - Consent forms must be reviewed and approved by the IRB before approaching potential subjects.
  - Must be written in lay language at ~Grade 8 level.
Protocol

▶ Discussion
  ▪ Confidentiality – how confidentiality will be maintained for records and when records will be destroyed.
  ▪ Qualifications of all investigators
  ▪ Role of any other participants listed in protocol.

▶ Abstract: written in lay person language

▶ Participant inclusion criteria:
  ▪ The proposed number of subjects, criteria and methods for recruiting, selecting, and excluding subjects.

Internet-based research

▶ Ways to obtain informed consent / assent and protecting participants include:
  ▪ Setting up a page that contains the required consent/assent form as a front page for study.
  ▪ This consent/assent page (page ‘a’) should indicate that by clicking on a link from page ‘a’ to page ‘b’, subjects are consenting to participate.
  ▪ Page ‘a’ should also include an e-mail address in addition to a telephone number(s) to withdraw consent and remove data, to the extent possible, upon request of the respondent.

▶ Protocol statement should include information on how contact information were obtained.

HIPAA

▶ HIPAA: Health Insurance Portability and Accountability Act of 1996
  ▪ Objective of Act is to protect privacy of medical records.
  ▪ Required to be in place on April 14, 2003
  ▪ Concerned mostly with medical information and is particularly important in clinical trials and in retrospective studies of medical records (but weight/height is in there)

Responsibility of Investigators

▶ Conduct the project as approved by the IRB
▶ Promptly report any revisions or amendments for review and approval by the IRB prior to commencement of the revised protocol.
▶ Promptly report any unanticipated problems involving risks to subjects or others to the IRB.
  ▪ Serious: must be reported in writing within 24 hours
  ▪ Others: must be reported in writing within 72 hours
▶ Request an extension prior to the expiration date if data collection is not complete.
▶ For student investigators, to keep the faculty sponsor in the loop.