CS160 Low-Fidelity Prototype and Usability Testing Assignment

Grading Guidelines (55 points total)

<u>Preamble (1 point)</u>

Like the contextual inquiry assignment, we added 1 point to the grading guidelines to emphasize that it *is* important to read the instructions in the assignment handout, and to seek clarification *ahead of time* in case of doubt. This point may be deducted if any of the following was not provided:

- URL (We were not expecting an online submission using the Swiki for this assignment, although that would be appreciated.)
- name and role(s) of each group member
- 6-page limit
- good writing with professional presentation
- a complete appendix that includes consent form, demo script, task descriptions and raw process data

It was also greatly appreciated that the group number was specified on every assignment report, even though it wasn't stated in the assignment handout.

<u>Introduction and Mission Statement (6 points, ¼ page)</u>

We awarded the 6 points as follow:

2 points –   introduction to the system
2 points –   purpose and rationale behind lo-fi experiment
2 points –   mission statement, with common purpose and goal of project clearly communicated

This is a lo-fi experiment, so we deducted ½ point if the purpose does not clearly communicate why lo-fi prototyping, as opposed to other human-centered techniques, was carried out. We also deducted ½ point if the project goal is not a specific performance goal that a third-party can use to evaluate the product's success. Please re-read "The Discipline of Teams" reading if you do not understand what is a performance goal.

Prototype Description (12 points, 1 page)

We awarded the 12 points as follow:

6 points –   description of prototype
4 points –   sketches *and* references to the sketches
2 points –   picture or DENIM/SUEDE screenshot

We were lenient in grading this section, and awarded full points for the prototype description if the system's overall structure (e.g. how does one part of the interface relate to another?), mid-level functionality (e.g. what is the function of each major section of the interface?) and low-level interface details (e.g. using post-it notes to build dialog boxes) were clearly communicated.

But we took off points for descriptions that were difficult to understand because they did not make adequate references to the interface screens, or if sketches were too blur to be read. We also took off points for descriptions that did not clearly convey that the prototype is a lo-fi paper prototype.

Method (12 points, ~ 2 pages)

We awarded the 12 points as follow:

1 point –   description of participants
1 point –   brief discussion of how participants were selected

1 point –   description of test environment
1.5 points –   brief discussion of how prototype and other equipment were set up

1.5 points –   all 3 task descriptions

3 points –   description of testing procedure, including the role(s) of each group member, how the Wizard of Oz technique was applied, and possibly additional precautions taken

3 points –   test measures, including the process data (i.e. what is happening in the big picture) and bottom-line data (i.e. time or # of errors) collected, and the rationale for using them

Several groups did not elaborate on how they carried out Wizard of Oz under the testing procedure sub-section. Remember that Wizard of Oz is an essential component of a lo-fi experiment – you cannot perform a lo-fi usability test without this technique.

Results (12 points, ¾ page)

We awarded the 12 points as follow:

5 points –   adequate coverage of findings, especially in terms of test measures
5 points –   adequate coverage of critical incidents
2 points –   severity ratings

We were lenient in grading this section, and accepted most answers as long as they *summarized* the test measures and critical incidents reasonably, with average severity ratings given for most critical incidents.   We also accepted descriptions of severity ratings (i.e. "major usability problem", "catastrophic problem", etc.) in lieu of the numerical ratings.

But we could not award points for assignments that did not report results pertaining to the test measures. It is incomplete to decide on test measures (both process data and bottom-line data) in the previous section, and not to follow-up in this section by reporting on them.

Discussion (12 points, ¾ page)

We awarded the 12 points as follow:

2 points –   rationale for severity ratings assignment
4 points –   major lessons learnt
4 points –   how did results inform redesign of the interface?
2 points –   limitations of the experiment

We were also very lenient in grading this section.

But we were unable to award points for assignments that did not summarize and discuss the significant lessons learnt from the lo-fi experiment. One example of a major lesson would be one that pertains to how the interface impacted the user's performance of his/her tasks, although this was not something we were looking for explicitly. On the other hand, cosmetic interface issues are not necessarily significant, but we nevertheless awarded partial credit for write-ups that hinted at related insights, even though they were not explicitly presented.