User Testing Report Template

Created by Mark W. Newman for Coursera course
Evaluating Designs With Users
Updated: April 6, 2017

Executive Summary

The first section of every report should be an executive summary. It should be between ½ page long. Summarize your methods, your results, and your recommendations concisely. The Executive Summary should convey *all* of the critical information contained in your report, though in an abbreviated form.

Introduction

The introduction to your report is to orient the reader to your understanding of the system's goals and audience and what you were trying to accomplish with your study. It should be ½ - 1 page long and contain the following:

- a brief overview of the system you're evaluating
- the key questions you are seeking to answer with *this* study (e.g., interviews, survey, usability test)
- some motivation for why those questions are important

Methods

In your methods section, present the details of what you did to carry out this study. This section will typically be 1-2 pages long and contain details of how you implemented the study. These will vary according to the type of study, but will include such details as:

- your target population
- your recruiting methods
- a summary of your participants (omitting names, but including salient demographic details)
- your instruments (e.g., scripts, questionnaires, task descriptions), with pointers to the appendix if they are too long to include in the body
- your analysis methods (e.g., video and logging sheet reviews, statistical analyses, analysis of critical incidents)

Findings and Recommendations

This section will usually be the longest section of your report and should contain the most significant findings from your study, along with recommendations for the client about what they should do. Usually you will have to be selective about the findings you include here, as you will have more findings than it's useful to present in the body of the report. This section will include:

Summary Results

Here report baseline statistics such as task completion rate, task timings (if appropriate), error rates, etc. Also report questionnaire results and high level reactions. A general assessment of the outcome of the test is appropriate here as well (e.g., did the test reveal that the product needs a lot of work or just a bit of fine-tuning).

Key Findings

Select a handful of key findings (usually 5-10) to present. These can be further categorized into a few groups if that makes sense for your study. For each finding, be sure to be clear about the following:

- what the problem is and how it manifests itself in the system (e.g., what pages, screens, or widgets are affected). It can be useful to employ screen shots (See Fig. 1).
- the evidence you have that a problem exists
- your assessment of the severity of the problem

Figure 1. Including screen grabs from the system being evaluated is often a good way to clarify what a problem is and how it manifests itself. This screenshot could be used to illustrate a source of user frustration with Doodle.com, namely that users find it cumbersome to manually enter date and time ranges in the event editor's many text boxes. Captions should contain complete sentences and explain the relevance of the image to the report text. Also, all figures and tables should be numbered and they should be referenced from the text.

Recommendations

Provide some suggestions about how the client might address the problem. Recommendations might take the following forms:

- If you have a concise redesign suggestion, offer it. Where possible back it up by referring to best practices (e.g., other "superior" products identified in your comparative analysis) or principles (e.g., Nielsen's heuristics (Nielsen 1994) or the advice of other UX sages).
- If the redesign is more complex, you can suggest a process (e.g., a thorough redesign of the navigation architecture should be undertaken, with special attention paid to…")
- If the severity or prevalence of the finding is unclear, you may recommend further research (e.g., "a usability test could be conducted to determine whether users will find the issue to be a significant barrier to using the site or merely an inconvenience.")

Note that recommendations can be presented as a separate subsection or they can be bundled with the findings they refer to.

Limitations

Including limitations is optional, but is generally a good practice. This will be short, usually about ½ page. Here, you should address any shortcomings of your study. These might include limitations of your sample (bias, size) and how these limitations may have affected the outcomes. You might also include alternate interpretations of findings and justifications for the interpretation you chose. If there are significant questions that remain unaddressed by this study, you might highlight those and suggest how they could be addressed by future studies (even if you're not planning to conduct those).

Conclusion

The conclusion should be brief—about ½ page long. It's very similar to the executive summary in that it will summarize the key takeaways from the study (what you want the reader to remember after they've finished reading). This will include a brief restatement of the key findings, recommendations, caveats, and suggested next steps.

References

Nielsen, J. (1994) Heuristic Evaluation. In J. Nielsen. & R. L. Mack (Eds.) *Usability Inspection Methods*. New York, NY: John Wiley & Sons.

Appendices

The appendices should include instruments, data, and other artifacts from your study that would help an interested reader delve deeper into your study to validate or reinterpret your results. These materials should be understandable by your target audience, perhaps including brief explanations about the source of the data and the steps taken to make the data more comprehensible. The appendix materials should not be essential for understanding the report, but should provide additional support for the intensely curious reader.

For user testing, appendices would ideally include:

- Your test script
- Your consent form (unsigned)
- Any screening, pre-test, post-test questionnaires that were used
- Filled-out logging forms, or a more readable transcription of the data from them
- Raw questionnaire responses
- A complete list of usability issues found, with severity ratings