Assignment M4:

Yasaman Vaghei
yvaghei3@gatech.edu

***Abstract*—Keyboards are currently the most common electronic input device people use when typing text on a computer. Despite the large number of users worldwide, the keyboard designs only target the English proficient speakers. Although it is possible to change the language of typing on a computer, the keys on the keyboard only show the English letters. This poses a huge burden on people who desire to type letters of a different language. The current project aims to investigate this issue, find the needs of the user, propose new ideas to efficiently re-design the keyboard layout, and finally evaluate them systematically.**

1 QUALITATIVE EVALUATION

For the qualitative evaluation, I am going to use a survey to evaluate my textual prototype. This prototype was fully explained as the first prototype in my M3 assignment. The design of this prototype is similar to the typical mechanical keyboard but allows for replacing the keys. Providing keys with different language-specific letters written on top, satisfies the need to type in different languages while addressing the user's needs, identified at the need finding stage.

1.1 Procedural Information

In the evaluation process, participants will be recruited from my classmates, friends, and co-workers. These participants are required to be at least bilingual and have experience to use a type of keyboard (including mechanical, on-screen, and virtual) to type in at least two different languages (including English).

The recruitment will be initiated by sending an email to all my classmates, friends, and co-workers, briefly explaining the study and inclusion criteria and requirements. Potential participants are asked to get back to me by email, mentioning their availability. I will then confirm the time with them.

The evaluation process will take place in a room where only me and the participant are present at my work office, or a library room booked at school.

No voice or video from the evaluation session will be recorded. The participants only fill out an anonymous survey of the textual prototype.

1.2 Content

Since a survey is going to be used as the method of evaluation, detailed survey questions are outlined below:

How satisfied are you with the visualizations that this design provides?

- Highly Satisfied
- Satisfied
- Neutral
- Not Satisfied
- Highly Dissatisfied

How satisfied are you with the current design for typing special characters?

- Highly Satisfied
- Satisfied
- Neutral
- Not Satisfied
- Highly Dissatisfied

How comfortable is it physically to use this keyboard design?

- Very Comfortable
- Comfortable
- Somewhat Comfortable
- Neutral
- Not Comfortable
- Hard
- Very Hard

How easy it is to switch between languages with this design?

- Very Easy
- Easy
- Somewhat Easy
- Neutral
- Not Easy

    - Hard
    - Very Hard
- What is your overall satisfaction rate with this design?
    - Highly Satisfied
    - Satisfied
    - Neutral
    - Not Satisfied
    - Highly Dissatisfied
- If you have any additional comments and suggestions, please provide them below.

1.3 Inventory and Requirements

The requirements identified in the need finding assessment are as follows: 1) the need for being able to visualize the language-specific characters on the keys, 2) the requirement for eliminating the need to press multiple keys at a time to type special characters, 3) the requirement for the design to be physically comfortable, and 4) the requirement for the design to provide for easy switching between multiple different languages. For each of the needs, a specific question is included in the survey. The assessment acquired using surveys can evaluate whether the identified needs are addressed.

2 EMPIRICAL EVALUATION

The prototype that I selected to evaluate empirically is my second prototype introduces in my M3 assignment. This prototype is similar to the design of a typical keyboard; however, a knob is embedded in the right corner to change the language. Turning the knob results in changing the language-specific letters written on the keys. The keys have a cylindrical shape and roll by turning the knob. Rolling a cylindrical-shaped key results in the appearance of a different language-specific letter. More details were previously provided in my M3 assignment.

The goal here is to compare the aforementioned design with the current typical design of the mechanical keyboard.

2.1 Control and Experimental Conditions

At this evaluation stage, the goal is to test whether the users are more satisfied with the new design (or prototype). For this purpose, I am planning to run a

satisfaction survey for both the current and new designs. I will then investigate whether the satisfaction rate difference is statistically significant between the two designs. This way, the current design can be identified as control condition and the new design can be identified as the experimental condition.

The satisfaction survey examines the user's satisfaction rate in terms of visualizations, ease of use, inclusiveness, compatibility with different modalities (e.g., Windows and Mac), and ease of use. The overall satisfaction rate will be quantitatively measured across the aforementioned terms. This enables one to compare the new design to the current typical keyboard design for each of the terms defined previously.

2.2 Dependent Variable

In this experiment, the goal is to measure and compare the satisfaction rate in different categories. These categories are highly correlated with the previously identified needs and are explained as follows: 1) satisfaction of the user with the visualizations (e.g., being able to visualize the language-specific letters when typing), 2) satisfaction of the user with ease of use (e.g., How easy it is to use the keyboard physically? How easy it is to learn using the newly designed keyboard), 3) satisfaction of the user with compatibility (e.g., is it possible to for the user to employ the keyboard in Mac and Windows settings?), 4) satisfaction of the user with ease of use (e.g., How easy is it for the user to learn working with the new interface?).

2.3 Hypotheses

The null hypothesis is defined as there is no statistically significant difference between the satisfaction of the users with the current typical mechanical keyboard and the new design.

The alternative hypothesis rejects the null hypothesis and specifies that there is a statistically significant difference between the satisfaction of the users with the current typical mechanical keyboard and the new design.

2.4 Design

In order to test the null hypothesis, a within-subjects design is considered. This means that all subjects test both designs (i.e., current typical keyboard and the new design).

2.5 Method

In the experimental protocol, the recruited subjects are assigned to examine both designs in a random order. This way, biases regarding testing one design prior to the other will be eliminated.

The task defined for each of the users is to sit at a desk, in front of a PC monitor and a keyboard attached, change the language for typing and start typing a few sentences in different languages.

A satisfaction survey will be given to the users to assess their satisfaction in each of the previously mentioned categories. For each category, five possible answers will be provided (i.e., highly satisfied, satisfied, neutral, dissatisfied, highly dissatisfied). In order to obtain the satisfaction rate for each of these categories, each of these answers will be given a weight respective to their order (i.e., 100, 75, 50, 25, 0).

2.6 Lurking Variables

The lurking variables in this experiment include the environmental noise (i.e., there might be noise coming from the surrounding environment that cannot be controlled for), disturbances created by the user's cell phone (i.e., notifications appearing on the user's cell phone cannot be controlled for in the real environment), and the mental state of the user at the time of experiment (i.e., some users might be fresh at the time of the experiment and therefore be able to interpret and work with the interface better. However, a user who is tired is more prone to making slips or mistakes and therefore not being satisfied with the new design for instance).

2.7 Analysis

The satisfaction rate is measured for separate categories. For each of these categories, the answers are defined as ordinal (i.e., highly satisfied, satisfied, neutral, dissatisfied, highly dissatisfied). Therefore, the appropriate test used for the analysis is Chi-Squared method.

3 PREDICTIVE EVALUATION

For predictive evaluation, I am going to evaluate the third prototype defined in my M3 assignment. The design of this prototype is similar to the current typical

mechanical keyboard. The main difference is that a small LED screen is provided on each of the keys to demonstrate the language-specific letters. The screen changes when changing the language for typing. Additional keys are also provided to type in special characters. The details of this wireframe prototype are given in my M3 assignment.

This wireframe prototype will be examined by means of the cognitive walkthrough to better understand the user's thought process while using the newly designed interface.

3.1 Task Description

In this section, more details are given on the task. The task at hand is to interact with a keyboard's interface to type in different languages. The method used for changing the typing language remains the same as current.

The task starts after changing the language. In order to perform the task, user requires to first identify the desired key, then identify the requirement to press it to type the desired letter, and finally press the previously identified key. These can also be considered as the elements in gulf of execution.

After that, the user visualizes the typed letter on the screen. They then interpret what the letter typed is and finally identify if the letter typed is the one desired. These can also be considered as the elements in gulf of evaluation.

In this evaluation procedure, I will conduct experiments to evaluate a user's thought process and navigation around the keyboard interface. It is important to understand how the users find the desired letters or identify and learn the adjustments made (i.e., additional keys) to type special characters.

3.2 Goals

In this task, the goal of the user is to type in their desired languages using the newly designed keyboard. There are two main categories emphasized when evaluating the prototype: efficiency, compatibility, and learnability. The design requires to effectively address the user's needs with visualization and physical comfort. It also matters that the system is compatible with both Mac and Windows systems. In addition, the design requires to be easy to learn. The user is expected to identify the items in the gulf of execution and evaluation easily.

3.3 Operators

The operator available to the user is the keyboard, which has keys each embedded with a small screen to visualize the language-specific letters. Changing the typing language on the keyboard results in changing the letters shown on the small sized screens on the keys. These keys can be pressed to type in the desired letters.

4 PREPARING TO EXECUTE

The evaluation methods that I select to perform are qualitative evaluation and predictive evaluation.

The reason for selecting the qualitative evaluation is to better understand "why" the users perform specific tasks or subtasks or "how" they perform those tasks or subtasks. This evaluation method is also a great way to not only understand the user's train of thoughts, but also develop a better understanding of the user's satisfaction rate. Using this evaluation method also gives an advantage of getting feedback from actual users.

The main reason for selecting the predictive evaluation is to better understand the user's thought process. It also saves time because it does not involve participant recruitment. This evaluation method also enables one to gain more information about ongoing design decisions.

The reason for not selecting the empirical evaluation is that with the limited time I have, I am not able to recruit a large number of participants for the results to become more reliable. Also, my prototype is not ready for empirical evaluation.