

P6: Pilot Usability Test

Due 3 May 2013, 11:59 PM

Overview

The goal of this project step is to learn to perform a simple usability test with your working prototype, in order to evaluate your prototype and to inform a last iteration of changes before submitting your final project report on Dean's Date. In the real world, this kind of "pilot" study would be used to redesign your experiment before running the study with a larger pool of participants.

The Assignment

1. Refine your 3 benchmark tasks, if necessary. You should choose 1 easy task, 1 medium task, and 1 difficult task that allow you to evaluate your prototype in a meaningful way. These tasks don't have to be the same ones you used in your lo-fi prototype study, and they don't have to enable you to evaluate all aspects of your interface. However, they should allow you to obtain useful insight into how well your proposed idea (or some significant component of it) will work, and into how your current prototype could be improved as you move toward a final product. In many cases, these will be the same 3 tasks you described in P5.

2. Recruit volunteer participants and plan how you will obtain consent and demographic information. You will need to find three unique participants (volunteers who are not in the class and who aren't your friends) to work through your benchmark tasks. These subjects should **not** be participants who participated in your lo-fi evaluation. (If there is a compelling reason to re-use participants, please discuss with the instructors first by posting a private message on Piazza.)

Participation must be voluntary. You should get the participants to sign an informed consent form. If you plan to use names, photos, and/or video of your subjects in your documentation of the usability test, be sure to ask for (and receive) consent first. You should also employ a short paper questionnaire to obtain other demographic information (age, sex, education level, major, experience with your type of tasks and application, etc.).

3. Think about what you want to evaluate in your usability test.

Decide which dependent variables you will want to measure. Task time? Number of errors? Self-reported measures of satisfaction, confusion, or other relevant phenomena? Although it will be hard to get statistically significant bottom-line data with only three participants and a rough prototype, you should measure some important dependent variables to get a feel for how this is done.

In most cases, we recommend focusing on *process data*. You will probably want to instruct your participant to think aloud as they perform the tasks. You should make a log of critical incidents (both positive and negative events). For example, the user might make a mistake, or they might see something they like and say "cool." Set up a clock that only the observers can see (one or more of you should be only observing during the evaluation, not wizard-of-oz-ing/sleeping/etc.), and write down a log containing the time and nature of each critical incident.

If you have access to a video camera, it is fine to use it (as long as you have explicit consent). Note the time you start taping so that you can find your critical incidents later on tape. You might use an audio recorder if you don't have access to a video camera.

If there is data that would be best obtained through a questionnaire and/or interview, create the questionnaire and/or interview script following the guidelines discussed in class. We recommend sharing your questionnaire with one or more students in the class (not on your team) for them to catch any mistakes, to check that it is not too long or confusing, etc., before using it with a real participant.

3. Run the evaluation with each participant.

Procedure for each participant:

- A. Obtain informed consent.
- B. Give the participant a short demo of your working prototype. Do not show them exactly how to perform your tasks. Just show them how the system works in general, and give an example of something specific that is different enough from your benchmark tasks. You should write up a script of your demo and follow the same script with each participant. Make sure to prompt them for any questions, and answer any questions they have before proceeding.
- C. Then, give the participant task directions for the first task, telling them what they are trying to achieve, **not** how to do it. If you are using a think-aloud study, be sure to explain to participants how the think-aloud protocol works (see Resources section below).
- D. When they are finished, you will give them the directions for the next task, and so on. Each participant will perform all 3 tasks. You will want to keep the data separate for each task and participant.
- E. Do any post-survey questionnaire and/or interview with the subject. Answer any final questions the subject might have.

4. Analyze your results.

You must report your results (values of dependent variables, summary statistics, and summaries of the process data) in your report, and in your “Discussion” section (see submission instructions below), you will use this data to draw some preliminary conclusions with respect to your prototype. You should think about how your system should change if those results were found to hold with a larger user population. This will be the most important part of the write-up. We want to understand how you would fix or change your system as a result of what you observed.

Submission Instructions

1. Create a new blog post on the course blog.
2. **Add this blog post to category Project6.** (We will take off points for not doing this.)
3. Include on your blog, in order, and with section headings:
 - a. Your **group** number and name.
 - b. First **names** of everyone in your group.
 - c. A 1-sentence project **summary**.
 - d. **Introduction (5 points)**. 1 paragraph. Introduce the system being evaluated and state the purpose and rationale of the experiment.
 - e. **Implementation and Improvements (5 points)**. Provide a link to your P5 submission. If you have changed your working prototype at all since submitting P5, supply a brief bulleted list of the changes made since P5. (It is not necessary to change your prototype from P5 before doing P6.) This section should be no more than 1 paragraph.
 - f. **Method (10 points)**.
 - i. **Participants**: 1 paragraph. (Who they are, how they were selected.)
 - ii. **Apparatus**: 1 paragraph. (What equipment did you use, where did you conduct the test.)
 - iii. **Tasks**. ~1/2 page. Describe the tasks you have chosen to support in this working prototype (3 short descriptions, 2–3 sentences each; should be one easy, one medium, one hard). If you have not changed the tasks from P5, you can re-use your text from P5 here. Otherwise, if you have changed the tasks, explain how and why. In any case, explain why you have chosen these tasks.
 - iv. **Procedure**. 1 paragraph. Describe how you conducted the study.
 - g. **Test Measures (5 points)**. Describe what you measured and why. Bullet points are encouraged.
 - h. **Results and Discussion (25 points)**. ~4 paragraphs. Provide results of your tests. Describe what you learned from the user study. Document any changes that you plan to make in your prototype as a result of the study.
 - i. **Appendices (5 points)**.
 - i. Provide all things you read or handed to the participant: consent form, demographic questionnaire, demo script, post-task questionnaire and/or interview script.

- ii. Also provide raw data (i.e., your merged critical incident logs, questionnaire responses, etc.)
4. Fill out the Google Form at <http://tinyurl.com/cos436P6> with the URL of your blog.

Grading

Your grade will be based on the quality of your writeup, with points allotted to each section according to the outline above. A substantial portion of your grade will be based on the design and execution of your experiment, especially on your ability to measure variables that are meaningful in assessing the quality of your design and useful in pointing to areas for future improvement (this information will be found in the Introduction, Implementation, Method, and Test Measures sections). Another substantial portion of your grade will be based on your thoughtful analysis and clear presentation of the study results (the Results and Discussion section). As always, your grade will also reflect the completeness, conciseness, and quality of writing.

Resources

More about the think-aloud protocol:

A brief summary and instructions: <http://www.hu.mtu.edu/~njcarpen/hu3120/pdfs/thinkaloud.pdf>

See also section 5.5 of *Task-Centered User Interface Design* by Lewis & Rieman, available online:

http://grouplab.cpsc.ualgary.ca/saul/hci_topics/tcsd-book/chap-5_v-1.txt

Acknowledgements

Much of this assignment is adapted from Manish Agrawala's CS 160 course at UC Berkeley.