

## M4

Dajiba Patil (dpatil32)

[dajiba@gatech.com](mailto:dajiba@gatech.com)

**Abstract.** When someone goes to store with a shopping list, they end up going through different aisle searching for the product they want and end up wasting time and energy. My objective is to make the shopping comfortable and efficient by creating an interface which can take shopping list and store name as inputs and output the shortest path to shop. In this assignment I am setting up the plan to evaluate my three prototype that I designed using qualitative, empirical and predictive evaluation

### Q 1: Qualitative Evaluation

I am going to use post-event protocols to evaluate my **textual** prototype.

My participants will be few of my **friend and their spouses**. I will be also including my wife in this evaluation process. I will be sending a request to them with following questions. Based on the answers, I will select 4-6 participants for the evaluation procedure

1. Will you be shopping in any of big shopping stores in next 10 days?
2. I will have mixed gender ratio
3. I will have few people who shop while coming from office and in hurry

Evaluation will either take place at **their house** or at my house depending on the scenario. I will be using **note-taking** method to record their answers and comments.

Following Qualitative data will be gathered

1. Participants ease level with the design
2. Challenges faced by them while using the design
3. Suggestions from them to make it the design easier
4. Why did they choose specific method over other methods?

The participants would be first told about the problem space and abstract of the exercise we are doing. I will ask them to go **through** the textual prototype first so that I meet the post event protocol's context. Since we do not have

working model, I can only simulate post event protocol. I will not offer any help during the process. Once they are done, following questions will be asked

1. Were you able to understand what the app was for?
2. What were you thinking when they were going over the prototype?
3. What did you like and what did you did not like about the textual prototype?
4. After going over the prototype, what did you think when you saw buttons “+”, “List”, “Stores”?
5. How to create a new shopping list?
6. Can you tell me how to copy an old shopping list by copying it from old one?
7. Can you tell me how to sort the list by stores?
8. Can you tell me how to add/delete items to the list? What are the different ways you can add items?
9. What button to click to get the shortest path?
10. How to compare the path from other stores?
11. Can you find out how many items have you already shopped? How many items are still to be shopped?
12. How to get your shopping path?
13. How to get the alternative item?
14. How do you know which item you are currently shopping?
15. What is the overall difficulty and challenges on this prototype
16. What improvement or design they would like to have to make it easier for them
17. How you think this app will help you in shopping? If yes, where and when would you use this app?
18. Even if you have ample time at your disposal for shopping, how do you think this app will help you in enjoying your shopping episode
19. How do you think this app helps in every task/subtask you perform for your shopping? Can you explain how? If it does not, can you explain how it does not?

This evaluation will be help me answering the following data inventory items

1. This evaluation will help me to understand who my users might be? If they answer that this app does not help them at all , so based on their profiling, I would know if that profile will be user of my app

2. Based on the Q17, I would know which places they would use the app
3. I am not sure if I can get the context of the task here, I can though call the user while he is going through the textual prototype and see how that context affects him
4. I can definitely understand the goal of the user and if that aligns with what the app is trying to achieve. If the user who has ample time, still says that app will help them to ease out the frustration of finding the item they want to buy, that will give them more time to window shop
5. Some of the answer that I will get , like answer to Q16, will give me idea what else do I need to put in the interface to empower the user to perform their task
6. Task and Subtask: Since I am not observing the user actual doing the shopping, the task and subtasks cannot be observed, however, Q-19 does ask user to explain if the app helps him in his tasks and if it does not explain how it does not. There could be follow up question to understand the task/subtask.

This evaluation will be help me answering the following Requirements

1. If user is understand how to get the shortest shopping path?
2. Q8 will help me in answering requirement of adding items to the list
3. Q13 will help me in answering requirement of alternative items

Most of the questions were drafted keeping the requirement and data inventory in the center, so most the answer that I will get I should be able to map it to the requirements and then infer if the prototype is meeting most of my requirements

## Q 2: Empirical Evaluation

I am using the **Paper** prototype for my empirical evaluation. My interface/app has multiple ways user can add items to the shopping list. However, since my paper prototype was more geared towards smart watch platform, I only had voice enabled, manual entry and syncing the shopping list from the mobile.

Generally voice enabled method is more suitable and useful in platform likes smart watch and with that understanding my assumption is that voice enabled method will help the user in adding the items to the shopping list faster than manually typing the items name in to shopping list. This also makes the

interface usable as giving a manual entry method on a smart watch will be very difficult.

I had designed for a new task and do not have an existing interface to compare. As suggested in the instruction, I am going to slightly modify my prototype so that I will have 2 prototype that I can use to perform my empirical evaluation. Since I am evaluating the task of adding items using voice enabled method, I am going to modify my paper prototype and create one more version where there will not be a voice enabled method and will only have manually entry method to enter the items.

- **Paper Prototype 1:** Original prototype with voice enabled functionality available
- **Paper Prototype 2:** Modified Paper prototype with voice enabled functionality removed

I will be using Paper Prototype-1 and Paper Prototype-2 as my 2 treatments for my empirical evaluations. I am performing this evaluation to **test** if having voice enabled method in my app on a smart watch platform will help the user in creating the shopping list **faster** than typing it manually.

My **null hypothesis** is that having voice enabled method does NOT help the user in adding the items **faster**. As per the video lecture, I am setting my null hypothesis opposite of what my understanding is. My null hypothesis will be assumed true unless I can find conclusive proof for my alternative hypothesis which is mentioned below

My **alternative hypothesis** is that voice enabled method on a smart watch platform does help the user in adding the items **faster** to the shopping list.

I will be recruiting 20 people which will be comprised of

1. People from both the gender
2. Mix of people who shop in hurry and who shop in leisure

I will then split these 20 people in the group of 2 to perform **Within Subject** experiment. Order in which people are assigned the treatment can be one lurking variable which can induce noise in our data. To mitigate this lurking variable, I will be randomly assigning users to treatment groups and at the same time I will randomly assign the order in which they will do the

evaluation. So randomly one person from each group will be asked to start the task and he will be randomly given Treatment 1(Paper Prototype-1) or Treatment 2(Paper Prototype-2) to evaluate. I plan to have same list of 10 items that will be given to all the participant that they have to add to the shopping list. This way all the participants will be performing task on same items.

I will eventually be getting two set of data.

1. Interval/Ratio based on the how long the participants take to add items to the shopping list for both the treatment 1
2. Interval/Ratio based on the how long the participants take to add items to the shopping list for both the treatment 2

For the interval/ratio based data, I will be doing **Student's t-test**. I will be calculating the **average** of all the values for each treatments. I will be also calculating the standard deviation for each treatments. For this analysis, my independent variable will be treatment 1 and treatment 2. My independent variable will be averages and **standard deviation** that I have computed for each treatments. The difference in the values should be big enough to prove the hypothesis.

There will be lurking variables which might affect my test and I have to be cognizant of them. One of the lurking variable can be **typing ability** of participants on the mobile. Participants from millennial generation might be expert in typing on mobile and they could enter the names faster in treatment 2 and it might not be true reflection of real world phenomenon. Another lurking variable could be expertise in mobile usability. Some participants might not be very good with mobile usage and hence expertise in mobile usage can be other lurking variable.

Item names can be short or long and this could be another **lurking** variable which can create some noise in my data that I am measuring. However, I was able to mitigate this variable by giving the same list of items to all the participants

### Q 3: Predictive Evaluation

I have chosen GOMS model to perform the predictive evaluation. The task I am choosing is to create a shopping list and adding items and store name to the shopping list. My goal is to find out how efficiently an **expert** user can perform this task. I have 2 GOMS model created for the same. First GOMS model (Figure 1) explains the different method available to create a new shopping list and criterion for the same. Since this is just the planning, I have not added time to it. Time will be added when we execute it. Second GOMS model (Figure 2) explains different methods to add items and store names to the list

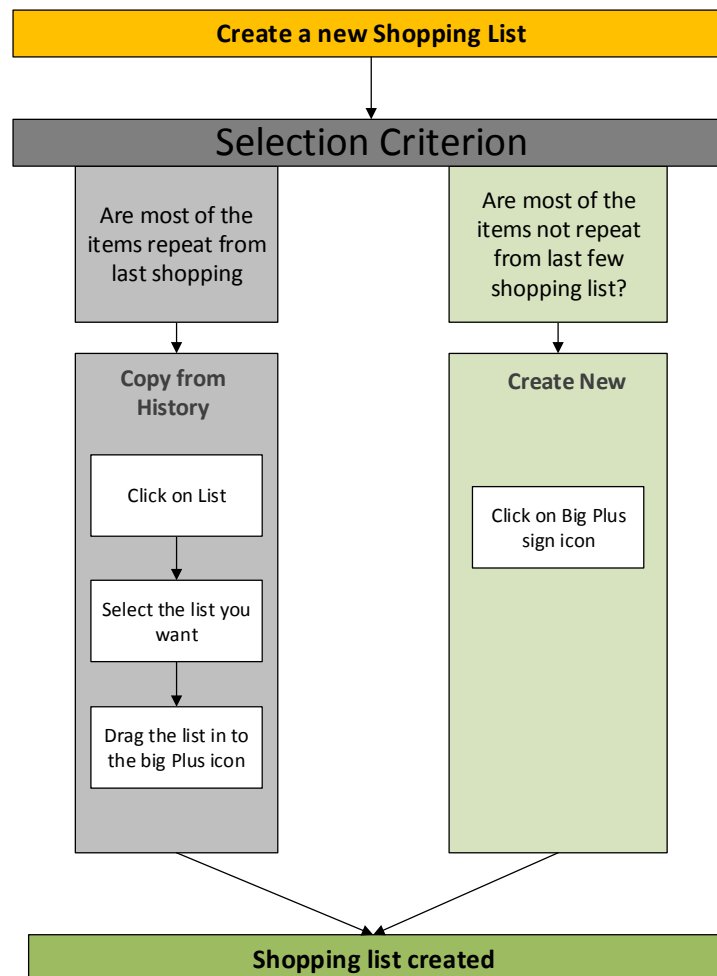


Figure 1. GOMS Model for Task1a

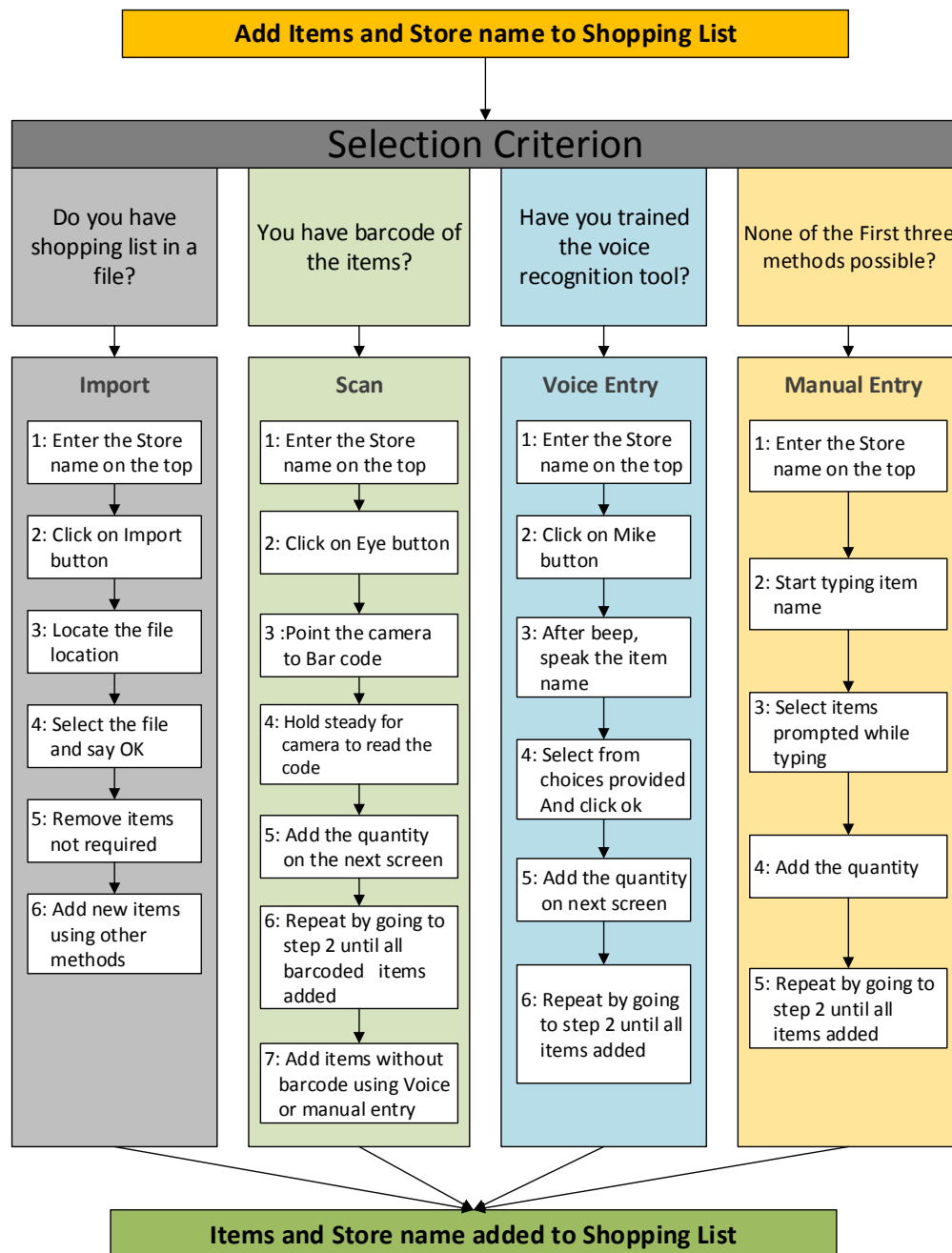


Figure 2. GOMS Model for Task1b

## Q 4: Preparing to execute

I will be selecting Qualitative and Predictive evaluation for my next assignment.

For my empirical evaluation I have to time the participants for performing the shopping task and since I do not have a working prototype yet at this stage, I am not choosing empirical evaluation since it is not feasible to measure time in non-working model. Moreover, empirical evaluation is more suited towards the end of the product life cycle.

Qualitative evaluation was chosen because it is more adept and useful in the initial phases of the user centric design cycle. It will also help me understand the task better by analyzing the answers from the users in this phase. For example if the result of the qualitative evaluation tells me that the interface is not very usable and user spend too much time understanding how to use the interface, then I will have time to modify the design and make it more usable. Since there is not actual working design, these kind of formative and informal type of evaluation is more suitable at this stage

Predictive evaluation was chosen because it is cheaper to perform and is not time consuming. Also, I do not have to recruit any participant for this evaluation. Also, since this is the first phase of evaluation (out of multiple we will eventually have), I felt it is more adept to first go through all the methods, operator and selection criterion before I employ broader user group to do empirical type of evaluation. Also since the working model is not available, this was the best evaluation to do.