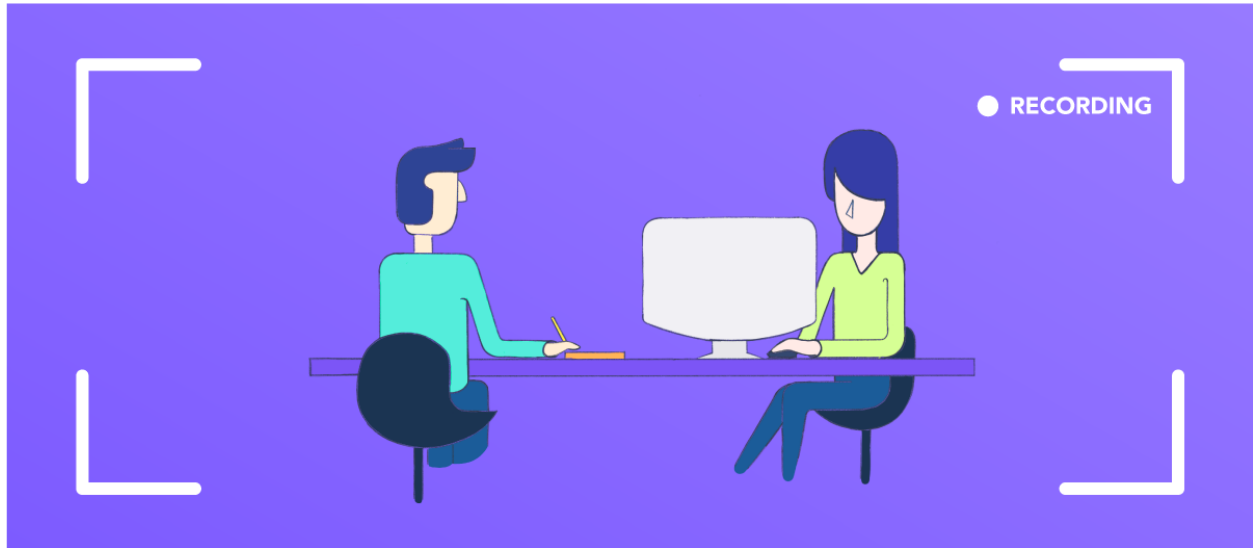


Dialogflow γ (3 Points)

Usability Testing



[Image source](#)

In this assignment, you will design and carry out a *mini* usability test of your Dialogflow deliverable, *the shopping assistant*, in three parts:

Part 1—Designing A “Mini” Usability Test (0.8 Point): In the first part, you will make some decisions on the *why*, *what*, *how*, and *whos* of the study and write a two-page test plan that reflects your decisions.

Part 2—Executing Test Plan (1.4 Points): Next, you will recruit two volunteers from among classmates, family, and friends who can help you with your testing, and you will execute your test plan, over videoconferencing, to collect quantitative and qualitative data on the use and experience of the shopping assistant.

Part 3—Analyzing & Reporting Findings (0.8 Point): Finally, you will analyze your data and translate your findings into design insight.

Submission Details

Your deliverables for the assignment will be your test plan from Part 1, the data you collected in Part 2, and a report of your findings and a discussion of their design implications in Part 3, all as a single PDF document submitted to Canvas.

Note: Your assignment will be graded on the contents of this report and not the usability of your system. If you find that your agent is hard-to-use or unintuitive, you can be honest with your outcomes.

Part 1: Designing A “Mini” Usability Test (0.8 Point)

In this part, you will make some decisions about the format and design of a brief *formative* usability test and develop a *test plan*. First, you will determine two desired outcomes for your study. You can choose from five Es we have discussed in class (*effective*, *efficient*, *engaging*, *error tolerant*, and *easy to learn*), the three dimensions of the ISO definition of usability (*effective*, *efficient*, *satisfactory*), or related concepts or outcomes (e.g., desirability, learnability, discoverability) that best fit to what you would like to evaluate. These will serve as your desired outcomes. Next, for each outcome, you will develop *questions*, *tasks*, and *scenarios* that will guide your testing. Then, you will choose two metrics: one performance, one self-report. Your deliverable will be a test plan that communicates these decisions and serves as a guide for the moderator (you) to run the test. Your study should be in the form of a remote *moderated* usability test conducted over videoconferencing, e.g., Zoom. The steps in the checklist below will help you in your decision-making and writing of your test plan and the form below that will help you draft your test plan. Your test plan should not exceed two pages.

Usability Test Design Checklist

- ☐ Choose two intended **outcomes**, e.g., effective, efficient, engaging, error tolerant, easy to learn, usable, satisfactory, etc.
 - ☐ For each outcome, formulate a **question**, e.g., “To what extent are users satisfied with the shopping assistant” or “What is the overall usability of the shopping assistant?”
 - ☐ For each question, devise a **task** using your shopping assistant that can help you assess how well your design meets the outcome. The task description should capture what you expect the users to do to successfully perform the task.
 - ☐ For each task, develop a **scenario** that will provide context and guidance to the user. The scenario should prompt the user to perform the task you developed.
 - ☐ Choose two **metrics** for measurement: one performance, one self-report. Examples of performance measures include task success (e.g., number of task substeps completed), time (e.g., seconds), or errors (e.g., number of deviations from expected use). For self-report measures, you can use the SUS questionnaire or all or part of the USE questionnaire.
 - ☐ Templates for [SUS](#) and [USE](#).
 - ☐ Write out your **test plan** using the form on the next page. Your plan should have three sections: (1) overview, (2) study design, and (3) test procedure. The overview section will briefly describe the context (including the “what” of the usability test, i.e., the scope of your interim or final design), the general goals for the testing, and the intended outcomes of the test. The study design section will outline your questions, tasks, and scenarios and your metrics. In test procedure, you will provide a step-by-step plan for the test in the form of a checklist.
 - ☐ You can see an example usability test plan from Barnum (2011) [here](#). Your plan will not be as detailed as this example and should be *at most* two pages.
-

Usability Test Plan¹

Overview

The goal of this Usability Test will focus on the Efficient and the Error tolerant dimensions. Two people will participate in this experiment by zoom. They will be asked questions and operate certain functions in the test. It will test the efficiency of the dialogFlow compared to traditional mouse and keyboard operation to the WiscShop website. Besides, it will find out what do users think about this shopping assistant? How will the user experience be? And what is the user's overall satisfaction of it?

Questions:

1. Is it more efficient to use the shopping assistant compared to just using the website with a mouse and keyboard? (efficient)
2. What is the user's overall satisfaction with the shopping experience? What part/function makes you satisfy, what does not? (satisfaction)

Study Design

1.Is it more efficient to use the shopping assistant compared to just using the website with a mouse and keyboard?

The user will be doing a series of activities. (1) login (2) add hat to cart (3) add a jogger to cart (4) check the review of a sweatshirt (5) see how much plushies are for sale (6) remove hat from cart (7) confirm check out.

Both users did not have experience using this website before. They will be doing the same activities using the shopping assistant and mouse. The time they complete the task will be counted and analyzed after the test.

2.What is the user's overall satisfaction with the shopping experience? What part/function makes you satisfy, what does not?

The user will be exploring the shopping assistant for 5 minutes freely. They can do whatever they think about and want to try. They will be told at the beginning to write down points they are satisfied with the assistant and the points they are not satisfied with. Finally they will give a self_report of how they feel about the shopping assistant and a final score they want to give (Max 10).

¹ Or use the [Usability Test Plan template](#)

Measurements:

1. Performance

Measured by time. The total time they take will be compared with the time they will mouse and keyboard to navigate the website.

2. One self-report

Not fully quantitatively. Users will describe how they feel about the assistant and give a final score. The score will only be an add on.

Test Procedure

<Describe the procedure you will follow in the form of a checklist for the study moderator>

1. Since my Dialogflow B was not fully functional. I hardcode some of the functions, also some of the functions only work in the Dialogflow chatbox. Users will need to open the Google Dialogflow and BadgerShop Website before the test starts.
2. Doing the first question. Ask users to do a series of activities and record the time.
3. Doing the second question. Ask users to explore the shopping assistant for 5 minutes freely and set a 5 minutes clock.
4. Final question and feedback

*Some of the function time will be based on Google Dialogflow reply time.

Part 2: Executing Test Plan (1.4 Points)

In this part, you will identify two volunteers to help you test your shopping assistant over videoconferencing, e.g., Zoom, Microsoft Teams, Webex, etc., choosing a system that allows remote control of your computer (see documentation on conducting remote sessions where you give control of your computer to your partner for [Zoom](#), [Teams](#), [Webex](#)). They can be your classmates, friends, or family members. It is acceptable to pair up with a classmate and trade taking each other's test. You can use any version of your shopping assistant as long as you have a working prototype and choose to focus on any aspect of it. You can capture performance measures during the test, e.g., by timing them, counting errors, taking notes, or by recording them and watching later. You can present self-report measures on paper or on a computer screen after they perform all scenarios. Finally, be sure to make qualitative observations and ask questions, e.g., "you seemed surprised by that response, what were you expecting," to your participant where appropriate during and/or after the study. The deliverable for this part will be your data in table and/or text format pasted below. For performance, questionnaire, and qualitative data, provide the raw numbers or text that you will later organize and analyze in Part 3.

User1: Steven, A graduate student at UW Madison, By Zoom

User2: Isa, My friend, Meeting in person

Task1:

(1) login (2) add hat to cart (3) add a jogger to cart (4) check the review of a swearshirt (5) see how much plushies are for sale (6) remove hat from cart (7) confirm check out.

	User 1	User2
Activity 1 total (by shopping assistant)	3: 15	*2: 26
Activity 1 total (by mouse and keyboard)	1:01	1: 10
Task1 by assistant	It takes long to import user name and password	User got stock, so I paused the time to explain a little.
Task2 by assistant	No problem	No problem
Task3 by assistant	User do not realize they do not need to go to the item page to operate	User did it successfully and relatively quick
Task4 by assistant	No problem	No problem

Task5 by assistant	User go the the tag quickly	No problem
Task6 by assistant	User go to cart first	User go to cart first
Task7 by assistant	No problem	No problem

Task 2:

User1: Over all score 8/10

User said he is satisfied with the assistant. He mentions it is better than he thought it would be. He is also good with the extra time to spend on the shopping assistant. He feels amazing that the assistant can do a lot of jobs by getting a single sentence. One thing he mentions is that a voice assistant better than a mouse and keyboard is that the assistant can access the other pages while the user stays on the same page. The reason why he took points off is that:

1. When doing some of the jobs, like confirm check out, there is no reason to use it instead of clicking a single button.
2. The login part is a little confusing. The assistant needs to have more response to tell the user how to input information. Or it could have an instruction at the beginning.
3. Some of the other features need to be explained more such as users do not need to go to the cart first to operate the cart.

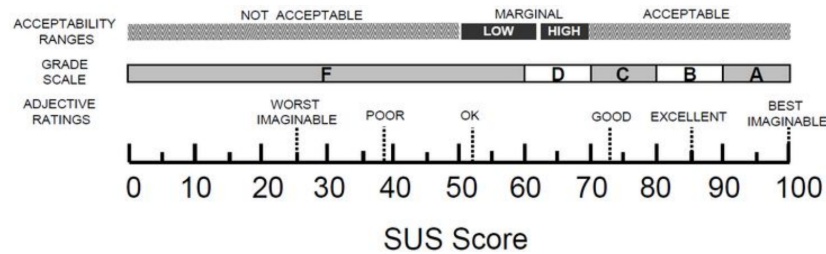
User2: Overall score 8/10

The user also thinks this shopping assistant works. She feels it has plenty of functionality. She likes the quick navigation using the assistant especially when you are not familiar with the website layout. She went to do some complex input in the last part of the free test section. Then she found the shopping assistant had only handled a single operation in one input. Like she tries if the user input "add the best reviewed hat in the cart" or "add this to cart and check out". The system will not get all the information. This is one of the reasons why she took points off. She mentions the other thing that can be improved was also the login task. She feels that part needs more instructions at the beginning.

Part 3: Analyzing & Reporting Findings (0.8 Point)

In this part, you will clean, consolidate, and analyze your results and translate them into design insight. For your quantitative data, calculate the average values from your metrics and report the averages. For self-report data, if you used SUS, follow the scoring method included in the template and give your shopping assistant a grade (e.g., "D") and level of acceptability (e.g., "high marginal") using the guide

below.² If you used a subscale of USE, such as “ease of use,” average out the scores for all items to arrive at a single value and average out the values for both of your test participants. For qualitative data, categorize your notes and observations into a minimum of two high-level findings. If the quantitative data or the qualitative comments from your two participants vary significantly, you can also comment on these differing views. Report your findings in narrative form and end your report with high-level design insight and recommendations for how your shopping assistant might be improved. Your report should not exceed a page.



² Based on Brooke, J. (2013). [SUS: a retrospective](#). *Journal of usability studies*, 8(2), 29-40.

Usability Findings

Quantitative Summary

Task1 result	User 1	User2
Activity 1 total (by shopping assistant)	3: 15	2: 26
Activity 1 total (by mouse and keyboard)	1:01	1: 10

User 1 takes **2 mins 14 secs more** on the shopping assistant than traditional operation. User 2 takes **1mins 16 seconds more** on it. Two users complete time has a **49 secs difference**. From the data we got, we can know that the time a user spent to complete a series of tasks using the shopping assistant is about **2-3 times** more than using the mouse and keyboard. This result was from two users who have not used this website and this dialogFlow assistant before. It also appears that different people can have a relatively large time difference on using the assistant.

Qualitative Summary

I recorded more users' feedback on the top section. They also point out some common problems. To summarize, I think the two most import find are:

1. The login function is a little confusing. It still needs more instructions at the beginning to tell users what information they need and how they input the information. For example, see the username and the password together or separately.
2. Users can only do one thing at a time with this shopping agent. They cannot say “add this item to cart and check out” to do two actions at the same time.

Conclusions

Thanks to Google, the agent can get most of the information the user says or entered. Users' overall feedback to the shopping assistant was not bad. The assistant has some advantages when users go to the website for the first time and have a hard time to find some functions. Also, users will be more comfortable to use them if they are using them more often. They will also be more likely to use it if the efficiency could be improved to get more close to the mouse and keyboard input.

The improvement will focus on three parts. First, try to improve the overall efficiency of the agent, focus on the tasks voice assistant has advantage over the traditional input method. Second, fix the login function. Add more instructions at the beginning and the feedback. Third, make users can have more than one simple movement in one input, which could also improve the efficiency.