W14 Report (Project Part II): Usability Study Report

Jesse Woo 39616545
Sadia Khan Durani 43762962
Kelsey Chua 85997864
Arshvir Bhandal 60148079
Calvin Zhang 69875094

Team: Team 5
April 10 (Thursday), 2025

Evaluation and Analysis Results (3~4 pages)

a) **Topic**:

An interface to improve the experience of searching and/or browsing of restaurants.

b) **Evaluation questions**:

1. Learnability: Is the interface easy for users to understand and navigate?
2. Learnability: How long does it take a user to know what features are available to use and how?
3. Frustration: Does the user think there are too many steps (clicks, scrolling, etc.) between them and the goal they wish to achieve?

c) **Summary of evaluation**:

Our final usability study methods and protocol involved assessing learnability and frustration levels related to our medium-fidelity prototype. We conducted each study remotely over Zoom (using UBC accounts), using both think aloud observational and structured interview methods. Each session lasted approximately 15 minutes and was recorded for further analysis to ensure that the coding sheet can be filled accurately. The recordings were stored on Canadian servers (UBC OneDrive). Participants interacted with the med-fi prototype, specifically completing 3 tasks: searching for and adding an Italian restaurant to an existing list, updating tags to a personal list, and browsing restaurants to select a potential restaurant.

We recruited five participants for the study, all 19 years of age or older. Each participant confirmed eating out at least 2 times a week, making them representative of our target users, which are students who explore restaurants. During the observational portion of the study, we recorded metrics such as the number of navigation errors made, time spent locating features, and time spent browsing lists. The interview included asking rating questions on the ease of use and satisfaction, as well as the ideal number of steps for certain tasks. Overall, these methods helped us gather insights into how well the interface in current design phase supports the users and where improvements will be needed.

d) **Evaluation rationale (Step 2):**

Our evaluation was designed to assess 2 main aspects of our interface, as stated above, which are learnability and frustration. Our evaluation questions were selected because they directly relate to how well users accomplish certain tasks. Learnability was important to measure to ensure users could understand and navigate the interface with minimal guidance, while frustration helped us evaluate the interface's potential limitations.

We decided to use a think aloud observation method because it allowed us to gain insights into how users approached the tasks. This method produced data in the form of screen and audio recordings which were converted into transcripts, and responses on a coding sheet. Observing the users in real time, and reviewing the recordings, gave us a clear view of how easily users navigated the interface and possible parts of the interface design that caused problems.

To follow up with the observation, a structured interview was conducted. This gave us qualitative data on the user's satisfaction, and their sense of how many steps certain tasks should ideally take. These responses help us understand their navigation process or possible reasons for errors made with the interface.

e) **Prototyping rationale (Step 3):**

We made a prototype that is mostly horizontal with some verticality, so that users would be able to visit many pages to see the different lists and restaurants available to make a choice about what they are interested in. The high degree of horizontality was in the many potential options illustrated and pages given. The verticality was largely limited to the potential depth of browsing a user could do through the lists and restaurants, as the buttons and other functionalities were not fully implemented. It was made using Google Slides because it has just enough functionality for browsing

and required the minimum effort to operate for a med-fi prototype. The data is all fake, but the pictures of the food give the prototype a realistic view of what kind of foods each restaurant could have. The prototype still requires supervision because it is not robust. For example, if the user were to click on somewhere that isn't a button, it will move to the next slide, or if they misclick anywhere else, they would most likely not be able to return to the previous page.

f) **Summary of data, findings and analysis (Step 7)**:

Throughout our evaluation, we found various problems with our interface. First, we had problems directly related to the prototype implementation. During our observation, two participants failed to locate the relevant buttons for their task as they were blended in with the rest of the interface. A response from the interview noted in particular, that more colors would help distinguish different elements. These visual fidelity-based problems may be specific to the medium-fi prototype, which may not have been clear enough with its options. Then, in terms of implementation fidelity, one participant complained that the selection box in front of the tag name was not clickable. Another participant mentioned they could only return to the home page, but not to the last search page they were on. These once again appear to be local to the particular medium-fi prototype, and not the overall design.

Our decision to use Google Slides as the prototype medium contributed to many user mistakes during the evaluation. We categorized these problems as related to the prototype medium itself. From our observations, 3 participants clicked on unclickable places/buttons, or scrolled mouse wheel which caused the slide to move forward. On average, each participant made 2.6 mistakes during the observation (Figure 1). It could be different types of errors behind the scenes. One type of error is caused solely by our choice of prototype medium. The second is the fidelity issue, where users were trying to interact with the features that had not been implemented yet. The third type of error is a design issue, making users misunderstand how our interface works. Currently with the data collected in our evaluation, we cannot distinguish which reason caused each erroneous click or mistake. During our evaluation, we should have asked follow-up questions to participants to explain the errors they made, to clarify the reason behind their mistakes.

In our observations, some participants met understanding issues when using the interface. Among them, the most common issue encountered by 3 participants was that they did not know they needed to click the search button to apply the filter after checking the filter box. One other area of confusion is when told to edit the tags, one participant clicked on existing tags, instead of our intended "clicking on edit list first, then clicking on edit tags". It also relates to our interview question, where 3 participants answered they want to limit the steps to add a tag to under 3 steps. One of the participants complained there are too many steps for adding a tag. Another participant explained he preferred to have two steps, one being able to search the tag, then add to list. Or only with one step, clicking from the recommended tags to add to the list. From all the feedback collected, we clearly need to simplify the process of adding tags. Lastly, we have recorded the time spent on each of the tasks as in Figure 2. Although there are some participants taking more time than the others, the difference is not great enough to draw a conclusion. For the "browse list" task that shows the greatest variation of time, it is mainly because some participants liked to spend more time exploring different restaurants, whereas some participants made quick decisions. In summary, the use time data was not particularly useful for our evaluation.

Finally, in terms of positive feedback, participants gave an average of 4.2 out of 5 on ease of use for our interface, which can be seen in Figure 4a. Some participants remarked in particular that adding a restaurant to a list was easy, because the first things that pop up after clicking are the lists to be added to. Likewise, the restaurant pages also contained enough information for the users to make a decision, which may include price, hours, or reviews. This then connects to the level of frustration users had, which was fairly low despite the earlier noted concerns. As per Figure 4b, when asked about their experience, with 1 being satisfied and 5 being frustrated, most participants reported a 4. As such, it can be seen that the overall experience was still well received by participants. It must be

noted however, that the prototype was made explicitly to support the given tasks, and that an actual implementation with features that do not relate to the tasks, such as more tags, may change how these tasks would be completed due to potential cluttering.

g) **Conclusions (Step 8):**

To begin with learnability, our interface would appear to be fairly easy to understand and navigate. The number of errors made seems reasonably few, and the high ratings for ease of use also corroborate the fact. Likewise, the time spent on each task seems fair, with the ease-of-use ratings once again showing that the users agree that it was not overly difficult to learn. In terms of frustration, while it did seem that users thought that there were too many steps for adding a tag to a list, the overall experience was regarded as non-frustrating. These high ratings are likely due to users generally having a smooth experience, save for the problems discussed later. The smoothness was found to be provided by the low cluttering and logical flow of the design which users could follow, as per participant comments. In addition to these aspects, users also generally thought the interface had a good amount of information regarding restaurants. That said, there were also numerous problems that users had with the interface, as seen earlier in the analysis.

In terms of design deficiencies, there were two parts that were found to have unclear and unwanted steps. The of these is the requirement of pressing "edit list" before adding tags. We opted for this to prevent immediate clutter when accessing one's list. This was to support users who would mainly use their lists for their own reference and thus have a lower need to update the tags. As we learned with our evaluations however, it made the steps to add a tag unclear. This design choice increased the level of potential frustration through adding more unneeded steps as it seemed most users would rather just have the option more immediately available. It also added a layer of confusion for learning how to edit a list.

The second area that had this problem was the search functionality, where users must press the search button in addition to pressing the filters to get their results. We made this decision to better serve more in-depth searching, which may involve multiple filters and/or keywords. However, as noted in the analysis, users who are looking for quick results for singular criteria are then limited from finding them. The interface reality did not align with some participant's expectations and left them unable to figure out how to get to their desired results. As such, the learnability of how to find and use new features for the searching was lower than expected. Aside from these problems, there were also some design limitations to be considered.

The main limitation to be noted is users generally having to use lists to browse restaurants. We chose to implement it as so since we consider lists a core feature of the design. However, we recognize that this may cause problems for users who do not care much for lists, as their immediate view upon opening the interface is limited to showing lists only. This concern was addressed in the case of one participant, where they chose to forgo browsing many lists for the third task, instead just choosing the first list and a restaurant from there. As such, it was still quite easy for a non-user of lists to still easily benefit from this design and navigate as per their needs. To further help these users, however, the use of search tools to have only restaurant results could have been made a more apparent option.

Recommendations and Critique (1 page)

h) **Design recommendations (Step 9):**

Through our user studies, we feel our design is of high quality overall, though we have received feedback for improvements. Our studies found that users ran into at most four errors (Figure 1), including errors caused by the prototype implementation/medium, and weren't frustrated when performing tasks (Figure 4b). This tells us that the interface could still be improved for guiding users to perform tasks correctly but for the most part users are happy with the design as is. Users found that adding restaurants to lists worked well. One participant was happy that "the first thing that pops up [when clicking "Add to List"] is which list to add it to" instead of a default list being selected like in other apps. This tells us that the design for this feature can be kept moving forward as it satisfies the task of adding a restaurant to a

list in a non-frustrating way. During our observations, we found that users struggled to see buttons. Users thought that buttons/features looked alike or blended in, mentioning that "some colours" to "differentiate between different features" would be nice. This is something we should look to improve on for future designs as it would help reduce errors made and frustration when doing tasks, as it would be quicker to spot what is needed for them. This system may not work well in a more natural setting as our results are based on a prototype that was not done with the final medium. As mentioned earlier, some errors were caused by the prototype implementation/medium, and others were caused by clicking in the wrong place. If implemented with HTML/CSS our results could have been more representative of a natural setting, as Slides related errors could be avoided. As mentioned before, participants thought buttons/features were hard to find, which would be an issue in a natural setting. By adding colours and differentiating buttons/features users will be less likely to make mistakes on their own and will be able to guide themselves through tasks without frustration. Overall, the process of doing the tasks was fine but the design could have benefited from being implemented with a different medium and aesthetic.

i) **Critique of process (Step 10):**

We learned a lot from this course as a whole and from this project. The main thing is to never assume what the user is like but rather try to understand them and their needs. From this project, we learned just how important thorough piloting is and how it can be used to avoid issues during the real study. Our pilot study revealed some important insights such as missing information and clarification issues, which we were able to clear up before the real study. However, the one pilot study did not reveal how much of an issue the Slides implementation would be, as it ended up being quite confusing to our study participants resulting in issues unrelated to the interface design. This may have skewed our results by causing our participants to interact with the prototype differently. In hindsight, these problems could have been avoided by implementing the prototype with a different medium. If we used HTML/CSS, users wouldn't have run into issues with Slides, scrolling when they shouldn't for example, and users would be able to interact with the prototype in a way that is familiar to them. The methods we used for evaluation did seem to get us what we wanted. The think-aloud observations for the most part were quite insightful when the participants were talkative. When they were not talkative, we still benefited from seeing how users executed tasks and could somewhat interpolate what they were struggling with. The interview also was overall quite insightful, though in hindsight it could have been beneficial for it to be less structured. With probing, we could have made up for losses in thought during observations with less talkative participants. In the end there is a lot we can improve on with evaluations if we were to ever do another one.

Appendix A

A.1) **Medium Fidelity Prototyping Video Reference**

https://youtu.be/lLonXDlSsX4?si=hi-_w_i7Xy7yQubP

A.2) **Figures and Tables**

Figure 1: # Errors Made by Each Participant

Figure 2: Time Spent on Different Tasks

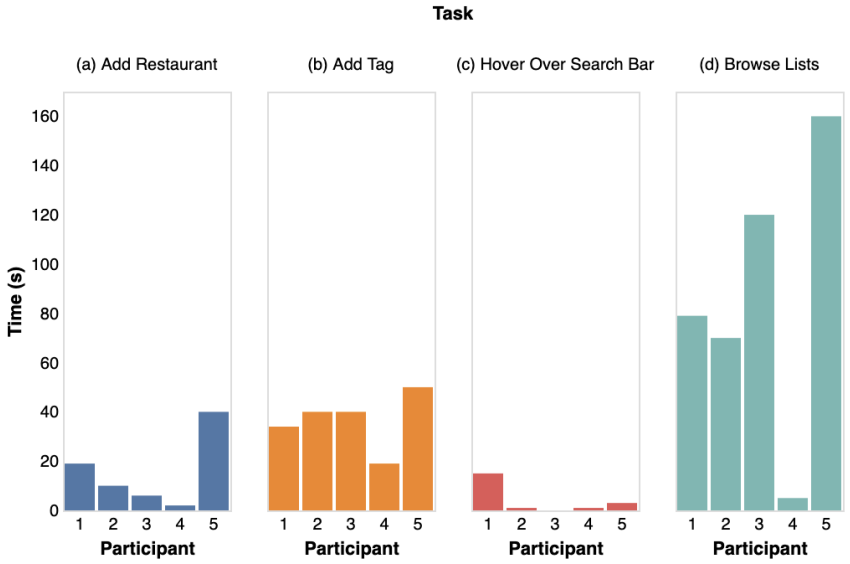


Figure 3: Ideal Steps per Task by Participant

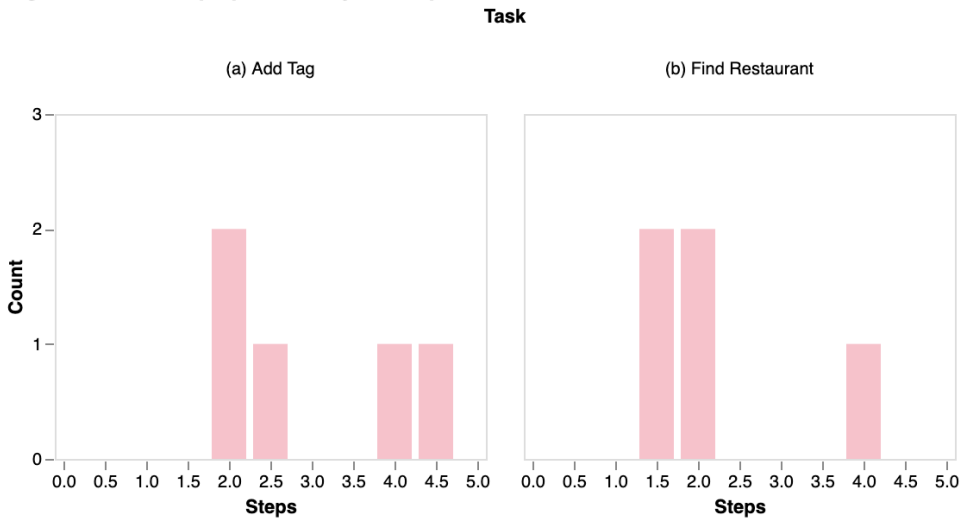
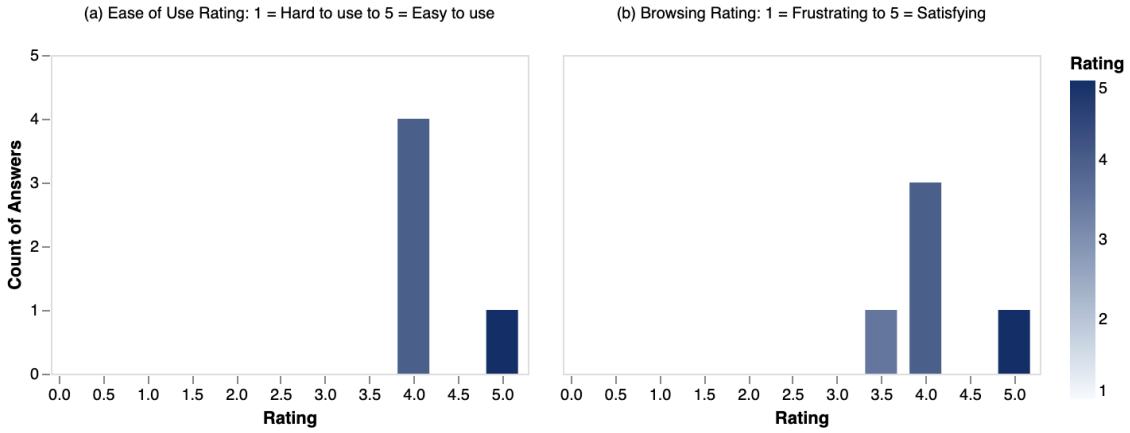


Figure 4: Ratings



A.3) **Evaluation Instruments**

Coding Sheet + Interview Questions

Participant #:	Date:
Evaluator Name:	Tasks order:

Question:	Value:	Other comments:
# of errors made while navigating the interface (clicks that lead to an unintended location)		
# of seconds spent to find how to add to a list (from entering restaurant page to alert)		
# of seconds spent to add a tag to a list		
# of seconds spent to use/hover over search bar (#misclicks)		
# of seconds spent browsing lists before selecting		

Interview question:	Answer:
On a scale of 1-5, with 1 being hard to use and 5 easy to use, how would you rate how easy the interface was to use?	
For the task of adding a tag to a list, how many steps do you think would be ideal? A step may be counted as clicking or any other discrete action (Starting from the list screen, currently 4 clicks).	
For the task of finding a restaurant, how many steps do you think would be ideal? A step may be counted as clicking or any other discrete action (Starting from the home page, currently 2 clicks, but requires browsing).	

On a scale of 1-5, with 1 being frustrated and 5 being satisfied, how would you rate the process of browsing through the lists to find what you wanted? How many lists did you feel you had to look through?	

Extra notes:

A.4) **Supplementary analysis** *(Optional)*

Participant	1	2	3	4	5
# of errors made while navigating the interface (clicks that lead to an unintended location)	4	1	3	3	2
# of seconds spent to find how to add to a list (from entering restaurant page to alert)	19	10	6	2	40
# of seconds spent to add a tag to a list	34	40	40	19	50
# of seconds spent to use/hover over search bar (#misclicks)	15	1	0	1	3
# of seconds spent browsing lists before selecting	79	70	120	5	160

Question	1	2	3	4	5
On a scale of 1-5, with 1 being hard to use and 5 easy to use, how would you rate how easy the interface was to use?	4	4	5	4	4
For the task of adding a tag to a list, how many steps do you think would be ideal? A step may be counted as clicking or any other discrete action (Starting from the list screen, currently 4 clicks).	4	2	3-6	2-3	2
For the task of finding a restaurant, how many steps do you think would be ideal? A step may be counted as clicking or any other discrete action (Starting from the home page, currently 2 clicks, but requires browsing).	2	4	1-2	1-2	2
On a scale of 1-5, with 1 being frustrated and 5 being satisfied, how would you rate the process of browsing through the lists to find what you wanted? How many lists did you feel you had to look through?	4,2	3.5, 2-3	5	4, 1	4, 3-4

Prototype Medium Issues:

- Was trying to scroll down on the restaurant page, and so changed the slide
- Clicking on the wrong location caused it to go to the next slide. In some cases, this was the correct slide to go to anyways.
    - Double clicks also caused this
- Confused by google slides setup expected it to be just an order slide deck without working buttons (scrolled by mistake, incorrect clicks led to different slides).
- Clicked search bar so next slide was displayed, and then clicked the filter

Prototype Implementation Issue:

- When choosing a tag to add, only the text is clickable but the box is not.
- Did not notice “save edits” (Button blended in)
- Buttons look too similar to each other, hard to tell what's where
- Was a little bit confused after clicked into self-profile page. Wasn’t sure which list was owned by him (may because we already had some restaurants in his favorite list that made him thought it is not his favorite list)
- Thought more colours would help differentiating features
- There are too many similar lists on the home page

Understanding Issues:

- Did not press search button (did not know to)
- Clicked “save list” on a list page when trying to find a restaurant to save
- Clicked on tag and description to edit tags and not on “Edit List” button when editing tags
- Didn’t know where own lists would exist, “Favourite spots” was on profile
- Confused with task #3, wasn’t sure what to do when picking a restaurant
- Confused by black lines on home page but realized from arrows they’re scrolling bars
- The user did not correctly use search functionality and was confused if a search had even happened when the filter was selected. (x2)

Moments of Frustration:

- Thought the screen was cluttered on the restaurant pages
    - “Like it's it wasn't frustrating at all to look through stuff, but again, I think the screen is like maybe a bit too much going on and that can be like a little overwhelming. But other than that it's very good.”
- Was a little bit confused after clicked into self-profile page. Wasn’t sure which list was owned by him (may because we already had some restaurants in his favorite list that made him thought it is not his favorite list)
- Thought there were too many steps when adding a tag for first time user

Positive Aspects:

- Liked that adding a restaurant to a list was easy, liked that list options were displayed instead of some default.
    - "I like one thing that I specifically like is that when I try to add something to a list, the first thing that pops up is which list to add it to. I've had other apps do this where like I'll try to save a like a post or a video or whatever, and it'll save to like one of my folders automatically and not let me choose which one I want to save it to so."

  Liked being able to see number of reviews for a restaurant
- Liked that the interface was compartmentalized
    - Likes categorization of lists on the home page
        - "The fact that they like the home page has, like I said, like these compartments of different categories or like different popular lists or whatever it is, is pretty good."
    - Had assumption displayed lists would be based on a user tailored algorithm
- Felt like restaurant pages had the information that they needed (hours, price)
- Liked the home page setup, seemed clear to them
- Said the size of everything was good.
- Thought that "Home Page" wasn't cluttered
- Thought editing lists was easy though executed the task incorrectly (mentioned in prototype issue)
- Liked seeing overall rating for a restaurant when deciding for what restaurant to visit
- Liked seeing number of bookmarks when deciding for what restaurant to visit
- Liked seeing number of likes a review has when deciding for what restaurant to visit
- Seemed to use the list owner review during descension making
    - *on vegetarian delights list* "Well, this one seems like it was bad, so we're not going to select 'salad barn'. 'Pizza Town' is one that I already selected beforehand."