

Assignment M4

Nicholas Ohl
nohl3@gatech.edu

Abstract—Launched on November 12, 2019, Disney Plus (Disney+) is a streaming service housing titles from "Disney, Pixar, Marvel, Star Wars, National Geographic, and More" (Hobson, Casano-Antonellis, and Schreibfeder, 2021). Plagued by launch issues (Span-gler, 2019), the Disney Plus platform lacks the maturity and user experience of other streaming platforms. In order to watch a program on Disney+, users have to be able to easily search for and find it within the platform. My research focuses on redesigning and improving the search interface of Disney+.

1 QUALITATIVE EVALUATION

For the qualitative evaluation, I plan to conduct a post-event protocol for my card prototypes. These prototypes are shown in appendix 6.1, and based on interviewee feedback I may rapidly create additional mockups during these interviews.

1.1 Evaluation Plan

My initial participant base will be friends and family members with varying levels of technological savviness. If necessary, I will expand to include classmates but will start with a convenience sampling approach. Because I will be interacting with people I know well, I need to be mindful of both social desirability and confirmation biases in my evaluation.

I plan to conduct these interviews in person for any non-OMSCS participants. This will allow for participants to interact directly with the card prototypes, more closely simulating interaction with the interface. With the consent of participants, I will video record a top-down view of their interactions with the prototypes so that only their hands are visible. These recordings will include audio to capture participant feedback, as well as the instructions I give so that I can analyze if my directions introduced any bias.

1.2 Evaluation Content

In addition to my current card prototypes shown in appendix 6.1, I will create a card mockup that reflects the current design of the interface. My intent here is to obtain qualitative feedback on both the existing layout and my proposed layout. To combat potential bias introduced by this, I will make sure that the current design is drawn with the same fidelity level as my other prototypes. I will not inform participants that one of the cards reflects the current system design, and will randomize which card sequence is shown first to each participant.

At the start of the evaluation session I will inform participants that I will be taking notes throughout the session on my computer or phone. I will also ask participants if they consent to having their actions and speech recorded. Whether I record or do not record the session based on their response, I will provide the following instructions:

- "I am evaluating different designs for Disney+'s television display and would like your help in providing feedback on proposed designs."
- "I am going to present you with a series of paper cards simulating what you would see on your TV screen in Disney+."
- "I would like you to pretend that you are using this physical remote control to navigate through the screens, pressing buttons as you would on your TV. For any buttons you press on the remote, please tell me what buttons you are pressing so that I can inform you what the output of your action is."
- "Your goal is to navigate to the search menu of the application and successfully search for the movie *Big Hero 6*".
- "As you perform this task, feel free to also speak any thoughts out loud in addition to your button commands. You can ask me for help if you feel stuck, and I will provide as much assistance as possible without explicitly telling you how to perform the task."

During this search task, I will gather empirical data on the amount of time and number of clicks it takes for participants to complete the task. While this is a qualitative evaluation, I want to get a quantitative sense of whether one layout is more efficient than the other. In particular, I am interested in the time it takes for the subtask of navigating to the search menu itself. The qualitative data I gather will be from responses to the following questions:

- How did you feel throughout the task?
- What did you like about the interface?
- What did you dislike about the interface?
- Was there any behavior in the interface that surprised you? For example, if the press of a button did not result in the output you expected? Is there anything you would change about the system?
- On a scale of 1 to 5 with 1 being very easy and 5 being very hard, how difficult was it to perform the task? What made it easy or difficult?

I will then have them repeat the search task with the other interface layout, again asking the same questions after completion. Additionally, I will ask which layout they prefer, and why.

After gathering feedback on the navigation workflow, I will present the participant with all of the keyboard layout cards at once, asking for their thoughts on the different designs. These are the questions I plan to ask:

- Where do you think is the best place on the screen for the keyboard?
- Do you prefer an alphabetical keyboard, a QWERTY keyboard, or a different keyboard? Why?
- Of the designs shown, please rank them from your favorite to least favorite.

1.3 Data Inventory and Requirements

Of all of the requirements (see appendix 6.4), the most critical one for my design is functionality. If participants are able to successfully complete the task without substantial assistance, then it will meet the baseline functionality requirement. In particular, the question about whether or not the interface behaved in a way that surprised users will evaluate whether the functionality was consistent with user expectations.

As my data inventory defines the users as anyone who would use the Disney+ service (see appendix 6.5), I plan to have both novices and experts for this evaluation. Two of the main requirements of the final design are learnability and discoverability, so at least one of the participants will be someone who has never used Disney+ before.

By having users perform the task on both the existing interface and my design, I will be able to determine quantitatively if the new design is more efficient, and qualitatively if navigating to the search menu itself is easier. As a more

discoverable search menu is one of my requirements, this evaluation will help determine if that goal is met. I designed my prototype to also be more consistent with the computer and mobile interfaces, which is a "nice to have" requirement. If participants prefer the new design, that would also meet this consistency aim.

2 EMPIRICAL EVALUATION

For the empirical evaluation, I plan to evaluate my textual prototype shown in appendix 6.2. While I would not be able to test the exact prototype I describe—that would require significant mobile application development and access directly to the internal servers of Disney+—I could simulate this prototype using an existing smart remote application that possesses the same core functionality.

2.1 Evaluation Conditions

For this evaluation method, the control condition is using a physical television remote to navigate Disney+ and search for a program, as that is the current workflow within the service. The experimental condition is performing the same workflow using my mobile remote prototype.

In the control condition, the independent variable is the physical remote. The independent variable for the experimental condition is the smart remote application. The dependent variable that I will evaluate is the time it takes to complete the task. Efficiency, particularly for expert users, is one of my defined requirements. I want to determine if the remote application is provably more efficient than the physical remote through this experiment.

2.2 Experiment Design

As previously mentioned, my independent variable is the type of remote, making it categorical data. Since time is a type of ratio data, the experimental method I will use for this evaluation is a student's t-test. More specifically, I will use a one-sided t-test as I hypothesize that my remote app prototype will be faster than the existing physical remote interaction.

As presented in lecture, if this were a two-sided test the null hypothesis would be that there is no difference in time while the alternative would be that there is a difference in time. Because I am conducting a one-sided test, my null hypothesis is that the amount of time it takes to search for a program in Disney+ using the remote app is greater than or equal to the amount of time it takes to search

using the physical remote. The alternative hypothesis is that it takes less time to perform the search task using the prototype than it does with the physical remote.

For this experiment, I will utilize a within-subjects design where participants experience both treatments so that the number of degrees of freedom is consistent across both conditions. I will conduct this t-test using a significance level of $\alpha = 0.05$, and aim to have at least 20 participants so that the critical t-value is reasonable (1.729 or lower).

I will recruit all participants prior to administering the experiment, and all sessions will be individual and in person. This is so that I can randomly assign participants to two subgroups; the first group will receive the control treatment followed by the experimental treatment, and the second will receive the treatments in the opposite order. This experimental design is in place to combat treatment order as a potential lurking variable.

Another lurking variable that is related to treatment order is general familiarity with the interfaces. If participants have not used Disney+ on a TV before, they would not know the sequence of inputs needed to accomplish the task. To address this, I will provide a walkthrough of the Disney+ interface when giving participants initial instructions, particularly how the search menu is located in a hidden sidebar from the default view. Additionally, I will inform them which buttons on the remote they will need to use and which buttons on the app they will need to use to move about the interface.

For both conditions, participants will start at the exact same place in the interface: the default home screen of Disney+, with the cursor focused on the banner at the top. All configuration of the TV, physical remote, and mobile device with the app will be completed prior to the experiment so that subjects only have to complete the task as instructed. The task that they will be instructed to complete for both conditions is to search for *Big Hero 6*.

There are two other lurking variables that I need to be aware of. First, if participants make error(s) during the workflow, that will increase the total time to complete the task. Errors could occur for either the control or experimental condition, and is something to note if it significantly impacts total time. Possibility for input errors is one of the reasons why I am focusing solely on time as a dependent variable versus number of clicks.

Second, input latency is possible for both the physical remote and the app. While there would ideally be no input delays, realistically there is some level of latency for both. If the difference between the two devices is on an order of magnitude of milliseconds it likely will not have an impact on the results, but if there is a noticeable input delay difference between the two I will need to take this into account in my analysis.

3 PREDICTIVE EVALUATION

For the predictive evaluation, I will evaluate my Wizard of Oz prototype shown in appendix 6.3. As mentioned in my requirements in appendix 6.4, my focus is on increasing learnability and discoverability for novice users. Thus, for this evaluation I will perform a cognitive walkthrough as my task analysis.

To kick off this analysis, I will prioritize the same core search task outlined in my other two evaluations: searching for *Big Hero 6* from the default menu within Disney Plus. The goal is to perform this search using only voice command operators, and is known in advance.

Here are some of the questions I plan to think about while performing this task:

- Do commands work how I would expect, or are there functionality issues? Do you have to provide confirmation to execute certain commands?
- How does the system respond to commands that it does not recognize?
- Is the voice dictionary comprehensive? Can it handle synonyms?
- Can users customize and add voice commands to the system?
- How easy is it for users to recover from mistakes? Can they undo or cancel a command?
- How do you clear search results? Can you only clear one letter at a time, or can you clear entire words or phrases?

In addition to searching for the specific program *Big Hero 6*, I want to more generally consider the full experience using voice command operators in the interface. If my prototype were incorporated into the search functionality of the service, users would expect to be able to use voice commands in all parts of Disney+. Thus, the secondary task to perform is the more abstract "Navigate and explore the entire Disney+ TV interface using voice commands."

With that task in mind, other questions to consider about the general user expe-

rience during the cognitive walkthrough include:

- Would novice users know how to exit the search menu? Can they get to other menus like the main menu or their watchlist?
- When watching programs, would the voice functionality be available? Could you pause, play, fast forward, or rewind the program?
- Would you have to press a button to activate voice functionality, or could it automatically pick up a command similar to "Hey Siri"?
- Could you change display settings or turn captions on or off from within a program using voice commands?
- Are there any features that should be restricted to not be changed via voice command, such as passwords or parental controls?
- Should the voice control capabilities entirely replace the existing controls, or supplement them? What would happen if the voice recognition stops working?
- What are the television hardware requirements of the voice functionality?

Finally, as discoverability and learnability are priorities in my design, I want to think about how easy it is to find and remember available voice commands. As outlined in the prototype, the first time the voice commands are available a brief message will display to users letting them know how to access the list of available commands. If users dismiss this message or forget this command, is it easy for them to later identify available actions and cross the gulf of execution? Does the system provide adequate feedback to help cross the gulf of evaluation? Can novice users learn the system quickly?

4 PREPARING TO EXECUTE

The two evaluations I will actually perform are the qualitative evaluation and the predictive evaluation. As the defined users in my data inventory are any Disney+ subscribers, the interface should be usable by both novice and expert users. Both of these evaluations specifically target the learnability and discoverability requirements without sacrificing overall functionality. Comparatively, the empirical evaluation places a greater focus on task efficiency since it analyzes how much time the task takes with different designs.

It is technically infeasible to create the exact textual prototype I described, as I am not an employee of Disney+ and could not put my design into their system. Creating my own functioning smart remote app would not be possible given

the time constraints of this class, and would be a more complex and detailed prototype than would be reasonable for the first evaluation phase of the design life cycle. While I could simulate this with a different, existing remote app, I have concerns about how much technical setup and troubleshooting would be required. Additionally, since this evaluation method has to be performed in person there would likely be difficulties obtaining enough participants.

In contrast, the qualitative evaluation is easy to setup due to the low fidelity of the card prototypes. For each session, I merely have to put the card prototypes in the correct order based on which group participants are randomly assigned, and prepare to record and take notes. This evaluation and use of card prototypes also allows me to rapidly create alternate prototypes during sessions based on the received qualitative feedback, further iterating on my design.

Finally, my voice prototype has the potential to be highly interactive as it incorporates Disney's mascot, Mickey Mouse, in its design. Since part of Disney's target audience is children, this feature could improve the user experience for that demographic. Young children could be considered a subset of novice users—although there are certainly some young kids that are expert users and more technologically adept than their parents. Would a young audience appreciate and understand how to use the voice functionality? Would it improve their experience with the application? By analyzing this prototype through a cognitive walkthrough, I hope to answer these questions and get a better sense of how well voice commands would integrate into the system.

5 REFERENCES

- [1] Dean, Brian (Mar. 2021). *Disney+ Subscriber Statistics 2021: How Many People Watch Disney+?* URL: <https://backlinko.com/disney-users>.
- [2] Hobson, Karen, Casano-Antonellis, Jessica, and Schreibfeder, Lisa (Feb. 2021). *Disney+ Lifts Off, Ushering in a New Era of Entertainment from The Walt Disney Company.* URL: <https://dmedmedia.disney.com/news/dtci-disney-plus-launch>.
- [3] Spangler, Todd (Nov. 2019). *Here's what went wrong during Disney Plus launch.* URL: <https://www.chicagotribune.com/entertainment/tv/ct-ent-disney-plus-launch-problems-20191112-n6gmdzfaujcudlmovwsojeuqkq-story.html>.

6 APPENDICES

These appendices contain copies of my prototype designs from Assignment M3 and data inventory/requirements from Assignment M2.

6.1 Card Prototype

In this prototype, the user navigates Disney+ using a traditional television remote. Figure 1 shows a redesigned main menu where the search menu is accessible from a top toolbar that displays upon entering the service and takes two clicks to open. Comparatively, the current interface has the search menu hidden in a sidebar that at minimum takes four clicks to open.

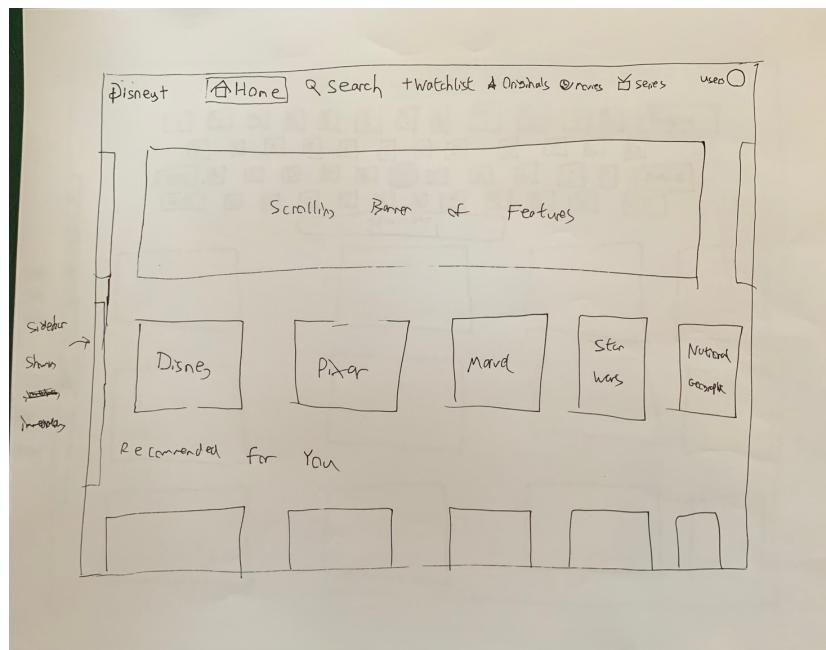


Figure 1—The updated Disney+ TV home screen with a toolbar at the top and the cursor focused on the home button.

Figure 2 shows the redesigned search menu with a QWERTY keyboard at the top, and figure 3 shows the outcome of the user performing a search for the word "big" and navigating to the program *Big Hero 6*. These screens are reached after clicking the search button shown in figure three. Some features that I would want to evaluate either through these renditions or through additional card prototypes would be which keyboard special characters are necessary, and keyboard designs for non-English languages. Alternate designs for these screens are shown in figures 4 and 5. These figures are pictured on the following pages.

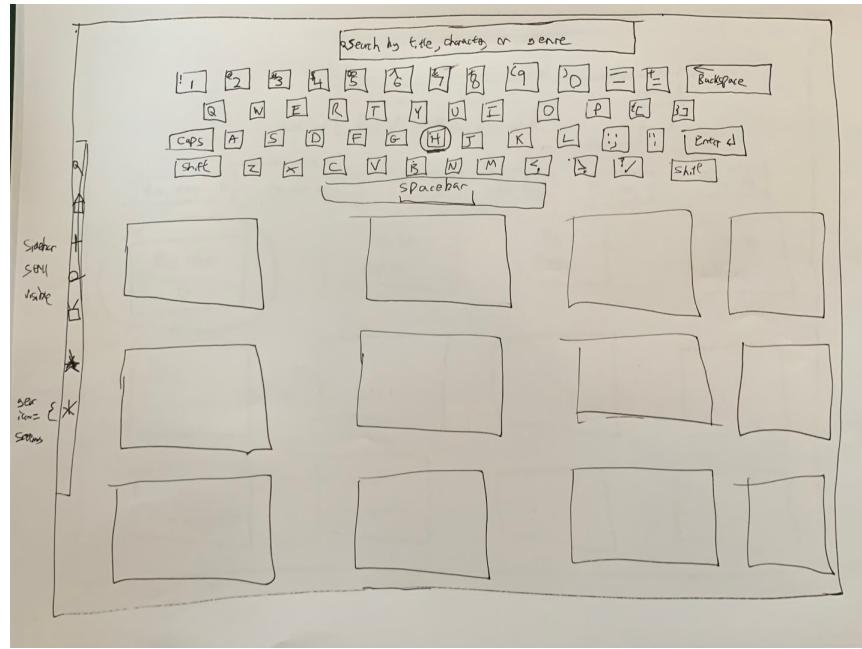


Figure 2—The updated search menu with a QWERTY keyboard.
Cursor focus starts at the center of the keyboard.

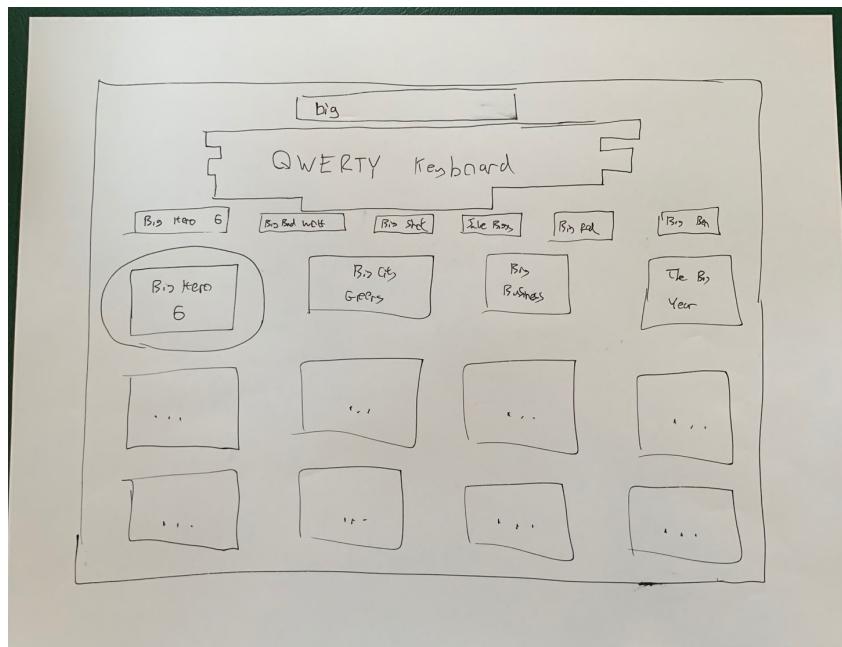


Figure 3—The results of searching the phrase "big" in the interface and moving the cursor to a program.

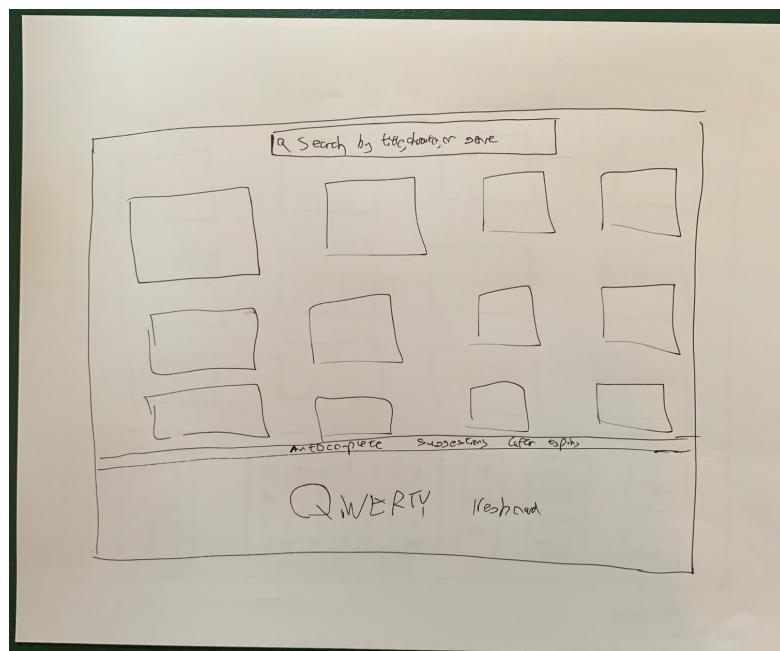


Figure 4—Alternate design with the QWERTY keyboard at the bottom of the screen.

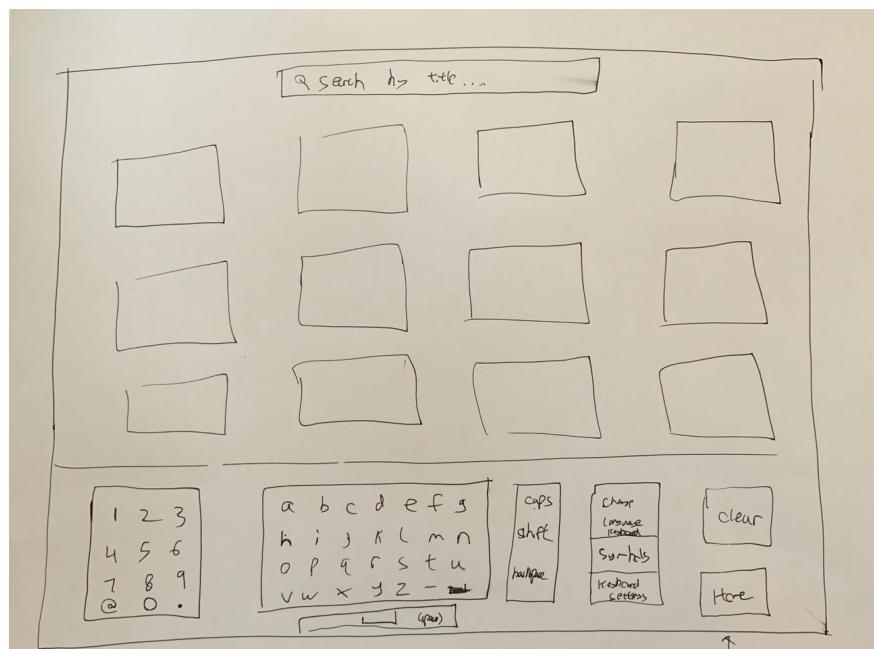


Figure 5—Alternate design with an alphabetical keyboard layout.

6.2 Textual Prototype

Rather than having to use a physical TV remote to navigate through Disney+, users can use the touchscreen of their smartphone (or tablet) as their remote. The layout will feature key navigation buttons, which include directional arrows, a select button, and a back button. Its design will be specific to Disney+, and it will be added to the existing mobile app rather than creating an entirely separate application. In addition to the main navigational buttons, there will also be hotkey-style buttons that allow users to quickly jump to different menus like the search menu. The overall design will be minimalistic to reduce clutter and increase tolerance, so that the user does not have to be as precise with their taps.

The remote app will also feature a secondary touchpad mode where the user can navigate with gestures instead of buttons. The user will be able to swipe directionally on their device to move the cursor on the TV in the corresponding direction (with the option to invert scrolling directions). The user can double tap to select and pinch out (zoom out gesture) to go back a menu.

Within the search menu itself, the remote will switch from whichever mode it is in to a keyboard mode. Rather than having to move the cursor to different letters on the screen, the user will type using their device's native digital keyboard. This will reduce total clicks and time spent typing in a search.

To connect your device to your TV, this design will leverage existing functionality within the Disney+ service. When you login to Disney+ on a TV, you are provided with a code to enter at a login website on either a computer or mobile device. This is so that you can use a full keyboard and enter your password securely. A similar code workflow would be used to sync your device to the TV. This could either be a one-time registration, or be embedded into the startup process for Disney+ when selecting your user profile.

6.3 Wizard of Oz Prototype

For this prototype, I present the script that will be provided to the user along with a list of recognized voice commands and their resulting actions. This list focuses on commands I anticipate users will make and is something I would refine through user testing. As Mickey Mouse is the Disney mascot, I incorporated him as the interactive agent, utilizing some of his catchphrases in the design.

The prompt to the user is as follows: "Your goal is to navigate to the search menu, and search for and play Big Hero 6 using only voice commands. Upon opening Disney+ on your TV, you hear Mickey Mouse say 'Welcome to the updated Disney+ with voice functionality! To see a list of available commands say <Voice Commands>. Enjoy!'"

- "Voice Commands" - Presents the user with a list of available commands for the menu that they are on, including a core list of actions that are available on any screen in the interface.
- "Help" - Brings up a help screen with FAQs, including voice info.
- "Tutorial" - Displays a brief tutorial to the user outlining Disney+ functionality.
- "Search" - Takes the user to the search menu from their current screen.
- "Back" - Takes the user back to the previous menu.
- "Home" - Returns to the home page of Disney+.
- "Big Hero 6" - After using the search command to move to the search menu, saying the name of a program, character, genre, etc. will search for it. Big Hero 6 is presented as an example of a searched program.
- "Clear" - Clear text in search field.
- "Play" - Play the selected program in the search menu.
- "Select" or "Enter" - Clicks on whatever the cursor has selected."
- "Up/Down/Left/Right" - Moves the cursor accordingly.
- "Toolbar/Sidebar" - Opens up the toolbar.
- "Play Big Hero 6" - Moves from the current menu to the video player and starts playing Big Hero 6.
- "Search Big Hero 6" - Moves from the current menu to the search menu with a search for Big Hero 6 already entered.

The interface would recognize commands that are similar to the ones listed. For example, "Look for" would function the same as "Search". For search commands that are not defined, Mickey's voice will appear saying "Oh gosh! I don't know what you mean. Can you try that again?" After two consecutive voice commands that do not match to an action, he will say "Aw gee, try saying 'Voice Commands' for a list of available actions".

6.4 Requirements

First and foremost, the final design for the search interface needs to be functional. As mentioned in the survey responses, the current search results can be

inconsistent or inaccurate. Having results that match user expectation is critical. From the three needfinding methods, I also identified inefficiencies in the current design. For the final design, I want to address this deficiency and will use number of clicks and time to perform the workflow as evaluation metrics.

While addressing these requirements will increase the usability for expert users, my focus is going to be on improving it for novice users. As shown through the observation of a novice user, there are currently large gulfs of execution and evaluation. For this redesign, learnability and discoverability of the search menu are priorities after functionality. My aim is to make it easy for all users to find the search menu, and subsequently use the keyboard to find their program. After using the interface the first time, it should be easy for users to remember how to use the search feature. I can also measure this using time and number of clicks as metrics.

Finally, if all of those requirements are met I will look at increased compatibility across devices. While my focus is going to be on improving the search interface for Disney+ on television devices, if possible I want to improve consistency of the interface across all devices. Having a similar interface on TVs, mobile phones, and computers would make it easier for newer users. This may be a challenge due to hardware limitations, thus I am treating this as "nice to have".

6.5 Data Inventory

6.5.1 Who are the users?

Disney+ as a streaming service has over 100 million subscribers in 59 countries (Dean, 2021). Based on the data from my survey results, combined with Brian Dean's compiled research, roughly one third of Disney+ users have children, and most users are under the age of 40. While my respondents were mostly male, this is representative of the OMSCS sample population rather than the overall population, which is closer to a 50-50 split. Based on the survey results most users of Disney+ subscribe to multiple streaming services, and use Disney+ fewer than 6 hours per week.

6.5.2 Where are the users?

From the survey results, most users utilize multiple devices for streaming. Almost all users use a television or computer, with over half also using mobile devices. Since TVs are not highly portable, users would be streaming on these

devices from their own home on their television. For computers, desktop streaming would occur at home while laptop and mobile streaming could occur at home or while in public. As I did not gather explicit data on the "Where", these conclusions are based on my own experience. To better address this question of the data inventory, I should have included questions on where users are located when they use the service in my survey. I could also conduct interviews with participants to ask them about what physical locations they stream in.

6.5.3 What is the context of the task?

This question was not answered during my needfinding. Similarly to the previous question, I can speculate about user context based on my own experience as an extension of participant observation. While most of the time when using Disney+ I will exclusively be watching the program, there are also times where I am answering email at the same time. I may put a program on in the background while I am cooking, or while I have friends over and we are playing a board game. Thus, there are times where other tasks or interfaces compete for user attention. To answer this question, I could have included questions about participant context in my survey, such as "When using Disney+, what other tasks do you perform at the same time?". To answer this, I would either do this additional surveying, or conduct naturalistic observation to observe user context.

6.5.4 What are their goals?

The goal for a user when using Disney+ is to watch a program. This question was answered through participant observation and survey needfinding. As a user of Disney+, when I login to the service, I am almost always planning to stream a specific program. Rarely, I will login without any specific plan of what to watch and will select something randomly from the service, albeit still with the goal of streaming some program. Based on the survey results, users of Disney+ have a variety of TV shows or movies that they watch, all with the same common goal when using Disney+ of watching a program.

6.5.5 What do they need?

To use Disney+, users need a physical device, such as a television, a computer, a mobile phone, or a tablet. If using a television, users also need a remote. The needed physical objects for this question were answered via participant observation. From the observation of a novice user, users need to be able to easily

discover how to navigate the platform, and they need feedback to help when having difficulties. From the survey results, users need consistent searches, age filters that work correctly, and efficient search tools. The results of the evaluation of other interfaces will help inform prototypes that address these user needs.

6.5.6 What are their tasks?

As the goal for a user of Disney+ is to watch a program, the main task is to load and start playing a selected program. In order to play a program a user has to find a program they want to watch, which is typically done via the search menu. Physically, a user is using either their hand or voice to enter in a search query. Depending on the device, they are interacting with either a remote, a phone screen, or a computer keyboard. Cognitively, the user is thinking about what program to watch, what letters/words to type or say, and if the program appears on their screen. Socially they may also be talking to other colleagues that they are watching a program with. This question was primarily answered by participant observation, but I could also conduct additional observation of users or interviews to see other tasks that users perform when using the service, such as using Group Watch for social interaction.

6.5.7 What are their subtasks?

The subtasks for Disney+ were answered through participant observation. In order to find and load a program to watch it (the task), the user has to perform subtasks: login to Disney+, select their user, and navigate to the search menu. From observing a novice user, I found that the subtask of navigating to the search menu can be difficult due to the sidebar's initial hidden state. Once a user is in the search menu, they find their program and then perform the subtask of hitting the play button to start the program. While this question was answered, in my needfinding I focused on a linear workflow going directly from login to searching for and playing a program. Some users may use other features like the watchlist for streaming, so I would want to perform additional observation or interviews to see other workflows users perform and if any workarounds exist.