CS 6750 Assignment M4

Xu Zhang
xzhang947@gatech.edu

***Abstract*—**Nowadays many jobs are being replaced by machines. The self-checkout machine is one of them. It is commonly seen in many places such as wholesale stores and gas stations. It provides convenience but also raises issues that require human intervention. You may have noticed that there is always at least a cashier standing by the self-checkout machines and being ready to assist the customers that have issues with the machines. I have experience of being stuck in the process of the self-checkout and end up having a semi-automated checkout with the help of the cashier. If you have the same experience, join me to find out the reason for using the HCI principle and methods, and hopefully redesign the current interface to maximize the capability to fulfill the user's needs. In this paper, we will be focusing on the evaluation plans of the prototypes in the previous paper.

1 QUALITATIVE EVALUATION

I would choose a survey as the method for the qualitative evaluation for my paper prototype (see figure 1 in appendix).

Who will be the participants?

The participants are my friends, classmates, and families.

How they will be recruited?

For friends and families, I will contact them by phone or emails and tell them this project and the purpose of the survey. For classmates, I will create a post in the piazza to briefly introduce my prototype and provide a link to the survey.

Where will the evaluation take place?

I will create the survey on peersurvey website and share the link with the participants. The participants can do the evaluation wherever they are.

How will it be recorded?

The survey data will be recorded by the peersurvey website which consolidates the participants' responses into a CSV format. I can then use excel to do some quantitive analysis as well.

Survey Questions

1. Select your age: under 18, 18-29, 30-39, 40-49, 50-64, 64+
2. Which one of the following checkout methods would you prefer when you shop at the supermarket? Cashier, self-checkout machine, other
3. If you choose "other" in the question above, please provide your option here:
4. From 1 to 5, rate your satisfaction with the overall interface design? Highly dissatisfied, dissatisfied, neutral, satisfied, highly satisfied.
5. Which part of the interface confuses you?
6. Which part does not meet your expectation? And why?
7. Are there any other features that you would like to see from the interface?
8. Please share your feedback about this survey. We appreciate your time.

how the evaluation will address the requirements in the data inventory and/or requirements definition phases. How will this evaluation help you gauge whether or not the prototype has actually met the requirements?

The responses to the survey can help me to understand if the user is comfortable with the interface overall. Because the interface is designed similarly to other online shopping interfaces such as Amazon and Ikea. If there is any part that confuses the user, it means the interface cannot meet the requirement of being easy and efficient. The requirements also include that users can use the phone camera to scan barcode and access to details of the product from the interface. If the user has higher expectations or wants to add more features, I would like to know and add to the design requirements.

2 EMPIRICAL EVALUATION

I would choose to use empirical evaluation for my textual prototype where a new type of handheld barcode scanner is assigned to each self-checkout machine.

What are you testing?

My prototype is built based upon the existing interface that is being used in the large grocery stores. The new interface has some add-ons or additional features compared to the existing one. My goal is to investigate the comparison between the new interface and the existing interface and then come to a conclusion about whether the new interface provides a better user experience.

What are you using as a point of comparison?

I would use the time that the user to finish the scan process as a point of comparison. The less time it takes, the better the interface is.

Control and experimental conditions:

I would make sure the time difference from the experiment is only caused by the new features to the handheld barcode scanner and not influenced by other factors. The design and layout for the other parts of the new interface should be the same as the existing interface. I would target the same group of participants and randomly assign them to use the two interfaces. Half of them try the existing interface first and the other half test the new interface first.

Null and alternative hypotheses

Null hypotheses: $\mu_{existing\_interface} = \mu_{new\_interface}$

Alternative hypotheses: $\mu_{existing\_interface} \neq \mu_{new\_interface}$

Experimental methods

The method in this experiment is to compare the difference within-subjects. Each of the participants experiences two treatments, the new interface and the existing interface. I would evaluate this with a paired t-test. It compares the means of two paired sets of data. The important thing here is that individual data points came from the same participant. I will randomly assign participants to different treatments because the order in which each participant tries the two interfaces impacts the results.

How will subjects be assigned to groups?

I would recruit ten participants and divide them into two groups evenly. Each group has five participants of mixed gender and age. Again, I would randomly

assign them to try the two interfaces in different orders. Each participant experiences both the existing interface and the new interface.

What will they complete as part of the condition?

They will be asked to finish the same task which is to scan the same number of merchandise items using the two interfaces.

What data will they generate?

The time each participant does the task on both interfaces will be recorded.

What analysis will you use on this data?

I would calculate the average time it takes for each participant to use each interface (new and existing interfaces). First, I would identify the outlier and then use the rest to calculate the mean. Then, I would confirm if the difference is big enough to accept the null hypothesis. According to the lecture, the null hypothesis can be accepted if there is more than a 5% chance that the difference could be caused by random chance. This is what we called "statistically significant". I would then do a T-test for the ratio variable and Chi-test for the nominal variable.

Lurking variables that confound your data

There are some such lurking variables. For example, a major lurking variable could potentially be which treatment each participant sees first, so we still have to randomly assign participants to treatment groups. Another example is the machine's speed to recognize reading from the handheld barcode scanner. This machine responding time might be long enough to affect the user's efficiency and thus mess up the time data.

3 PREDICTIVE EVALUATION

I would choose predictive evaluation for my verbal prototype which is in the form of a loose conversation script between me and my friend. This prototype is about having a "smart" shopping cart equipped with sensors that can recognize the merchandise items as the user places them into the cart. I would use the GOMS model-based evaluation.

Which type of task analysis?

I would select the merchandise scanning as the task. How many steps and how much time it takes for each user to scan all the merchandise items he/she wants to purchase. Therefore, I would use the GOMS including the goal, selection rules, methods, and operators as well as the estimated time for each operator.

Describe the specific task to be addressed

The task is for each user to scan all the merchandise items he/she wants to purchase.

User's goal

The user's goal is to use this "smart" cart to scan all the merchandises he/she wants to purchase and the ultimate goal is to pay for them.

Operators available to the user

The operators include getting the shopping cart, pushing the cart, locating the merchandise items on the shelf, placing the item into the cart, or taking the item out of the cart.

Will you be evaluating a user accomplishing a single goal they know how to do in advance, or will you be evaluating a user's navigation around the interface to figure out how to accomplish their goal?

I will be evaluating a user accomplishing the tasks which he/she knows how to do in advance. First, I am not a novice user and I am familiar with the steps to complete the task. Secondly, the interface is straight forward and even the novice user knows how to do it. I will record the steps and the estimated time for each step to accomplish the task.

4 PREPARING TO EXECUTE

I would select the qualitative and predictive evaluations for the next assignment M5. The survey can be administered asynchronously for at a pretty low cost. Also, I have created a survey in the need-finding phase, and I received a significant amount of responses including the feedback to improve the survey. As for the predictive evaluation, I would be able to create the GOMS model by myself rather than spending time and effort recruiting the real users. I will ask someone

else to review the GOMS model to avoid me over emphasizing my own option since “I am not my users”.

5 APPENDIX

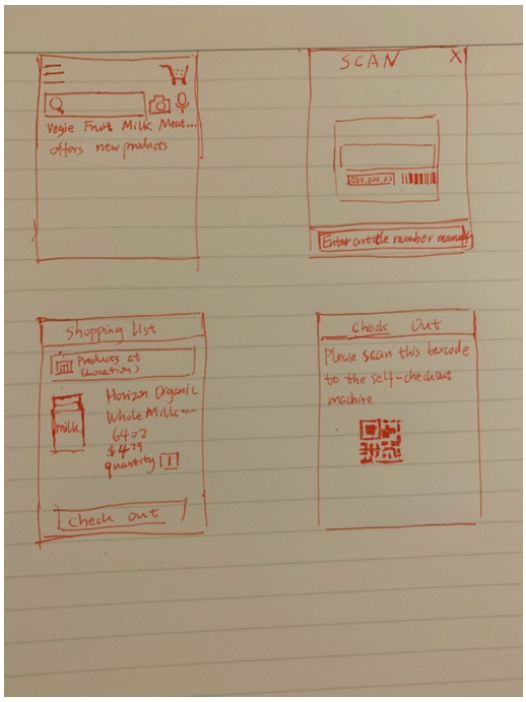

Figure 1