Kyle Schwartz

IS420

Dr. Calongne

4/29/2024

IS420 Final HCI Project: Splash Dash Usability Evaluation Report

Name

Project Name: Splash Dash

Developer Name: Kyle Schwartz

Avatar Name: Tippy Tractor

Project Type and Introduction

My project is a 2D horizontal side scrolling runner game made with the Unity game engine. Although testing all the features in the game would require more than three tasks, I am planning on analyzing three of the most essential tasks to playing the game in my usability evaluation. I will be able to use this data to improve upon the game's existing user experience by adjusting the interface and by modifying how the game progresses as the user interacts with it.

**Project description - what does it do?**

The game focuses on helping to fight pollution by keeping our oceans clean. The user can play as multiple sea creatures that are controlled in an up and down motion. The game uses a time-based mechanic that allows the player to swim longer by avoiding obstacles and collecting time boosters. The user's score is increased by collecting trash. Multiple menus allow players to choose between various playable characters and levels. Overlays appear when the game ends or if it is paused. While the game is playing, all player characters are animated. The user also hears confirming sounds if they run into obstacles, collect things, or get a time boost.

Tools used to develop and test the project.

To develop the game, I used the Unity Game Engine Editor. The game uses thirty separate C# scripts to run, and due to it being a Unity based game, it runs on the Unity game engine. The C# scripts were written in Microsoft's Visual Studio with the Unity extension. All the artwork used in the game went through a developmental process and required the use of several different tools. Most of the characters were hand drawn with a pencil and paper, and then scanned onto the computer using a Canon scanner. They were then edited using Microsoft photo editor, and they were further edited and colored in Microsoft's Paint 3D. All sprite animation were hand

drawn in Paint 3D. The UI elements such as the buttons were also developed in Paint 3D. The background music was downloaded from Pixabay.com, but the other sounds used in the game were recorded and edited using the Audacity sound editor. In short, the entire interface was put together in Unity, but each individual element was usually created in Paint 3D or Audacity. To test the project remotely, I used Zoom with the chat feature. To test the project in person I used papers with the test task printed on them.

Interface images (what we see on screen)
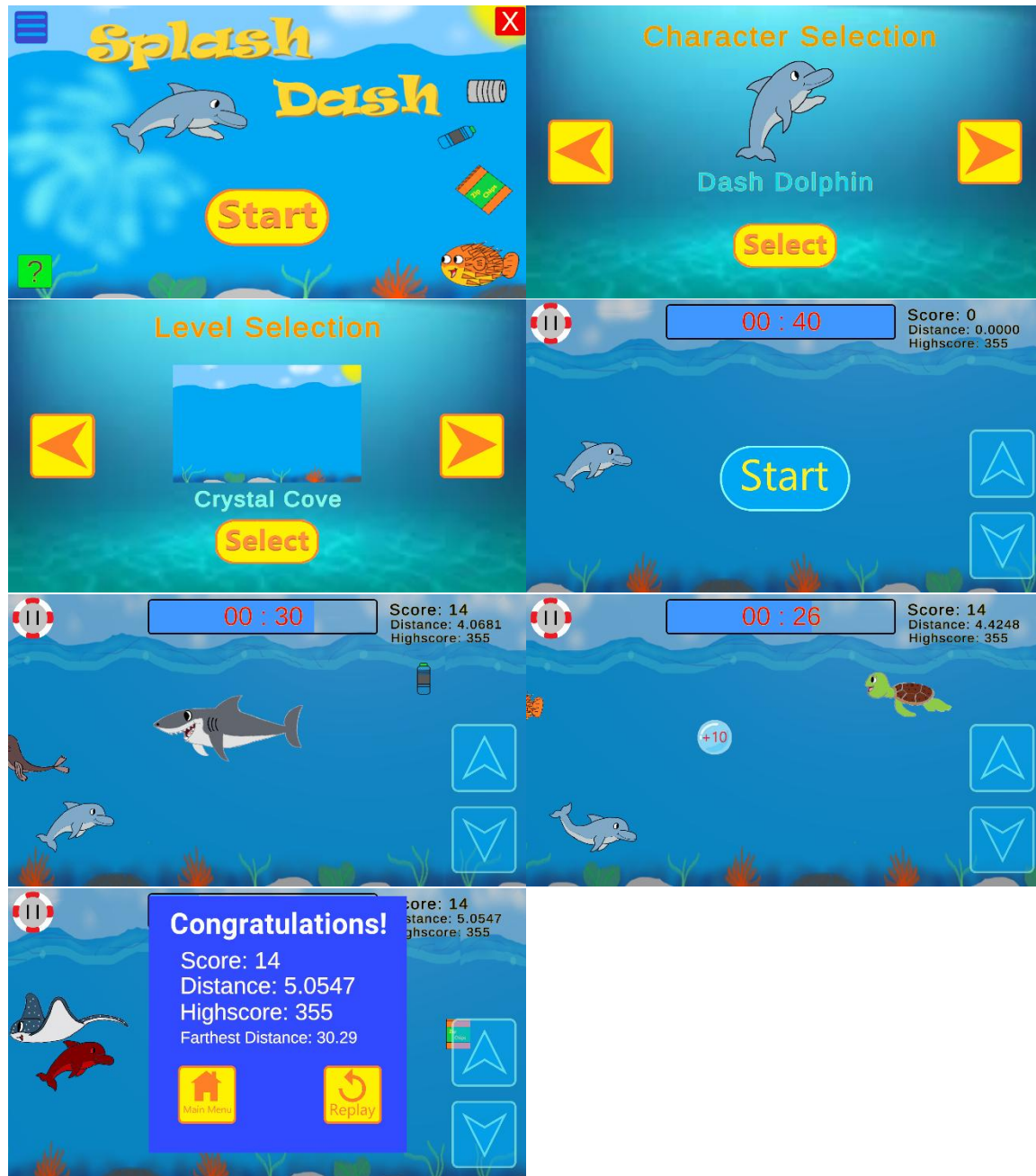
Test Tasks / Test Plan

Introduction:

Hello, thank you for taking the time to participate in this usability test. Today you will be testing a 2D runner game. I will be giving you three tasks. I will give you a written copy of each task and I will read the task to you. Please wait until I say, "you may begin" before starting a task so I can write down the start time. Please talk aloud and tell me what you see, what you are doing, and how the experience is going as you complete each task.

Test Tasks:

> Task 1: Start the game, select a character and level, and proceed to the game scene

> Task 2: Start the game and play it as long as possible

> Task 3: Take note of your score and return to the main menu

Usability Criteria and Goals

For this project, the usability criteria were automatically set "Easy to Learn and Use" and "Subjunctive Satisfaction" since they are generally the easiest criteria to test. To test this criterion, I used two measurements. The time to perform a task for easy to learn and use, and a subjunctive satisfaction survey to test the subjunctive satisfaction of the project. The goals for the tasks and survey questions are outlined below.

Usability Test Goals:

Task 1: Start the game, select a character and level, and proceed to the game scene

> Good (average time for someone who has never played the game) = 25 seconds.

> Better = 15 seconds average.

> Best = 10 seconds average.

Task 2: Start the game and play it as long as possible

> Good (average time for someone who has never played the game) = 2 minutes average.

Better = 4 minutes average.

Best = 6 minutes average.

Task 3: Take note of your score and return to the main menu

Good = 20 seconds average.

Better = 15 seconds average.

Best = 10 seconds average.

QUIS Survey Questions Goals:

QUIS Question 1: Did you find the software to be engaging? On a scale of 0 to 9, with 0 being dull and 9 being stimulating, how would you rate the software?

Good (starting goal) = 7.5

Better = 8

Best = 8.7

QUIS Question 2: Can tasks be performed in a straight-forward manner? On a scale of 0 to 9, with 0 being difficult and 9 being easy, how would you rate the software?

Good (starting goal) = 7.4

Better = 8.2

Best = 8.6

QUIS Question 3: Based on a scale of 0 to 9, with 0 being terrible and 9 being wonderful, what is your overall reaction to the software?

Good (considering this is a demo) = 7.3

Better = 7.9

Best = 8.8

Test Results

Test Data: Test Task Times and QUIS Survey Responses

Participant #1			
Date: 4/11/2024			
Time: 9:12:55PM – 9:20:53PM			
Participant Description: This participant is a 66-year-old female gamer who plays mostly 3D games as opposed to 2D casual games. She interacted with the interface once during a prior evaluation but did not remember all of the technical details of the application.			
Device Used: Windows 10 desktop computer on Simmer.io game hosting server in a chrome browser version 123.0.6312.122.			
Test Task Times		**QUIS Survey Responses**	
Task #	Times	QUIS Question #	Scores
1	Start Time: 9:12:55 Stop Time: 9:13:44 Total Time: 49 seconds	1	9
2	Start Time: 9:15:38 Stop Time: 9:16:46 Total Time: 1 minute 8 seconds	2	8
3	Start Time: 9:20:22 Stop Time: 9:20:53 Total Time: 31 seconds	3	8
Additional Comments: The participant commented on how easy the up and down movement was to use. The participant also commented on how she wasn't completely sure what objects she should run into right away. Finally, the participant commented on how, with a little practice, the game would become more engaging.			

Participant #2			
Date: 4/25/2024			
Time: 11:35:16AM – 11:41:42AM			
Participant Description: This participant is a 54 year old female that primarily uses the technology that this game is played on for web browsing and online shopping. Device Used: Tested in a Google web browser on an iPad OS version 17.			
Test Task Times		**QUIS Survey Responses**	
Task #	Times	QUIS Question #	Scores
1	Start Time: 1:35:16 Stop Time: 1:35:40 Total Time: 24 seconds	1	9
2	Start Time: 1:36:45 Stop Time: 1:40:26 Total Time: 3 minutes 41 seconds	2	9
3	Start Time: 1:41:35 Stop Time: 1:41:42 Total Time: 7 seconds	3	9
Additional Comments: This tester saw the application several times before during development, so this has to be factored in to the data.			

Participant #3 (Not Counted Toward Averages See Comments Below)
Date: 4/25/2024
Time: 11:13:43AM – 11:17:52AM

Participant Description: This tester is a 86 year old male who never uses computers and never plays video games of any kind.

Device Used: Acer Chromebook 315 at Simmer.io server using a Google Chrome browser.

Test Task Times		QUIS Survey Responses	
Task #	Times	QUIS Question #	Scores
1	Start Time: 11:13:43 Stop Time: 11:14:15 Total Time: 32 seconds	1	2
2	Start Time: 11:15:20 Stop Time: 11:16:20 Total Time: 1 minute	2	5
3	Start Time: 11:17:00 Stop Time: 11:17:52 Total Time: 52 seconds	3	5

Additional Comments: This user was more familiar with mobile devices, but the only device I had available to test on was a Chromebook with no touch screen. I believe that the interfaces inability to support this user was mainly caused by not taking into account the user's abilities and preferences to minimize the effects of the unfamiliar technology. I had to help the user with the laptop, and the time for the interface was okay, but the QUIS questions were based on the technology used as opposed to the interface itself.

   For this reason, I believe it would be beneficial to omit these results from the final analysis as they are more centered on the usability of a laptop as opposed to the usability of the interface that is being tested. I still included this data, however, as it did provide real world insight into someone who may be attempting to use the interface on an unfamiliar device.

Participant #4			
Date: 4/25/2024			
Time:			
Participant Description: This user is a 82 year old female that has limited technical knowledge mainly geared toward shopping online and surfing the web online. Device Used: Acer Chromebook 315 at Simmer.io server using Google Chrome web browser.			
Test Task Times		**QUIS Survey Responses**	
Task #	Times	QUIS Question #	Scores
1	Start Time: 10:43:24 Stop Time: 10:43:33 Total Time: 9 seconds	1	9
2	Start Time: 10:44:07 Stop Time: 10:44:33 Total Time: 26 seconds	2	9
3	Start Time: 10:45:10 Stop Time: 10:45:13 Total Time: 3 seconds	3	9
Additional Comments: This user was right handed but had an injured right hand so using her left hand to complete the tasks did affect the times slightly.			

Participant #5
Date: 4/25/2024
Time: 7:04:18PM – 7:06:22PM

Participant Description: This user is a 54 year old male that uses and repairs technology and devices daily for work, but the user never plays any video games.			
Device Used: Tested in a Google web browser on an iPad OS version 17.			
Test Task Times		**QUIS Survey Responses**	
Task #	Times	QUIS Question #	Scores
1	Start Time: 7:04:18 Stop Time: 7:04:40 Total Time: 22 seconds	1	8
2	Start Time: 7:05:14 Stop Time: 7:05:50 Total Time: 36 seconds	2	7
3	Start Time: 7:06:20 Stop Time: 7:06:22 Total Time: 2 seconds	3	9
Additional Comments:			

Average Task Times

Test Task Averages	
Task #	Average Time
1	26 seconds
2	87.75 seconds
3	10.75 seconds

QUIS Survey Responses

QUIS Question 1: Did you find the software to be engaging? On a scale of 0 to 9, with 0 being dull and 9 being stimulating, how would you rate the software?

> Participant 1: 9
>
> Participant 2: 9
>
> ~~Participant 3: 2~~
>
> Participant 4: 9
>
> Participant 5: 8

QUIS Question 2: Can tasks be performed in a straight-forward manner? On a scale of 0 to 9, with 0 being difficult and 9 being easy, how would you rate the software?

> Participant 1: 8
>
> Participant 2: 9
>
> ~~Participant 3: 5~~
>
> Participant 4: 9
>
> Participant 5: 7

QUIS Question 3: Based on a scale of 0 to 9, with 0 being terrible and 9 being wonderful, what is your overall reaction to the software?

> Participant 1: 8
>
> Participant 2: 9
>
> ~~Participant 3: 5~~
>
> Participant 4: 9
>
> Participant 5: 9

Average QUIS Responses

QUIS Question Averages	
Question #	Average Score
1	8.75
2	8.25
3	8.75

Average Scores and Goal Comparison

Test Task Averages		Test Task Goals	
Task #	Average Time	Goal Times	Comparison
1	26 seconds	Good = 25 seconds Better = 15 seconds Best = 10 seconds	This time did not fall into my goal range. It was 1 second over my good time goal.
2	1 minute 46.25 seconds	Good = 2 minutes Better = 4 minutes Best = 6 minutes	This time was under my goal range. The difficulty of the game affected this task, and it was more difficult than my original prediction. This goal was 13.75 seconds under my good time goal.
3	10.75 seconds	Good = 20 seconds Better = 15 seconds Best = 10 seconds	This goal fell into the better category. Most testers completed this task quickly with the only delay being to read their score.
QUIS Question Averages		QUIS Question Goals	
Question #	Average Score	Goal Scores	Comparison
1	8.75	Good = 7.5 Better = 8 Best = 8.7	This score fell slightly above my best goal range. Most users found it to be more engaging than I had predicted.
2	8.25	Good = 7.4	This score fell into the better score

		Better = 8.2	range based on my goals.
		Best = 8.6	
3	8.75	Good = 7.3	This score fell into the better range of my goals. Almost reaching the best range.
		Better = 7.9	
		Best = 8.8	

Your analysis of the future work needed to finish or maintain the project.

Overall, testing went quite well. All participants were happy to test the application, and the process was quick and simple. The data I collected during these tests will help me to better fine tune my application based on user feedback.

I believe that the test introduction was clear and easy for users to understand. They generally had a good idea of what was going on, but it was sometimes difficult to make the users wait until I say, "you may begin". Most users attempted to start as soon as they receive a task. All the testers understood what to do when given a test task, and the test tasks did reveal both the strengths and weaknesses of my interface.

Something I might change if I were to continue this project is the goal on my first and third tasks. I originally based these goals on how fast the user could maneuver through the menus, but most participants stopped a little longer to look at the available characters and levels as well as their scores. I believe that someone stopping and taking note of these menus and numbers shows engagement with the interface, since they take their time in selecting a character and level and since they are interested in the score they received. I would increase my time for these goals based on this data since the software better supports the user not only if they can navigate through the menus, but also if they engage with them.

A final thing I might change would be the demographics that I test the project with. Although it supported the interactions of nearly any age group, the application itself was primarily targeted at younger audiences. This is why I believe that in a more formal testing situation, finding more younger participants under the age of 15 would be beneficial as it would provide me with a wide range of data based on both the participants technical skills and their age.

The manual recording of time was difficult especially when trying to record other data and inform the testers of their next task. This time recording process also creates lag that causes a slight distortion of data compared to the actual time and the recorded time. In the workplace, I would generally have a formal automated system that made recording times much simpler. This would allow me to better compare my test data to my goals.

Another tricky part with the testing was participant #3. I wasn't sure to include this data or not since it did provide insight into this user's experience, but this tester mainly based their

QUIS responses on the issue they had with the device they were testing on, not the interface itself. This is why I decided to omit participant #3 from the final analysis. I thought it was important to see how the interface performed minus the technical issue to help limit the outside variables as much as possible. The participant #3 data gave a "real-world" insight as some users may be using the interface on devices that they are also unfamiliar with, but it was not particularly useful when trying to determine how to improve the interface design.

Your code or personal notes on its behavior and your reflections.

Overall, the program performed better than I expected. Most users found it easy to use and engaging. I think being a game, most people answered regarding rather or not they liked the game which is to be expected. Some people enjoy 2D runner games or video games in general and some do not. I believe that I could have factored all 5 testers data into the results had I considered each testers need a little better. The tester should have had the option to test on both a mobile device and a laptop computer depending on the type of technology they are comfortable with considering the game would be deployed on both device types.

The game itself could have minor adjustments, but the interface performed generally well. The main issue that people commented on was the difficulty to distinguish between things to run into and things to avoid. Developing a solution to make bad obstacles and good obstacles more distinguishable would help the learnability of the game, but this is one aspect of the game that people do learn over time after hitting the obstacles. Below is an example of three obstacles that the user must determine to hit or to avoid. Roughly half the users had trouble determining what to hit or to avoid starting out, but they quickly learned after testing the obstacles. Since the user did not play the game a second time during the test, there was no data on if they remembered what obstacles to hit or to avoid. Asking the same testers back to play the game again would be one way to see if the users remember how to play without modifying the interface itself. Adding another usability criteria test such as retainability or memorability would allow me to test this aspect of the interface.

Overall, I am pleased with the project and how testing went because I learned about not only how to develop an interface, but also how to test it to ensure that it is meeting the needs of its target users as development is progressing through the use of formative evaluation.