Usability Testing

Project Name: Element Being Tested

Digital Corps
Month Year

System Overview

This is where you describe the current or proposed website/app/experience. Describe the most essential goals of a project and then describe the sections of the system that accomplish these goals. Once you define the goals, refer to this overview to decide which elements of the system need to be evaluated. Then determine how best to define tasks that would naturally interact with these elements. This should be comprehensive and thorough, yet concise; avoid bullet points if possible.

Location

<URL> or how to access the app/system

Methods

Disclaimer (optional)

If there is anything unique to this project testing, write it out here. This would most likely be added after testing has occurred, because you may not be able to predict how testing will go. Examples include:

- Tasks were designed to be used with a functional prototype but users were instead given a paper prototype due to time constraints.
- Users had account type a, while tasks were designed for users with account type b. This could apply to the fact Digital Corps students are testing a site meant for faculty (e.g. Writing Challenge).
- Essentially, anything that could have skewed or altered your results.
- **It is important you are honest with your methods, including any limitations, because if, for example, we suggest things that a client does not agree with, they may look for ways to invalidate our suggestions.**

Goals of Testing

Using bullet points, list the primary purposes of this testing. The way we've written goals previously can be used here. As stated in the system overview section, these test goals should

directly relate to the system's overall goals. Try to use action words when describing the goals - we want to *evaluate*, *understand*, *identify*, *assess*, *review*, etc. By stating clearly defined action-oriented goals, it defines our purpose and we can more easily determine if we met those goals.
Example Goals:

- Evaluate layout of ___
- Assess navigation redesign for expert users
- Identify weak points in information architecture

Participants

Describe your user pool. Ideally, these would line up with previously created personas, but in most of our projects, they just need to meet a few base requirements:

- The participant has not worked on this project
- The participant has not participated in other testing for this project
- The participant has minimal background knowledge of the project/system

Pre-Test Questionnaire

This section is closely linked with the participants section. Essentially, take all of the requirements for your participant pool and turn them into questions. Example:

- Have you, in any capacity, worked on this or any other related project?
- Have you participated in any other studies related to this project?
- Before you can participate in this study, please describe your knowledge of this project.

These questions can be sent via Slack or asked in person. They should be asked before someone is scheduled for in-person or remote testing.

Tasks

How your tasks are written depends greatly on the system you are evaluating. If the system is designed for a specific type of user or scenario, you should try to incorporate that into your tasks. If the system is for "all Ball State students", then your tasks may not include a scenario (since most of our participants are Ball State students). Here is an example set of tasks:

1. Your boss just dropped off this last-minute addition for the lesson. Add this question to the lesson you created earlier.

   Expected screens: xxx

   Steps

   - Do this
   - Do that

2. The training department created several videos for salespeople, and requested that they are used within the manual. Add the video from your company YouTube channel to the current section.

   Expected screens: xxx

   Steps

   - Do this
   - Do that

3. Your boss asked you to take a series of outdated Word documents and create a lesson in Lessonly. Use the content below to get started.

   Expected screens: xxx

   Steps

   - Do this
   - Do that

For some usability testing, we may implement the **"lostness index".** With this reporting method, we define how many screens a task should take and then compare it with the actual number of screens that a participant used. **While there may be multiple routes to the desired destination, it is important to identify the minimum number of screens/pages required to complete a task.**

Definition of a screen

The definition of a "screen" includes:

- When the browser loaded a new page and/or the URL in the address bar changed

- When the page layout and available form elements changed

A new screen does not include:

- Pagination of a list of items
- Conditional input box (EX: “yes/no” radio button, if yes, enter a comment)

Post-Test Questionnaire

The post-test questionnaire will be the last data you get from your participant, so it should be thorough and effective. We implemented a Google Forms template of the Systems Usability Scale (SUS), which can be found here. When you create a usability plan, you should **make a copy of that template** and add any questions you would like to ask. For the purposes of this document, you can just briefly describe what the SUS is and then link to your form. If you would like to add any other questions to the survey, consider both qualitative (free form responses) or quantitative (Likert scale). For Likert scales, use the following format: 1 = strongly disagree, 5 = strongly agree. Write out all question for this part of the plan.

1. This is the first Post-Test question
2. This is the second.
3. This is the third.
4. This is the fourth. Testing to see if Jolee is still reading and will remove this one.

Script

Write a word-by-word script of what you will say to each participant. By using the same wording with each participant, we prevent some bias or skewed results. Here is an example script:

> Thank you for coming today. To begin, I have a script that I will read to ensure all participants receive the same information. Are there any questions before I continue?
>
> The purpose of today’s activity is to test the learning management system Lessonly. We want to observe whether Lessonly is easy to learn and understand for people who may not be familiar with the site or familiar with learning management systems. Based on what we observe and learn in today's activity we will advise Lessonly on how to improve the system.

I will be asking you to try and complete various tasks. We want to see if Lessonly makes it easy or difficult for the learner to find what they need to complete each given task. It is important for you to understand that we are in no way testing your ability to complete these tasks. We are only wanting to understand the usability of Lessonly.

This activity will take about 30 minutes to complete. I will be taking notes the whole time and will ask for your comments about Lessonly. The information you will be giving us is going to be kept confidential and your name will not be stored with the information.

Please read this consent form and sign it if you agree. Are there any questions about what I have gone over so far? Are there any questions about the consent form?

Hand a consent form to the participant.

Before we get started, I would like to ask you a few questions about yourself. Once again, let me point out that your name will not be stored with the information.

For this test, we will be using a series of handouts that represent the type of content a creator would use within Lessonly. The lesson should match, as close as possible, to the handouts, including headers and content. Each task will have its own handout, along with the task printed at the top. As you complete the tasks I would like for you to use the Think Aloud method. This method requires you to narrate what you are doing and your thought process as you navigate the site to complete each task. By using the Think Aloud method, I will be able to better understand the thought process behind how a user might navigate Lessonly.

If you find that a task is too difficult to complete, that is alright, just let me know and we will move onto the next task. At the end of each task, I will ask you to rate the difficulty of the task. Do you have any questions?

Are you ready to begin?

Administer each task in turn by reading the task, then offering the handout to the participant.

You have now completed the tasks. Your participation is appreciated and it has been very helpful with learning what tasks are more difficult to complete on Lessonly.

I would now like you to answer some general questions about Lessonly and the tasks you completed on the paper provided.

Hand post-session questionnaire to the participant.

Thank you for completing the post-session questionnaire, now I have one last form for you to complete about Lessonly's usability.

Hand the System Usability Scale form to the participant.

That completes today's session. Thank you for your participation.

Escort the participant from the test area.

Task Success Ratings

Task success ratings are used to evaluate how successful a user was in completing a task. The ratings are as follows:

- **S** = success. The user was able to complete the task with no problems
- **P** = partial success. The user was able to complete the task, but they needed some clarification
- **F** = the user abandoned the task or was unable to complete the task.

Procedure

Chronologically describe the steps of testing, but not your prep work. For instance, don't describe when you will print materials or schedule testing. Assume that is all done. You can, however, briefly describe the sample size.

Example

We scheduled 5 participants for testing. Participants went through in-person testing one at a time. They were not compensated for their time.

1. The participant is briefed on the structure of testing.
2. The proctor confirms that the participant meets all requirements

3. If scenario-based testing, the proctor reads the participant the scenario.
4. The participant's screen is recorded.
5. The note-taker begins taking notes.
6. The proctor gives tasks on a note card, one by one.
7. The proctor asks follow-up questions, as needed, and the note-taker records the participant's answers.
8. The participant uses the system for each task.
9. The proctor asks post-test questions

NOTE: Keep scrolling, this is not the end of the document!

Quality Assurance Report

Brief introduction to the QA testing. A spreadsheet of the results of this testing can be found here.

As with the Usability Report, provide a high-level overview of the results and identify any specific problems that were discovered. This would eliminate the need to leave the document and parse through a list of (mostly) successful cases (fingers crossed).

Usability Report: Round 1

Overall Task Success Rates

This can be reported with a variety of different visualizations, but you essentially need to report the percentage of participants who were able to successfully complete a task. To calculate this percentage, count every 'success' with a value of 1, every 'partial success' with a value of 0.5, and every 'failure' with a value of 0. Then average out across all participants, and you have a percentage to report. Below are a couple of examples as to how you can report this percentage.

It is important to include additional text with the charts and table to explain what the data means. If one task had a poor success rate, explain why. If a task was not applicable, explain why. If there was anything of special interest, explain that. Typically, the chart and table do not provide enough context without additional details.

Task Number	Success Rate
1	100%
2	100%
3	87.5%
4	100%
5	62.5%
6	100%
7	100%
8	100%
9	100%
10	100%
11	n/a
12	75%
13	100%

Table X. Brief description of the data presented in the table – this should be thoughtful and relevant, do not just repeat the name of the table. Tables and Figures (charts, graphs, screenshots, etc) each have their own sequential numbers - so start with 1 for Table and 1 for Figure.

Captions should be concise but comprehensive. They should describe the data shown, draw attention to important features contained within the figure, and may sometimes also include interpretations of the data.

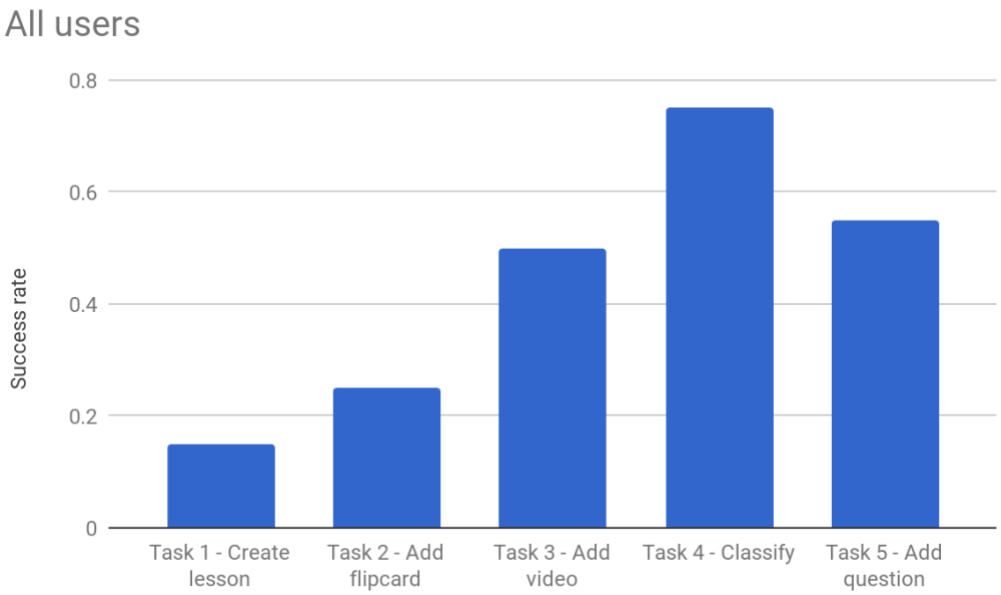

Figure X (replace X with number, in sequence). Brief description of the data presented in the chart.

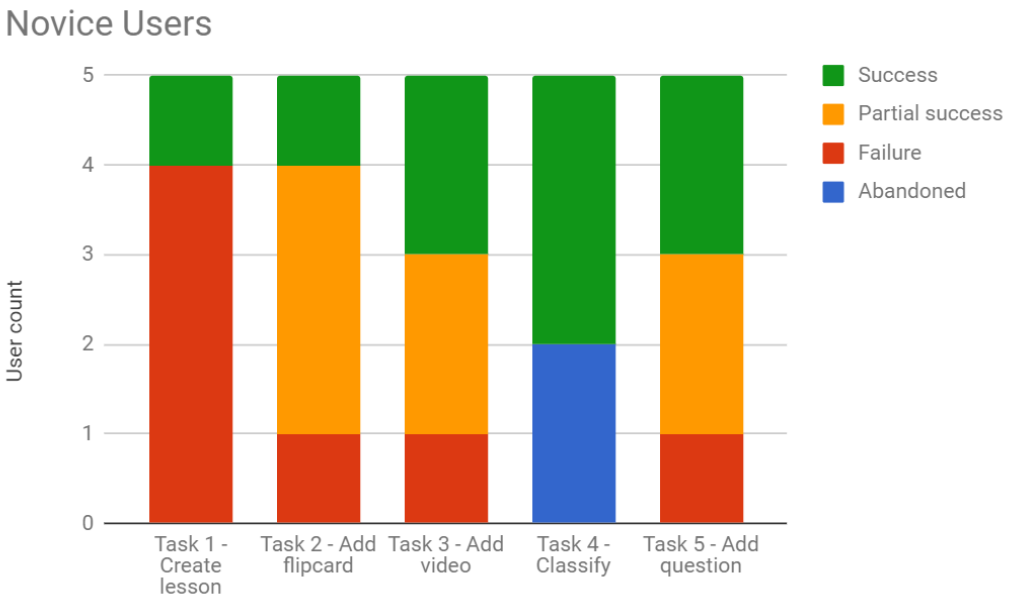

Figure X+1. Brief description of the data presented in the chart.

Lostness Index

To calculate the lostness index, it takes a bit more math than the task success rates. See this article for a step-by-step walkthrough and explanation. Use this formula:

$$L = \sqrt{(N/S-1)^2 + (R/N-1)^2}$$

- L = Lostness
- N = The number of unique screens visited during the task.
- S = The total number of screens visited during the task.
- R = The minimum number of screens that must be visited to complete a task.

The score ranges from 0-1. The higher the score, the more lost a user is. **Therefore, a higher score is a bad thing.** Report these findings in a bar graph (example shown below).

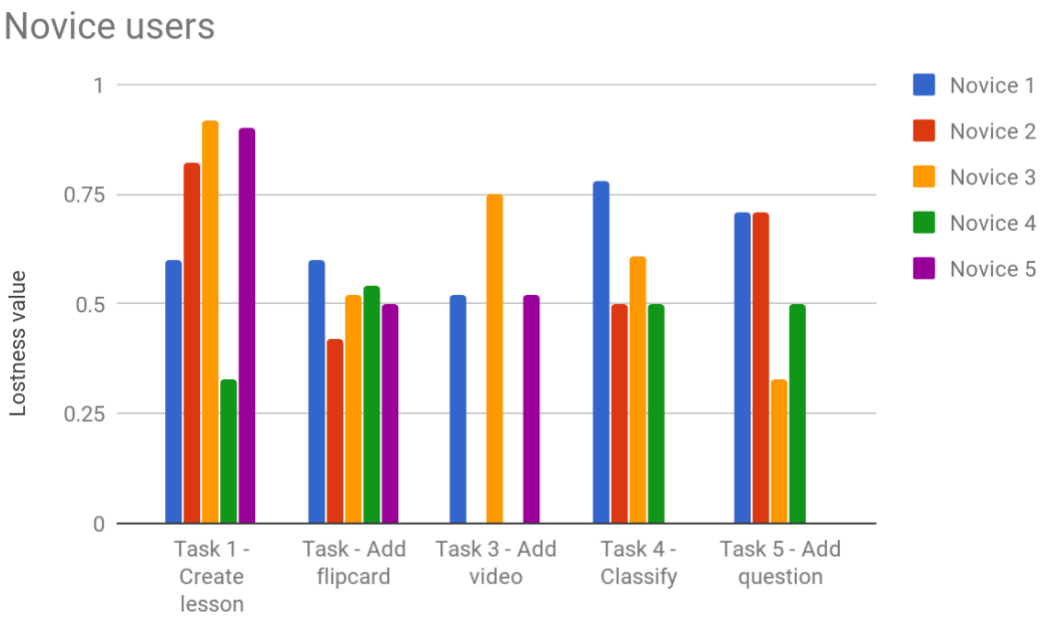

Figure X+2. Description.

System Usability Scale

Use a bar chart to report the results of each participant's SUS score. Report on a y-axis of 100, as that is the maximum score.

If you added any questions to the post-test questionnaires, report them in a table in a separate section (shown below).

Post-Test Questionnaire

In a table, report the Likert scores of any additional questions you had in your post-test questionnaire.

Question	All users
Completing the tasks assigned to me was easy	3.4
The tasks assigned to me were easy to understand	3.1
I felt overwhelmed when trying to complete the tasks	2.4
I understood how to use the handouts to complete the tasks	4.0
Using the handouts to complete tasks was clear and simple	2.8
Lessonly was easy to navigate	2.9
The directions given by the moderator in the beginning were clear	4.5
Creating a new lesson was easy	3.8
Adding multiple elements to the lesson was easy	4.4
I didn't feel the need to ask questions in order to complete tasks	3.9
The moderator allowed me to complete the tasks on my own without much guidance	4.8
I felt comfortable expressing my thoughts out loud when completing tasks	4.5
Most of the time I felt nervous that I would make a mistake	3.9
I understand the purpose of Lessonly and would use it in the future	3.9

Table X+1. Description.

Task-By-Task Ratings

These should serve as a general overview of how a task went. A table works best. Example below.

Task 1: Example Task Title

Description	Your boss asked you to take a series of outdated Word documents and create a lesson in Lessonly. Use the content below to get started.
Success rate	All users: 0.15 *Average of all users based on the following rating: Successful completion: 1.0, Partial: 0.5, Failure: 0.0*
Time on task	All users: 5:35 *Average of all users, in minutes and seconds*
Efficiency rate	4% *Based on all users, ratio of the time taken for successfully completed tasks in relation to the total time taken by all users*
Lostness index	All users: 0.69 Expert: 0.67 Novice: 0.71 *All users had significant issues in finding what they needed to complete the task (less than 0.4 indicated a low lostness value)*
Post task difficulty rating	All users: 3.9 Expert: 3.4 Novice: 4.4 *Average response from the users. The rating scale was 1 through 7, with 1 representing a "very easy" task and 7 "very difficult."*

Problem Statements

This is the meat and potatoes of a usability report. This is where you get to give your expert input as to how to fix the issues discovered during testing. Tables are a good way to share this information. **Format this section so there is one problem statement per page. The intention is a developer or designer could print off this one single page (or two if you're long-winded) and have all the information they need to implement/address your feedback.** If there are numerous problems in a system, consider organizing them by subcategory using nested sub-headers.

Example on the following page (each problem should appear on a separate page):

Severity Ratings

- 1 superficial errors - does not need to be fixed unless time is available
- 2 minor - low priority
- 3 major - important and should have high priority
- 4 catastrophic - should be fixed immediately before release

Problem 1: Multiple Tabs

Description	Describe the problem as encountered by a participant. This is typically 1-3 sentences.
Assessment	Assess the problem and describe why this is an issue. This may be a longer explanation - 1-3 paragraphs.
Severity	2 - minor
Design defect	Explain why this is a design defect - this is the technical portion of what went wrong with the element/feature/section. This can be 1-3 sentences or longer, as needed.
<include a screenshot of the affect area/problem>	
Effect on user	Explain how participants feel and how the problem changes their experience in the app/website.
Recommendations for improvement	Explain what should be done to fix the problem.
Participants who encountered this problem	1 novice (#1) 3 experts (#2, 3, and 4)
Representative comments from participants	Provide some quotes from the testing that highlight the issues faced by participants
Link to audio/video example	Include a timestamp of the video/audio recording along with the link

Testing Summary

Depending on the scale of your usability testing or the system you're evaluating, you may find it beneficial to write up a brief, couple paragraph summary of all aspects of testing. Consult with your teammates and team lead if you are unsure.

Usability Report: Round 2

TBD - need to determine how to discuss changes made to the app/website and the supplemental testing conducted, along with results.