Lab Usability Study (Team)

Due: **Tuesday, January 31, 2017 (1 hour before class)**

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

Prototypes
You will perform this test using last quarter's prototype after you've fixed any issues that you know you needed to change (i.e., implemented key missing features), and addressed issues the teaching staff raised.

Participants
You will find three-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 worked on last 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 interface at this point; if they don't you better redesign them in advance. Make sure they are not fragmented tasks and represent real things people want to accomplish. **If you are developing a non-CS147 project or making a major change to your project, run your proposed tasks by the teaching staff well in advance.**

Measures and Observations
Although it will be hard to get statistically significant bottom-line data with only three-five participants and a rough 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*. 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." Set up a clock (phone) that only the observers can see (one or more of you should observe), and when a critical incident occurs, write down in the log what happened and the time.

You should also record the test using a digital camera or a phone – note the time that you start recording so that you can find your critical incidents later in the recording.

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 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 lab setting or other quiet setting (e.g., an office, apartment, or dorm room). This will make it easier for you to record the test participant's reactions and use of your application. In a later assignment you will do a field study where you will evaluate in a more 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.

Results
You must report your results (values of dependent variables, summaries of those values, and summaries of the process 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 a **twelve**-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 team web site**.

Presentation Guidelines
You will have **twelve 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. Prototype Changes
    - Major changes you made to your prototype in preparation for the test (1-3 slides w/ images)
3. Method
    - Participants (who -- demographics -- and how were they **selected** & **compensated**) (1 slide)
    - Apparatus (describe the equipment/software you used and where tests run) (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
4. Test Measures (1 slide)
    - describe what you measured and why
5. Results (multiple slides)
    - Results of the tests (process data, quantitative data, and logged data)
6. 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**
7. Summary of talk.
    - What did you do? What were the key things you learned?

CS 194H Winter 2017 web site
http://hci.stanford.edu/courses/cs194h

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
    - ___ Prototype Changes
    - ___ 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 12 mins (+ 5 minutes Q&A). Practice in advance.
    - ___ Ensure the presenter makes eye contact and projects well. (__ : __)

Group grade   (GROUP NAME: _____)

- Participants
    - ___ good range? demographics?
    - ___ recruiting/compensation clear?
- Apparatus & Environment
    - ___ clear on where the experiment took place and what equipment/software used?
    - ___ appropriate location?
- Tasks
    - ___ good range & real tasks?
    - ___ clear on what you were looking for?
- 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?
    - ___ good use of logged video & clickstream data?
- Discussion?
    - ___ recommendations for a larger study appropriate
    - ___ recommendations for design changes based on the results?
    - ___ recommendations for design changes good ideas?