

Field Usability Study (Team)

Due: Tuesday, February 21, 2017(start of class)

Goals

The goal of this assignment is to learn how to perform a field usability study and to incorporate the results of the study into design changes in your hi-fi prototype. In practice, this “pilot” study would be used to redesign your experiment before running the study with a larger pool of participants.

Prototypes

You will be performing this field test using the prototype you produced on the last assignment after you’ve fixed any issues that you know you wanted to change, implemented missing features, and made changes the teaching staff highlighted.

Participants

You will find at least five participants (i.e., volunteers who are not in this class) to work through your benchmark tasks (***the more you can run the better***). Remember, it must be voluntary. You should get the participants to *sign an [informed consent form](#)* and obtain other demographic information (e.g., age, sex, education level, major, experience with your type of tasks & application, etc.). Make sure they are *representative* of your target customers and that they are not participants that you have used previously during this project (unless you have pre-approval from the teaching staff).

Benchmark Tasks

Your test will use the tasks that you have been using this quarter (unless we asked you to update or you decided to make changes). They should include at least 1 simple task, 1 moderate task, and 1 complex task. These tasks should give good coverage of your UI at this point; if they don’t you better redesign them.

Measures and Observations

Although it will be hard to get statistically significant bottom-line data with only five participants and a prototype, you should measure and report on some important dependent variables to get a feel for how it is done (i.e., task time, # of errors, etc.).

You will *concentrate on process data* in your presentation. For example, you should instruct your participant to *think aloud (if co-located)*. You should make a *log of critical incidents* (both positive and negative events). For example, the user might make a mistake and you notice it or they might see something they like and say “cool.” When a critical incident occurs, write down in the log what happened and the time.

You should also record the test using a digital video camera or a phone – note the time that you start recording so that you can find your critical incidents later – as well as logging software (see below).

Procedure

You will give the participant a short demo of the system. Do not show them exactly how to perform your tasks. Just show how the system works in general and give an example of something specific that is different enough from your benchmark tasks (if necessary at all). You should write-up a script of your demo and follow the same script with each participant. **If your application is so limited in terms of what is possible for a user to do that this would bias the test considerably, do not demo your application.**

The participant will then be given task directions (e.g., on a single card) for the first task that *tells them what they are trying to achieve, not how to do it*. Make sure that you have given them enough scenario information (e.g., who they are, who their friends are, etc.) to put them into the right context for the tasks.

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.

Location

This study will be carried out in a **field setting**. That is, if your application is meant to be used outdoors, then you will test it outdoors. If your application is supposed to be used in a user's home, then you will test it in their home. If its use would happen in a public place, you should test it in a public place. If it is a combination of settings, try in different types of places. We want to see the tradeoffs of testing in a naturalistic setting.

Video and Clickstream Logging

We'd like you to integrate video and clickstream logging into your app. Use a tool that accomplishes this for you. For the iPhone and Android, Lookback (<https://lookback.io/>) seems to be the best. Other possibilities are Watchsend (<https://watchsend.com/>) for iPhone and Validately (<https://validately.com/>) for web apps. Make sure to report on this data and how it was helpful or not helpful in your talk.

Results

You must report your results (values of dependent variables, summaries of those values, summaries of the process data, and summaries of clickstream data). You should draw some conclusions with respect to your interface prototype. ***You should also say how your system should change if those results hold with a larger user population.*** We want to understand how you would fix your system as a result of what you observed.

Deliverables

Presentation

One member of your team will present your project during an **10-minute** slide-based presentation in front of the other students and teaching staff. See the grading guidelines for information on how to structure your talk. You must **make the slides available for download on your web site.**

Presentation Guidelines

You will have **10** minutes for this presentation plus up to 5 minutes for questions. Please practice as we will grade you on how close you are to the time limit (under and over). All team members are expected to work collaboratively on the presentation, though only one team member will deliver the presentation.

Talk Outline:

1. Project title & team (introduce yourselves)
2. Introduction
 - Introduce the system being evaluated (1 slide) & the purpose of the experiment (1 slide)
3. Outline of Talk
4. Method
 - Participants (who – demographics – and how were they **selected & compensated**) (1 slide)
 - Apparatus (describe the equipment you used and the environment where tested) (1 slide)
 - Tasks (1 slide each) - describe each task & **what you looked for** during each task
 - Procedure (1 slide) - describe what you did and how
5. Test Measures (1 slide) - describe what you measured and why
6. Results (multiple slides) - Results of the tests (include lots of images & data including clickstream)
7. Discussion (multiple slides)
 - what you learned from the pilot run what you might change for the “real” experiment
 - what **you might change in your interface from these results alone**
 - difficulties of running the experiment *in situ*? Any leanings particular to testing it in the wild?
8. Summary of talk.
 - What did you do? What were the key things you learned?

Grading Criteria

Your grade will be based on the thoroughness of your experimental design, the analysis of your results, and the quality of your presentation. The presentation grading will be broken into two components: the individual grade of the presenter and a group grade for the quality of the content itself. Each bullet/grading category below will be out of 4 points (4=+, 3=check+, 2=check, 1=check-, 0=missing).

Presenter's grades (NAME: _____)

- Organization
 - ____ Project & team introduction
 - ____ Introduction to the system being evaluated & the purpose of the experiment
 - ____ Overview/Outline of talk (1 slide) – don't read this, tell it like a story
 - ____ Participants
 - ____ Apparatus & Environment
 - ____ Task & what looked for
 - ____ Procedure
 - ____ Test Measures
 - ____ Results
 - ____ Discussion
 - ____ Summary of talk
- Presentation
 - ____ Use effective slides (easy to read, understand, good use of visuals/images)
 - ____ Cover required scope in 10 mins (+ 5 minutes Q&A). Practice in advance.
 - ____ Ensure the presenter makes eye contact and projects well. (__ : __)

Group grade (GROUP NAME: _____)

- Introduction:
 - ____ clear what the system is for?
 - ____ valid reasons for running an experiment?
- Participants
 - ____ good range? demographics?
 - ____ recruiting/compensation clear?
- Apparatus & Environment
 - ____ clear on where the experiment took place and what equipment/software used?
 - ____ appropriate location?
- Tasks & Prototype
 - ____ good range?
 - ____ clear on what you were looking for?
 - ____ prototype implemented sufficiently to test?
- Procedure
 - ____ clear on what participants needed to do?
- Test Measures?
 - ____ appropriate variables recorded/measured?
- Results?
 - ____ clear if participants succeeded?
 - ____ clear on what worked and what did not, and why?
 - ____ clear on what was learned from logged clickstream data?
- Discussion?
 - ____ recommendations for a larger study appropriate
 - ____ recommendations for design changes based on the results?
 - ____ recommendations for design changes good ideas?
 - ____ issues / advantages of running the experiment *in situ*?