

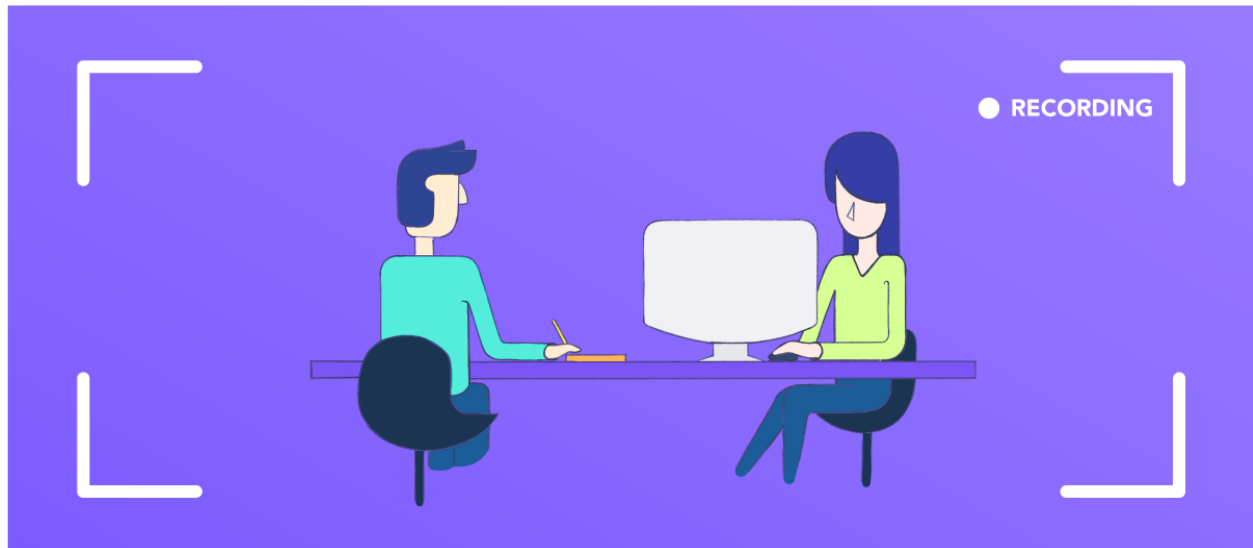
Dialogflow



(3

Points)

Usability Testing



[Image source](#)

In this assignment, you will design and carry out a *mini* usability test of your Module 3 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

<Describe the general context and scope of your study>

This study will be doing remote testing with two peer classmates. The scope will be to assess the functionality of the shopping assistant to help the user select and purchase products from the WiscShop website and providing a welcoming and supportive environment for the user to interact with.

<Outline study goals, including intended outcomes, in the form of questions>

1. What is the overall effectiveness and accuracy of the shopping assistant?
2. To what extent are users satisfied with the shopping assistant for the entire experience?

Study Design

<For each question, describe the task and the scenario>

1. The task is to let user purchase target products using the shopping assistant and check out the shopping cart successfully and accurately. The scenario is: the user is to purchase a pair of leggings, a plush with the bucky badger icon, and a red t-shirt for Thanksgiving gifts from WiscShop using the shopping assistant
2. The task is to enable the user to abandon certain products in the cart as well as providing a welcoming and supportive persona of the shopping assistant to talk to the user. The scenario would be the user is to buy four white visor for his or her squad, but decide to only take three hats before making the payment and want to buy something more.

<Describe your measures>

1. Evaluate the task on whether it is accurately completed based on the accuracy and correct number of task sub-steps completed at the end with or without undetected error
2. Use satisfaction interview questions (SUS questionnaire) at the end to learn from the user to what extent do they like the experience or not

Test Procedure

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

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

The test will begin with an overview briefing, followed by the scenarios, and then post-task questionnaires.

Overview

The moderator will welcome the participant and give him/her a general description of the shopping assistant. Next, the moderator will explain the designed scenarios and ask the participant to complete the tasks step by step. The moderator will be open to questions and proceed with the test.

Scenario 1:

You are graduating this December and this is your last semester at UW-Madison. You are going back home for Thanksgiving and won't return because of the covid-19 pandemic. You want to thank your family for their support by buying some souvenirs for them as Thanksgiving gifts. And you heard that WiscShop is a good place that exactly sell these stuffs. Your younger sister always wants a bucky badger plush and you also want to buy a pair of leggings for your mom and a red top for your dad.

Scenario 2:

You and your friends are forming a squad and you would like to wear the same hats for your squad. You have already added all four hats for your friends and yourself in the cart the other day, but today you just got a message from your friend saying that his brother has a totally same visor and you don't have to buy the one for him. Thus, you want to keep only three visors, at the same time you also want to buy something else using the money for the extra visor.

Closing:

The moderator will ask the participant to complete the post-test SUS questionnaire.

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.

<your data>

Scenario 1:

Time:

User 1: ~7min

User 2: ~ 4min

Errors:

User 1:

- There's no follow-up response from the greetings
- When user doesn't know the exact name of the product, the agent can't recognize it
- The user can't directly search for a product with both the tag and category specified e.g., instead of saying "show me all the plushes with a bucky badger icon", the user has to search separately for "all plushes" and "the ones with a badger icon"

User 2:

- When there's only one product under a category, the agent isn't able to add it to cart without user specifying the product name, e.g., the Jump Around Shirt under Tees

Scenario 2:

Time:

User 1: ~ 5min

User 2: ~2min

Errors:

User 1 & 2:

- When removing an item from the cart, user have to provide the number of items they want to remove and the product name separately, e.g., instead of saying "Remove 1 White Wisconsin Visor from my cart" and all set, user is asked "how many do you want to remove?" even though the number is already provided

Questions:

1. What would you expect the system to react when you provide both the number and the product name?

- For both adding to cart and removing from cart functions, the user can directly specify both the number and the name for the product they want to add or delete, it feels a bit cumbersome when the system has to ask twice when user has already provided the number in one sentence.

2. When adding one item to cart, what do you expect the agent to infer or understand your choice when you don't provide the full name of a product?

- When there's only one product shows up there after I set the filter, I assume the agent knows that my only choice is to pick the product on the screen
- When user says the name wrong for some reason, I think it's better for the agent to infer from the wrong name user provided and search in the database for a possible match and ask the user whether they are looking for any of the product in those possible matches

Satisfaction:

User 1: $24 \times 2.5 = 60$

# Questions						
1	I think that I would like to use this system frequently.	Strongly disagree	1	2	3	4 5 Strongly agree
2	I found the system unnecessarily complex.	Strongly disagree	1	2	3	4 5 Strongly agree
3	I thought the system was easy to use.	Strongly disagree	1	2	3	4 5 Strongly agree
4	I think that I would need the support of a technical person to be able to use this system.	Strongly disagree	1	2	3	4 5 Strongly agree
5	I found the various functions in this system were well integrated.	Strongly disagree	1	2	3	4 5 Strongly agree
6	I thought there was too much inconsistency in this system.	Strongly disagree	1	2	3	4 5 Strongly agree
7	I would imagine that most people would learn to use this system very quickly.	Strongly disagree	1	2	3	4 5 Strongly agree
8	I found the system very cumbersome to use.	Strongly disagree	1	2	3	4 5 Strongly agree
9	I felt very confident using the system.	Strongly disagree	1	2	3	4 5 Strongly agree
10	I needed to learn a lot of things before I could get going with this system.	Strongly disagree	1	2	3	4 5 Strongly agree

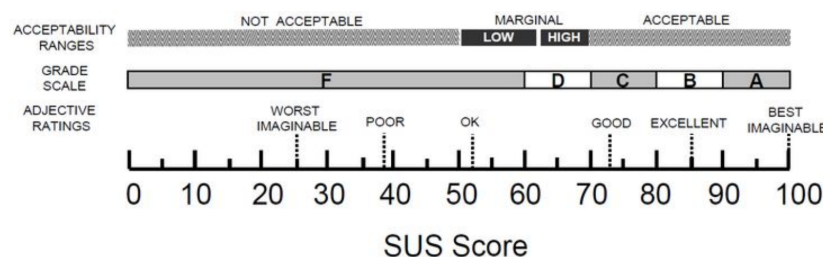
User 2: $35 \times 2.5 = 87.5$

# Questions						
1	I think that I would like to use this system frequently.	Strongly disagree	1	2	3	4 5 Strongly agree
2	I found the system unnecessarily complex.	Strongly disagree	1	2	3	4 5 Strongly agree

3	I thought the system was easy to use.	Strongly disagree	1	2	3	4	5	Strongly agree
4	I think that I would need the support of a technical person to be able to use this system.	Strongly disagree	1	2	3	4	5	Strongly agree
5	I found the various functions in this system were well integrated.	Strongly disagree	1	2	3	4	5	Strongly agree
6	I thought there was too much inconsistency in this system.	Strongly disagree	1	2	3	4	5	Strongly agree
7	I would imagine that most people would learn to use this system very quickly.	Strongly disagree	1	2	3	4	5	Strongly agree
8	I found the system very cumbersome to use.	Strongly disagree	1	2	3	4	5	Strongly agree
9	I felt very confident using the system.	Strongly disagree	1	2	3	4	5	Strongly agree
10	I needed to learn a lot of things before I could get going with this system.	Strongly disagree	1	2	3	4	5	Strongly agree

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

<Calculate averages and summarize>

Scenario 1:

- Average Time: 5.5 min
- Average # of errors: 2

Scenario 2:

- Average Time: 3.5 min
- Average # of errors: 1

Average SUS score: 73.75 -> C

Qualitative Summary

<Categorize your notes and report a minimum of two findings>

The agent is not interactive enough to continue the conversation before logging in, as the user can not resume on the small talk with the agent. And the search functionality doesn't infer from user's input but require the exact information stored in the database, which is sometimes lengthy for the user to type the exact product name.

The difference between the views of the two users might be: 1. User 1 is completely new to this project as well as DialogFlow and thus might spend more time in familiarizing herself with the interaction and more likely to try different routes in progressing the interaction. Whereas user 2 is also a student of this class who completed the same project and is quite familiar with the API as well as how the interaction is expected to progress, thus he spent less time but also got a fixed mindset instead of trying different paths.

Conclusions

<Report high-level insight and recommendations for how to improve the agent>

The agent needs to include fuzzy matching to better infer product names rather than having to ask user to type in the exact product name. The user is generally satisfied with the agent feedback but there's more to improve upon in terms of efficiency so that user won't be asked redundant questions and the agent always remembers every piece of useful information given by the user.