

FORMATIVE EVALUATION

Whiteboard: When Blackboard lends an ear

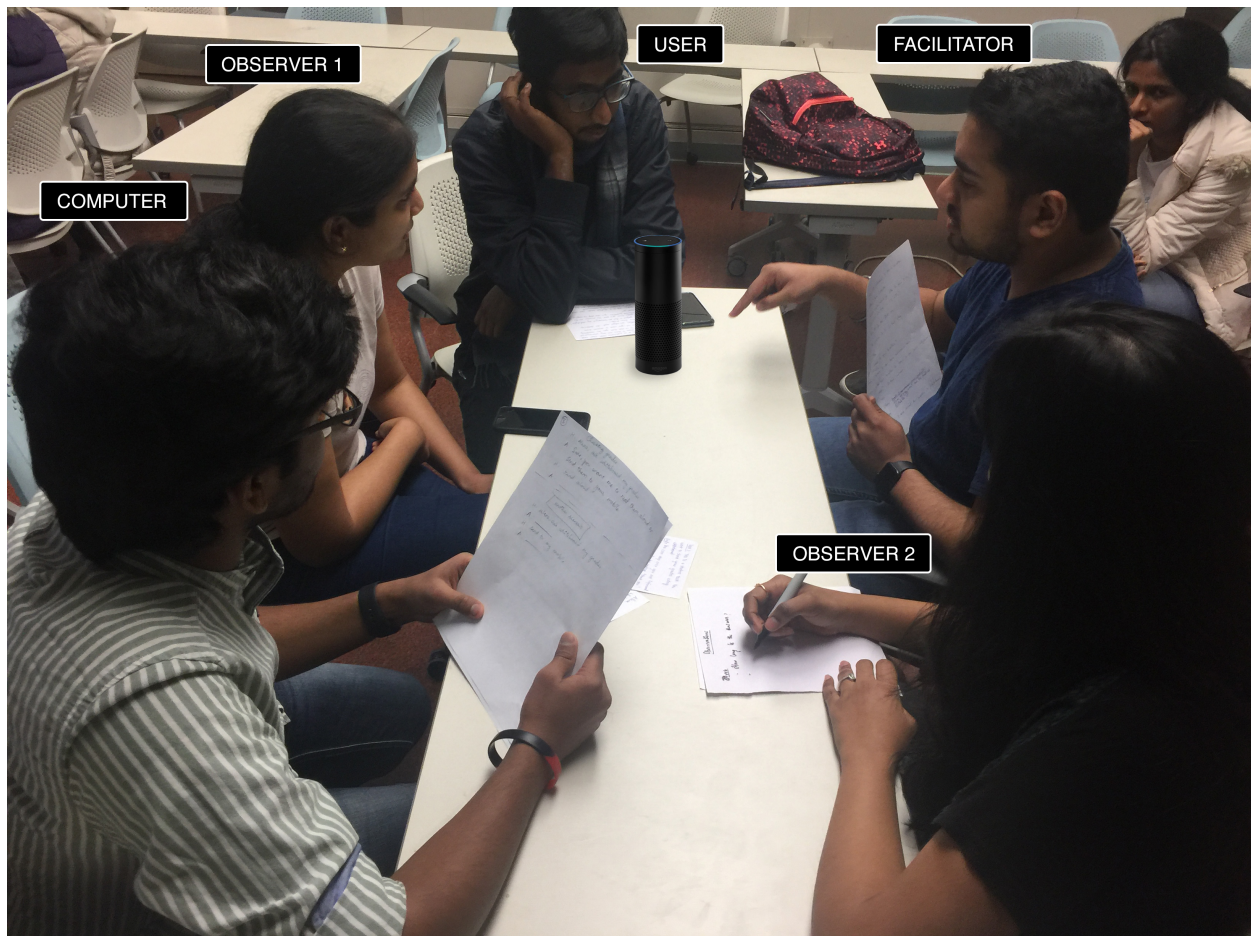
Team: Gautam Krishnan, Sai Priya Jyothula, Sumanth Reddy Pandugula, Varshini Sampath

INTRODUCTION

The formative evaluation technique that we used is **Wizard of Oz**, a technique in which users interact with a system that users believe to be autonomous, but is actually being operated or partially operated by an unseen human being. The Wizard of Oz evaluation is carried out with 4 people - the user to participate in the experiment, facilitator to explain the user what tasks to perform and take consent, observer to record the voice/video and take notes and computer to simulate the actions of the computer/system.

The application that is being evaluated is **Whiteboard**, a skill of Alexa (the voice assistant on Amazon devices such as Echo, Echo Dot or Echo Tap).

WIZARD OF OZ SETUP



The image shows the facilitator explaining the user how to interact with Alexa, introducing the skill 'Whiteboard' and what tasks to perform. Observer 1 is recording the entire conversation and observer 2 is taking notes. The computer is waiting for the user's initiation.

To give the user the illusion that they are talking to Alexa on an actual Amazon device such as Echo, we placed an Echo in front of the user (gives visual feedback by lighting up when the user says the initiation command "Alexa, ..."). We reduced the volume of the Echo to zero and the voice of a human answers the user (audio feedback) when the initiation command is uttered by him/her. For the tasks where the user might get the results on his/her mobile, we made a paper prototype of how the user's mobile display should look like and placed the paper on top of the user's mobile to simulate Alexa cards on their mobile. We performed the formative evaluation with 5 users (three of our classmates who are students, one user who is both a TA and a student and the fifth user is an instructor).

LIST OF TASKS

Task 1: This is an Instructor task. Your task is to ask Alexa if there are any unanswered questions in CS 522 that you are teaching.

If there are, you want to answer them.

Information you might need for follow up questions from Alexa:

"The project is due on January 6 by midnight"

End goal: Once you post the answers to the unanswered questions, you can stop by saying 'thank you'.

Task 2: This is a Student task. You want to post a question using whiteboard.

Information you might need for follow up questions from Alexa:

Course is 522.

Question you want to post:

"When is our next assignment due?"

End goal: Once you post the question, you can stop by saying 'thank you'.

Task 3: This is a student task. You want to check your grades using whiteboard.

End: You can stop once you are informed of your grades by saying 'thank you'.

Task 4: This is a student task. Ask if Prof. Debaleena posted a homework.

Information you might need for follow up questions from Alexa:

Course is CS 522.

End goal: Once you got to know whether the homework is posted, you can stop.

We performed two iterations of the task and evaluated both of them for Task 1. This is because in Task 1, Alexa provides the user with two options: read the unanswered questions aloud or send them to the user's mobile. We gave the user the freedom to choose whichever option he/she wants in the first iteration of the task. In the second iteration, the facilitator asked the user to select the option that he/she left out in the first iteration. Similarly, Task 3 also has two options, read aloud or send to mobile. We performed two iterations for this task as well. This way, we made sure that we didn't leave out any use cases in the evaluation.

INSIGHTS FROM THE FORMATIVE EVALUATION

- **Repeating the options:**

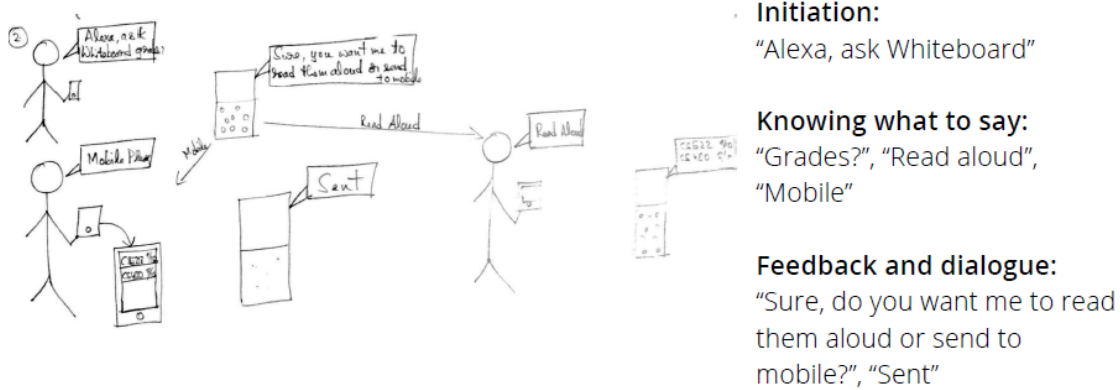
(Initial design) When a user asks to post a question (Task 2), Alexa helps the users by reading out the list of courses that are available for the user. The user then recognizes the option he/she wants and chooses one of them.

(Observation during evaluation) We observed in our evaluation that one of the users missed out on the options (momentary lapse in concentration) and asked Alexa to repeat the options. The same case applies when the user wants to check his/her grades (Task 3). The user might want to Alexa to repeat the available courses.

(Modifications after evaluation) We never considered this aspect in our design iterations, we will take this into consideration in our design so that Whiteboard can repeat the available options whenever needed.

- **Checking grades:**

(Initial design) We had three design alternatives for this particular design challenge. We assumed the simpler the better and finally chose the alternative where the user would simply ask for his/her overall grades as shown in the sketch below:



(Observation during evaluation) During the evaluation for Task 3, majority of our users were trying to find the grades for their recent homework or the overall course grade. Also, some of them wanted to know the class statistics as well.

(Modifications after evaluation) An alternative design best fits this purpose for this task rather than the chosen design sketch, so we will be modifying our design to the one shown in the next page:

User is a student, at home:



- **Checking for unanswered questions:**

(Initial design) When the user is checking for unanswered questions (Task 1 - Instructor task), Alexa asks whether to send the questions to mobile or read them aloud.

(Observation during evaluation) The above interaction is fine as long as the questions are small, but sometimes the questions can be longer (more than 100 words) which makes things a little tricky. A user tried to ask Alexa questions such as “how long are they?” and “how many are there?” before he/she made a decision about sending them to mobile or reading them aloud. We also looked at the questions from the Piazza discussion group of CS 522 and found that some of the questions are very long. Our designs did not cater for above situations before but we will let the user know of this information before asking the user which option they want.

(Modifications after evaluation) We will approach this problem with automation (CS 522 HW2 Question 6). Alexa checks for the number of questions and length of each question, and if it is long then Alexa tells the user “There are questions that are very long/There are too many questions (if more than 4 questions). Do you want me to read them aloud or send them to your mobile?”

INSTRUCTOR FEEDBACK

- **Make different tasks for read aloud and send to mobile:** We let the users to chose the option of their choice on their first iteration and the facilitator asked the user to repeat the task using the other option in the second iteration.
- **Do not introduce the computer to the user or tell who is playing which role:** We never introduced the computer (role) to the user, the facilitator just explained the user about Alexa, Whiteboard and the tasks. However, we took the user’s consent to record the voice/video and take notes before the evaluation.
- **Do not give the initiation command:** We stopped giving the initiation command as a part of the task. The facilitator explained every participant about how to initiate Alexa to use the skill Whiteboard. The users were only told what task to perform and they framed the commands themselves.
- **End goal:** We clearly provided an end goal for every task - when the task will be accomplished and the user has to stop.
- **Test what the user might want to do:** We did not give the user any commands but just told the users in a more abstract manner what they want to achieve and observed how they did it. We specifically observed what they did in addition to or in contrast to what we thought they might do.

TEAM INVOLVEMENT

Role Switching: We switched our roles for different users so that all of our teammates are familiar with the duties for all roles.