STARS

University of Central Florida
STARS

Electronic Theses and Dissertations, 2020-

2020

The Effects of Gesture Presentation in Video Games

Jack Oakley
University of Central Florida

Part of the Computer Sciences Commons

Find similar works at: https://stars.library.ucf.edu/etd2020

University of Central Florida Libraries http://library.ucf.edu

This Masters Thesis (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2020- by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation

Oakley, Jack, "The Effects of Gesture Presentation in Video Games" (2020). *Electronic Theses and Dissertations, 2020-*. 389.
https://stars.library.ucf.edu/etd2020/389

University of
Central Florida

STARS
Showcase of Text, Archives, Research & Scholarship

THE EFFECTS OF GESTURE PRESENTATION IN VIDEO GAMES

by

JACK P. OAKLEY
B.A. Collins College, 2008

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Computer Science
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Fall Term
2020

ABSTRACT

As everyday and commonplace technology continues to move toward touch devices and virtual reality devices, more and more video games are using gestures as forms of gameplay. While there is much research focused on gestures as user interface navigation methods, we wanted to look into how gestures affect gameplay when used as a gameplay mechanic. In particular, we set out to determine how different ways of presenting gestures might affect the game's difficulty and flow. We designed two versions of a zombie game where the zombies are killed by drawing gestures. The first version of the game is a touchscreen-based game where the gestures are drawn in 2D space on the screen while the second version utilizes 3D space to draw gestures in virtual reality. We performed two studies comparing gestures presented as symbols and names, one study using the two-dimensional touchscreen game and one using the VR version. We found that presenting gestures by name increases the game's difficulty in the 2D version of the game. Flow was unchanged by gesture presentation but flow increased with difficulty in our 2D game. We were unable to affirm these same results with any significance in the VR version of the game. We discuss the implications of our results and provide insights to help game designers make more informed decisions about gesture implementations as gameplay elements in video games.

I dedicate this thesis to my loving wife, without her support and recommendation, I would have never gone back to school nor accomplished as much as I was able with her support. I also dedicate this thesis to my dad and stepmom for without our arrangement, I would have struggled more with time and money throughout this process.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Video games are very popular worldwide with 78.61 billion US dollars worth sold in 2017 and expectations exceed 90 billion by year 2020 [39]. Gestures have been utilized in games since the Nintendo Wii first came out. The Wii Remote used an accelerometer and infrared sensors to determine its angle and location [7, 22]. Shortly after that, Microsoft created the Kinect which uses a multi-camera system to recognize movement and gestures as seen in games such as *Dance Central* [28]. These however are systems which are large and not easily transportable. As gesture recognition systems are becoming more ubiquitous and robust, they are also becoming lighter, simpler, and easier to deploy [37]. This makes it easier to include gestures in video games on mobile platforms. Nintendo's DS series used touch-based gestures in games and as more mobile devices started using touch-based interfaces [12, 23], touch-based gestures began to make a larger appearance in mobile games [7]. There is a multitude of research that looks into gestures as a menu navigation method within applications, but we are more curious about the research and use of gestures as a gameplay mechanic. Are game designers implementing gestures into games by following established standards for menu navigation? Do gestures or their presentation to a player when used as a gameplay mechanic affect the gameplay of the game? We set out to determine how gesture presentation might affect gameplay when gestures are used as a gameplay mechanic rather than a menu navigation technique.

We thought of two major areas that gesture presentation may affect a game, its level of fun and its difficulty. Fun is not a simple concept to narrow down, and why is it important for video games? The goal of any game, including board games, card games, pencil and paper role-playing games, and video/computer games is to be fun, engaging, and/or immersive. Engagement is the act of being drawn to an activity and becoming hooked in it [24]. Immersion is when someone is so engaged in an activity that they do not want to stop that activity and oftentimes forget about
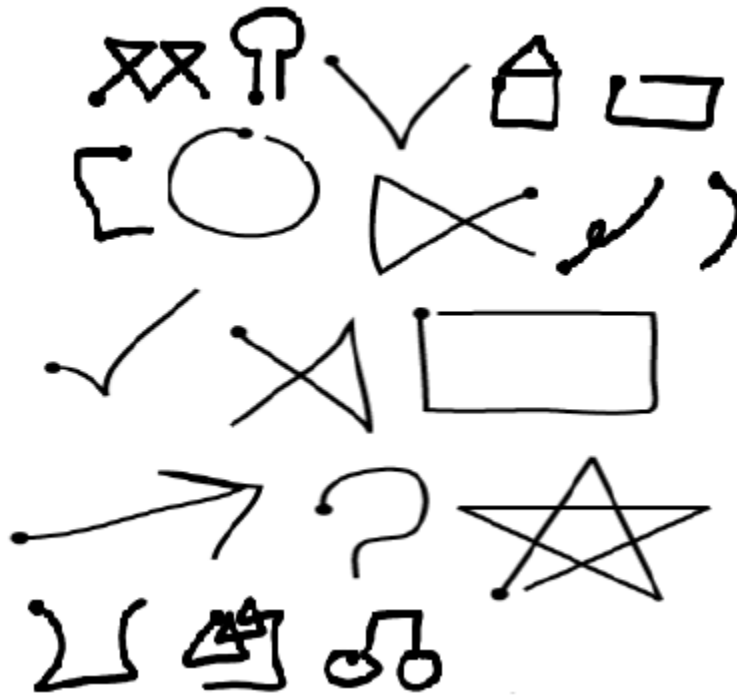
Figure 1.1: An example of gestures presented in symbol form.

the world around them [24]. Flow encompasses all of this as it is a multi-faceted dynamic that encompasses fun, engagement, and immersion. As defined by Gregory [11], flow is the experience of becoming engaged and fully immersed in an activity. In other words it is the process of being drawn to the activity, then becoming hooked in the activity and not wanting to stop the activity, oftentimes forgetting about the outside world. According to Bederson [2], flow is conceptually difficult to narrow down but it is an important aspect of human experiences and anything that can be done to improve the experience of interaction is a worthwhile investment of time.

When gestures are implemented in a game as a gameplay element, they are typically presented to the player using a visual method. This visual method can present the gestures in various forms.

Together X Arrow
House Rectangle Car Check
Delete
Tree V Question Triangles_Chain
Right_Paren Star
Circle Unistroke_X
Left_Square_Brckt
Reversed_Pi

Figure 1.2: An example of gestures presented textually as names.

One form the gesture presentation might take is the form of symbols that represent the motion of the gesture as seen in Figure 1.1. This form of gesture presentation does not take much cognitive load, or mental stress, to replicate, as the task imparts recognition and replication of what is shown [8, 10, 33]. Another form might be to present the gesture textually as its name, which takes more cognitive load because it instills the added task of recalling the gesture from the textual context that is shown. An example of how several gestures might be presented using their name as a textual cue can be seen in Figure 1.2. With a rising use of gestures in games, it makes one wonder if the gesture presentation might have an effect on the flow of a game. Does it make the game more fun? Possibly less fun? Some gestures are complex whereas others are simple. From this, one begins to ponder if there is some sort of balance between cognitive load and flow. At what point does it become too much load that it reduces flow? After extensive research, we determined that cognitive

load is the mental component of difficulty, as opposed to the physical component [34]. This brings us to our research questions.

- **RQ1:** How does gesture presentation style affect game difficulty?

- **RQ2:** How does gesture presentation style influence flow?

- **RQ3:** How does difficulty from cognitive load affect flow?

In order to test these questions, we set out to compare two levels of cognitive load based difficulty interventions in two separate video game scenarios. These two scenarios were compared to each other in user studies that incorporated two separate gestural interventions first via a touch-based game, then a second study using a virtual reality version (VR) of the game. In both games, the player tries to survive a zombie apocalypse by avoiding and killing zombies. The method of killing zombies is to draw specific gestures as presented above the zombies' heads. Both user studies had the participants try to survive as long as possible while killing as many zombies as possible, while being presented with the two interventions separately. One intervention presented the gestures as symbols. The other intervention presented them by their associated name. Each one of these memory interventions imposed a different level of cognitive load via gestural presentation style.

In the touch-based user study, the player played the game using a touchscreen device from a bird's eye view of their character. The player would utilize their finger to draw on the screen. They would draw the two dimensional gestures that they see or the gesture associated with the name they see above the zombies' heads. If the correct gesture was drawn on top of a zombie or the zombie walked into the gesture, it would be killed.

In the virtual reality study, the player would use a controller to draw three-dimensional gestures in the air in front of them. The gesture would then be projected to the closest object in the direction

4

the player is looking, as denoted by a targeting reticle that jumps to that object. As with the touchscreen version, a zombie that touches one of the gestures that matches what is above its head is killed.

The participants of the studies were questioned to determine flow within each of the interventions. They were also questioned about how much difficulty the task imposed upon them.

1.1   Overview

To design our study and formulate hypotheses from our research questions, we first reviewed the related literature. We provide a chapter sectioned by related material. Our first section shows that we researched gestures by examining how gestures have been designed, utilized, interacted with, displayed, and in what ways they have been researched in the past and provided a brief overview. We then provide an overview on our review of literature related to difficulty, followed by our overview of flow theory with a focus on how it relates to games.

Our next chapter goes into the details of our first study. In this study, we designed a two-dimensional touchscreen game. We detail the formation of our hypotheses, experimental design, the design of the game, our technical contributions, the technical specifications of the devices we used, then our study procedure. We then report our findings and discuss their interpretation and implications for designers.

Our next chapter goes into the details of our second study. In this study, we converted the touchscreen game into a full three-dimensional virtual reality game. We provide details on the game's conversion from 2D to 3D. We then go into detail about the technical specifications of the second study, the experimental design, the formation of our hypotheses, our user study procedure, our results, and a discussion of the findings.

Following that chapter, we provide some statistical comparisons from the first study to the second. We then provide a chapter with our conclusion and a brief overview of possible directions to take future work.

CHAPTER 2: LITERATURE REVIEW

We provide a review of relevant literature at the intersection of gestures and interaction techniques, difficulty, and flow. We section the literature review into categories by topic for ease of reference and readability. Material that spans across more than one of these categories are categorized into the category that it most significantly represents.

2.1 Gestures and HCI Techniques

The idea for this study came from helping with some development for a study by Taranta et al. [35] and helping run participants in a few other related, but as of yet, unpublished studies. From this, we further modified the game used within these studies and turned it into a zombie game with the capability of presenting gestures in two ways. Because of this, we already had a gesture recognizer implemented into the game, Jackknife, developed by Taranta et al. [36]. This was a good gesture recognizer for our purpose as it allows for both 2D and 3D gestures as well as is used for the training of the gestures which is done in real time instead of scanning documents of pre-drawn gestures. There are a great many gesture recognizers out there, and we looked into a few of them to be sure that the Jackknife recognizer truly fit our needs for this study.

GRANDMA is one of the gesture recognizers that we looked into, created by Rubine [29]. He created GRANDMA as a toolkit to rapidly add gestures to direct manipulation interfaces as at that time gestures were not common in interfaces and had not been researched much This gesture recognizer would work for the first study that we had planned to do, but would not work for the second study. Because of this, the Jackknife gesture recognizer was a better fit for our needs.

Wobbrock et al. [43] found that gesture recognizers were being created and implemented by ex-

perts in the pattern matching and artificial intelligence fields. This means the human-computer interaction field was not benefitting from this natural way of using technology. To remedy this, they created the $1 recognizer to be cheap, robust, and easy to implement even for novice programmers. The $1 recognizer recognizes unistroke gestures and maintains high recognition accuracy with very few sample templates loaded. Anthony and Wobbrock took the $1 recognizer and extended it to recognize multistroke gestures, calling their new creation the $N recognizer [1]. The $1 recognizer is comparable to Jackknife and would work for our needs for the first study, but would require some manipulation for the second study. The $N recognizer utilizes multistroke gestures, but we had planned to only do unistroke gestures. Therefore neither of these recognizers were ideal for our two studies.

In an attempt to make a more simplified gesture recognizer, Taranta et al. [37] created the Penny Pincher recognizer. This gesture recognizer, unlike its other lightweight counterparts, works off a single gesture template and has comparable if not better accuracy than the other lightweight gesture recognizers. They accomplished this by using only addition and multiplication in the template matching process rather than having expensive calculations for rotation, translation, and scaling. Because of this, they were able to make a gesture recognizer that is able to recognize gestures at five times the speed of its closest competitor. They took it a step further and tested the recognizer in video game setting where gesture input is often rushed and therefore input is less accurate. Their results showed that the Penny Pincher recognizer is on par with and often outperforms similar gesture recognizers. Although the Penny Pincher would likely have worked for our studies, it was more convenient to use the real-time recording of Jackknife. This especially made more sense when recording the virtual reality gestures in three dimensional space.

Most software systems lack formally expressed user-oriented properties during their design. This is a problem as defined by Duke and Harrison [9]. From their research, they have found that user-interfaces are often left to the final steps as an implementation detail rather than designed and

worked into the early steps of the design of the software. Because of this, artifacts are often created that function perfectly to specification, but are user-unfriendly. They performed a study looking into various perceptions of a user interface and its functionality to complete the task for which it was intended. They concluded that the process of design demands that user-oriented properties be considered early enough to be part of the specification and that some aspect of the presentation should be modeled as a prerequisite. With their research, we know it is important to know how presentation of a UI element affects the user's perception so that we can implement it the proper way in the design part of the specification.

Looking further into gestures, Wobbrock et al. [42] found that gestures created by system designers are not necessarily the best gestures to use for features that are controlled by such gestures. With the gestures not matching that of user behavior, they looked into methods for designing gestures to match what users expect. In their study, they presented an effect that a gesture should have to their participants and asked the participant to perform the cause that would have that effect. In the end, they found that users typically prefer one handed gestures over two handed gestures and they do not typically care how many fingers are used. They also found that users' thoughts are influenced greatly by computer desktop idioms.

Morris et al. [25] performed research into user preference when it comes to gestures for particular tasks. They ran a study in which they had a few different gesture sets created by varying amounts of people, some by researchers only and others by an end-user elicitation method. They then ran their study by having participants sit down at a Microsoft Surface device display and were shown the title of a particular task. The title was then read audibly and a brief definition was given audibly. The participant was then shown each gesture for that task one at a time and told to perform that gesture to ensure they give equal evaluation to each possible gesture. The participants then answered two Likert scale questions, "Is the gesture a good fit for the task?" and "Is the gesture easy to perform?" Their results determined that users prefer gesture sets that are created

by larger groups of people. They also concluded that user-defined gestures are more preferred to researcher-defined gestures.

Along those same lines, Nacenta et al. [26] performed a study into the memorability of gestures. By using gesture sets defined by researchers, user-defined gestures sets, and random gesture sets, they had their participants over the course of 4 days learn gestures, perform them, then on subsequent days repeat what they have learned. Their results showed that user-defined gesture sets are easier to remember than researcher-defined gesture sets. The user-defined gesture sets also took less time to learn than the researcher-defined sets. They concluded that the majority of differences between gesture sets are caused by association errors and not errors in the forms of the gestures.

With gestures often representing certain tasks in which they are used to perform, it is essential to maximize their guessability so that more users will naturally understand how to perform those gestures with little to no training. Wobbrock et al. [41] looked into this dilemma and designed a method for maximizing the guessability of gesture sets. Their method involves obtaining guesses from participants. Their method states not to show participants anything like text or previous gestures as this can bias the results. With that in mind, they ask participants to guess the gestures associated with particular referents without showing them any symbolic form of the referent. After performing this collection of symbols, the symbols are analyzed, categorized by similarity, and counted. Whichever symbol has the highest count per referent is the symbol that gets assigned to that referent. If there are symbol conflicts between referents, whichever referent has the majority of that symbol gets it assigned to them.

Benz [3] performed a 3x3x3 between-subjects study with three different video games of different genres and compared three different modalities of interaction with the game interfaces. The three chosen game genres were puzzle, strategy, and dueling. In these three games, he compared a gesture only interface with a non-gesture and a mixed interface in an attempt to determine which

has the best flow state for game interaction. Each participant played the game and filled out Likert questions in response. He then calculated the mean values for each interaction type and game genre. Though the mean values of each interaction type varied per game genre, overall Benz determined that gestures offered the highest flow state for players. From these findings, Benz concludes that gestures are an appropriate interaction technique for games.

With the numerous gestures that exist already, how does one choose which gestures are appropriate for a gameplay element? How difficult are the gestures and how difficult should they be? Cao and Zhai [5] designed a time-based model of determining unistroke gesture complexity using a combination of curves, line segments, and corners. They call the model CLC. They began their experiment with five phases, three to determine elemental models of arc, straight line, and corner, the last two to validate the models. They found that gesture production time is significantly impacted by length and orientation of straight lines, both sweep angle and radius for arcs, and corners of all degrees, with some angles being in the positive mean, some in the negative mean, and some hovering around zero. They also found that people tend to draw arcs instead of corners on well-practiced gestures. Because of this, gestures that are well-known tend to be completed quicker than those that are foreign. They hope their model can be a foundation point for designing and evaluating gestures for gesture-based interfaces. We used this model to help choose gestures with a variety of complexities for both versions of our game.

Mayer et al. [23] performed a study into how users perceive and perform gestures with which they are presented. They used 3 straight-line gestures with varying numbers of segments and presented them to users at various angles of rotation. They then asked the users to replicate the gestures. They found that users have a tendency to align the straight lines with the primary axes of the device of which is being used. They conclude that design implementations of gestures should include flick gestures and designing systems with improved recognition of compound gestures.

Tu et al. [38] compared gestures drawn by a pen with those drawn by finger. They found that gestures drawn by pen are often smaller and drawn slower than those drawn by the finger. The shape between the two gestures differs by aperture of closed gestures, the deviation of intersecting points, and the distance in the corner shape. Gestures drawn by fingers were similar to the ones drawn by pen when it came to proportional shape distance, indicative angle difference, axial symmetry, and articulation time. They conclude that it is vital when designing finger-based gestures to keep in mind the similarities and differences they have with pen-based gestures and to avoid using features of strokes that do not perform as well as the pen. The Jackknife recognizer used in our studies is able to pickup on curves that are similar to the corners and therefore recognizes gestures with a 'close enough' approach, thus heeding the suggestion by Tu et al. for sharp corners drawn by finger.

Rather than using a touch-based system, the Nintendo Wii used a separate controller device to implement gestures into their games. According to Lee [22] the Wii Remote uses a combination of infrared sensors and an accelerometer. Lee determined this by disassembling a Wii Remote and taking a look at its internal components. He tested each of the internal components individually to determine their functionality within the controller. Though this device differs quite a bit from our touch-based system, it is not unlike our VR interaction methods for the second study.

Cummings [7] explores the history of video game interaction devices. He begins his investigation with early gaming controls that were made in the lab by researchers and were made of simple switches and buttons from whatever parts were lying around at the time. He then goes into the creation of early systems in arcades and then onto early consoles such as the Atari and Nintendo Entertainment System. This was the birth of the controller. As he continues his historical account, he gets into touchscreens and motion sensing devices such as the Nintendo DS and Wii respectively. In the end, he concludes that early gaming spurred the creation of individual interaction devices, whereas later gaming has taken to conforming to systems that already exist and utilizing

those devices rather than inventing new interaction devices.

By reviewing ten studies systematically, Simor et al. [32] found that there is no standardization to evaluating gestural interfaces. When they did their study, they looked into the Wii and Kinect as devices, gesture-based games, and a method for evaluating the usability of user interfaces. In the end, they found that gestural interfaces are in need of a method for evaluating usability for serious games in order to offer users more confidence, comfort, and welfare. They make note that this is especially the case for older adults.

2.2   Difficulty

We originally set out to determine how gesture presentation style might affect cognitive load. As we researched cognitive load and how it might relate to flow, we discovered that cognitive load is a piece of difficulty. As opposed to physical strain, cognitive load is specifically the mental strain caused by all of the tasks being performed by a person within a scenario. Because of this, we have included our literature review of cognitive load as part of the difficulty section rather than providing it with its own section.

Langdon et al. [20] performed a study in which they looked into user interaction with various products. They found that not many products cater to the cognitive decline of older users. As we age, our brain begins to physically break down and our cognitive resources become less, causing many tasks to become cognitively more difficult than they are for younger people. This difficulty can be offset by experience as muscle memory can lighten the cognitive load of a task. By performing repeated actions, people move from novice declarative moves to knowledge-based procedural moves. From the results of their study, Langdon et al. determined key elements in establishing a strategy to assess the inclusiveness of created artifacts. They lay the groundwork for a proposed

framework for the quantification of predictive cognitive interaction that is inclusive in design for people of an age variety and background.

According to the research of Joo et al. [17], the mental strain from a task can be affected by emotional strain. They performed a study where they had participants perform certain tasks with various goals. Those that had strong goals felt more positive emotions and decreased mental effort from those that had weaker goals. The participants that had weaker goals had more negative emotions and showed increased response time due to the increase in mental effort. In other words, if you have a goal that is not very strong, you are less interested in performing the task toward that goal. This lack of interest causes you to be slower at performing the task and in a time-related scenario increases the difficulty of performing the task.

In our text driven world, Dewan [8] found that there is precedence for using pictures instead of words in many cases. She mentions that it takes less mental effort to recognize and process pictures rather than words. She continues on to say that pictures are easier to recall than words. The reason for this is that words go into our long-term memory with a single code, whereas pictures have a dual-code nature to them. Pictures are both visual and verbal. When it comes to remembering pictures, it is likely that we will remember one of the two codes or both whereas the words alone have half as much chance to be remembered.

The work of Grady et al. [10] corroborates these findings by doing a study using a positron emission tomography to map the active areas of the brain during pictorial and word encoding. In their study they used three types of encoding strategies in order to explore interaction possibilities between the processing types and specificity of the material. They found that although the underlying mechanisms of encoding strategies operate similarly for both pictures and words, the encoding of pictures showed greater bilateral visual and medial temporal cortex activity. Word encoding resulted in increased activity of the prefrontal and temporoparietal regions of the brain which are

related to language function. From their results, they state that pictures were overall remembered better than words.

Srivastava and Vul [33] focused on looking into the memory performance differences between recall and recognition memory. Based on previous memory research, it was originally thought that recognition is always faster than recall, but Srivastava and Vul show that this is not necessarily the case in all instances. They performed an experiment in which the participants were shown stimuli-to-cue mappings and then asked questions about what they saw. They determined that the speed at which we remember an item is not specific to recall or recognition, but rather the association of the number of cues that help us to remember them. Recognition is faster than recall when the number of items exceeds the number of cues. However, recall is faster than recognition when the number of cues exceeds the number of items.

Many studies have their participants go through a set or sets of experiences then fill out questionnaires about those experiences. Hassenzahl and Sandweg [14] did a study to determine if this "summary assessment" is an effective method for obtaining user experience data. In the study they had participants go through a set of experiences and had them fill out the Subjective Mental Effort Questionnaire after each experience. At the end of the test, they were to fill out another questionnaire of the overall experience. They found that user summary assessments do not reflect their overall experience, rather they reflect the users' most recent experience. Because of this, it is important when doing a within-subjects study with questionnaires to provide the questionnaires between each experience and not all at once at the end of the session. We took this lesson to heart and incorporated it into our first study as it was a within-subjects study and we gave the participants the questionnaires between each intervention.

Witvoet [40] performed a study to determine if cognitive load influences a user's performance in a game-based learning task. During the study, she measured cognitive load using an electroen-

cephalogram. She defines performance in her study as the number of levels that were completed as well as changes to the user's conceptual knowledge. Her results showed no significant conceptual knowledge changes across the board and that cognitive load did not have an effect on performance.

2.3   Flow

From over twenty years of experience, Csikszentmihalyi [6] reports on the psychology behind what makes the optimal experience in an activity, also known as flow. He breaks it down into eight major components and mentions that every time somebody reflects on a positive experience, they mention at least one of these components. The components that make up good flow are: the task when confronted has a chance of being completed; the person must be able to concentrate on what they are doing; the task has clear goals; the task provides immediate feedback; the person acts with deep, effortless involvement that makes them forget about worries and frustrations of everyday life; the person has a sense of control over their actions; concern for oneself disappears; sense of time is altered.

Gregory [11] reports on the various aspects that make video games engaging and have flow. He begins by informing the reader that humans understand their world through story telling and that video games are a digital form of storytelling. By his report, there are nine flow conditions, all of which are universal and are not affected by individualistic things such as gender, race, education, or socio-economic status, etcetera. These conditions are as follows; clear goals, skill to difficulty balance, immediate feedback, deep concentration, problems are forgotten, self-consciousness disappears, altered sense of time, control is possible, rewarding. He defines these important flow conditions and mentions that each one contributes to flow. Having clear goals means the individual engaged in the activity is aware of what they want to do. When an individual's skill level is in balance with the task at hand, they have skill to difficulty balance. Immediate feedback refers to

a person knowing how well he or she is doing at any moment. When an individual focuses all of their attention on the task at hand, they have entered into a state of deep concentration. The "problems are forgotten" condition refers to the individual being able to dismiss concentration interfering irrelevant stimuli. While engaged in an activity, if the individual becomes so enthralled in the activity that they forget ones sense of self, their self-consciousness disappears. When a person loses track of time or time appears to pass faster than normal, they have an altered sense of time. If the activity allows for the individual to gain a feeling of mastery, then control is possible. If the experience of the activity is worth partaking in for its own sake, then it is intrinsically rewarding. Video games utilize many of these flow conditions and engage the player on multiple levels which is why they are so popular.

According to Schoenau-Fog [30], "Engagement is an essential element of the player experience." He set out to gain further understanding in the desire to continue playing electronic games. As defined by Schoenau-Fog, engagement is the level of desire to continue an experience. Through the use of surveys, Schoenau-Fog was able to generalize the concepts of engagement into eighteen conceptual categories within four main components. He lists these categories and components as follows: accomplishment - achievement, completion, and progression; activities - interfacing, socializing, solving, sensing, exploring, experimenting, creating and destroying, experiencing the story and characters; affect - positive, negative, and absorption; objectives - intrinsic or extrinsic. In the end, Schoenau-Fog re-defines engagement as a process, one in which the player engages in the pursuit of objectives, performs a range of activities to accomplish objectives and feel affect.

Sharek and Wiebe [31] performed a study to investigate a new method of measuring engagement in video games. This new method uses little task interference by capturing behavioral data. They point out the inherent flaw in using self-report questionnaires at the end of a task to determine level of engagement as has been used in most previous flow related research. The questionnaire process can be subjective because they are retrospective by nature and thus are a summative of the

17

experience. Post-task surveys are difficult for the participant to report on different states of affect at various stages of a single task, whereas attempting to do the surveys mid-task disrupts the participant's engagement in the task. They designed their experiment to take noninvasive measurements of frustration, boredom, and flow, in an attempt to determine if this line of thinking is plausible for future explorations of a similar nature. They found that a game clock combined with the NASA-TLX, for measuring cognitive load, is a useful analytical tool for user-behavior studies related to engagement. Our studies utilized this concept by including a game clock and had a NASA-TLX questionnaire given after each intervention.

Hern [15] set out to determine what makes games fun and what causes them to lack the 'fun factor' as he calls it. When deciding on features to make a game more fun, he mentions weighing the pro et contra of each feature to determine which one will give the best improvement in terms of fun factor. He breaks fun down into a subset of categories; creativity, understanding, power, surprise, humour, addictions, evolution, identification, and defying the player. He mentions that creativity should be used to help make the game more fun. By engaging the creative process of brainstorming, a designer can approach the problem from many more angles rather than from the first thing that comes to mind. The player needs to be able to understand what it is that they need to do. Though a different name, this is not unlike other flow research that states an activity should have clear goals. Power is the idea that the player should have the capability of doing things they cannot in real life. Nobody likes to feel powerless and Hern states that the new experience of having power draws the player into the game. Being surprised at what happens next draws the player further into the game. If the timeline cannot be guessed, the continuous changes affect the player on an emotional level. He specifically notes that humour is inherently fun if done well and suggests adding humour to games to increase the fun factor. By affecting curiosity and the player's ego with game elements, a designer can design an addicting game. In addition, having evolving elements in the game such as a level up skill system adds another goal for the player to accomplish.

The game will be more engaging if the player can identify with the character. This same concept is seen in movie and book creation and should be utilized when designing a video game. Defying the player refers to giving them a challenge, which is not unlike other flow research related to the skill level to difficulty balance. Many of these 'fun factors' that Hern mentions fall in direct alignment with the various conditions of flow. Since fun is part of flow, this makes sense that the fun factors affect flow or are almost exactly like the conditions of flow.

Nordberg [27] writes on his personal experience with game difficulty and how it relates to fun. He relates his experiences with that of others whom he has spoken with and those that have posted review articles. He begins by wondering why video games are said to be easier today than they were in the past. After looking into it for a bit, he discovered that older video games were on arcade machines and needed to be difficult in order to make money by having its players keep dropping quarters into the machine. Games today are on home consoles and are not in need of this requirement for funding. He then focuses on whether or not video games should be difficult. As he progresses through his experiences, he stumbles across aspects of flow such that he enjoys playing easy games that have clear goals and rewards. From his accounts of others' experiences, difficult games are fun and engaging due to the sense of accomplishment and reward systems in place. He mentions that different games are enjoyable for different people as a balance between skill level and difficulty plays a large role in the fun factor of the game. These findings fall in line with Gregory's [11] work.

In an attempt to investigate how game difficulty and player performance impact the enjoyment of a game, Klimmt et al. [19] performed a research study in which they recruited participants that self claimed to be expert gamers. In their study, there were three difficulties of the first person shooter game, Unreal Tournament 2, and after playing each difficulty, the participants completed questionnaires to determine their level of game enjoyment, their personal performance satisfaction, and their perceived game difficulty. They found that the most enjoyment playing the game occurred

when the difficulty of the game is at its easiest difficulty, even though this seems to contradict flow theory. The results of the experiment showed that players were more focused on their kill to death ratio as their performance marker and did not take into consideration the other objectives within the game. Klimmt et al. also mention that the short amount of time playing the game within the lab setting may not be enough time for a person to become bored with an easy task, especially when they are focused on quick successes rather than overall success from a greature duration of time. They suggest that the skill level to difficulty balance required for flow is dependent on the amount of time allotted for playing a game.

Buncher [4] ran an experiment where he had three groups of participants play a game, those that were stressed or overstimulated as he put it, those that were bored or under-stimulated, and a control group that was in the middle. The participants would then select their level of difficulty and repeat the action of playing the game. They would then repeat these steps again. As Buncher anticipated, participants that were in the stressed group selected lower levels of difficulty than those of the bored and control groups, but these findings did not have a significant effect. Similarly, even though there was no significance found, the results followed along with his predictions that stressed participants engaged in lower rates of play than the bored and control groups. He found that there was no significance correlating difficulty selection with stressed participants. However he did find that people are more likely to select lower difficulty levels the more bored they are. There was no significance in correlating boredom to rate of play, but there was significance found when correlating stress with rate of play. Buncher reports the more stressed a participant was, the slower they played the game. Though again the means of flow were higher starting out with the control group as predicted, these results showed no significance. He did find that flow had increased from the beginning of the study when the participant was assigned a difficulty to when the session was completed and the participant was playing a difficulty level they had chosen. The final discovery that Buncher had from his study was that a participant's perceived skill when compared to the

game's difficulty level gets better after repeated selection of difficulty levels.

By focusing on more observable aspects of flow, Klasen et al. [18] took a more physical approach to investigating flow experience with video games than other researchers. They used a functional magnetic resonance imaging (fMRI) device to scan the brain while their participants played a video game. In this game, they had several areas where there were no enemies as well as areas with various concentrations of enemies. They had some of the kills by the player give immediate feedback and some lack the feedback. From five neural correlate content factors that were calculated, they got significance in four of them having meaningful neural network activity. Joint activity in somatosensory networks and common motor system activation was seen in conjunction with event related activities. They were able to determine that the somatosensory network was being activated during factors that contribute to flow such as the balance in skill level and difficulty, clearly defined goals, control, concentration and focus. Areas of the motor network showed similar activation patterns for three of these same flow factors; difficulty to skill balance, control, concentration and focus. This aligned with the simulation of physical activity within the game and when those events correlated to flow factor events. They concluded that the neural correlates of flow can be captured on brain imaging technology and sensory-motor network activity contributes to flow.

Bederson [2] did a review of related literature and consolidated many thoughts on the concepts of flow in an attempt to determine how user interfaces affect flow. He says flow is a difficult concept to narrow down, but defines it as the optimal human experience. In terms of interface design, Bederson mentions that there is a lack of research into how poorly designed interfaces interrupt the user experience, thus breaking flow and he suggests that researchers look further into the way interface designs affect flow so that designers will have guidelines to creating interfaces that do not interrupt flow. By suggesting such research, the need for our concepts of looking into how gesture presentation can affect flow gains some additional validity.

CHAPTER 3: TOUCH-BASED GESTURE EXPERIMENT

We were interested in determining how gestures might affect how the player perceives a game that utilizes gestures as a gameplay element. In particular we were interested in how the gestures' presentation to players might affect the game's flow and difficulty. After our review of literature, we established hypotheses and designed a study to determine such behaviors.

3.1   Hypotheses

Presenting gestures via their names requires the player to recall the symbol associated with its name, whereas presenting gestures via the associated symbol circumvents the need to recall it. Our research goes a step beyond Srivastava et al.'s [33] psychological work by incorporating both tasks into a video game that has other tasks and distractions that can affect additional cognitive processes or cognitive load and therefore increase difficulty. Additionally, by testing for how various gestural presentation styles affect things like difficulty and flow, further theory can be developed about the presentation of gesture related interfaces. Therefore, we hypothesize:

- **H1:** *Presenting gestures by name will be significantly more difficult than presenting them as symbols.*

Using the lessons learned from the literature related to flow theory, a game was designed to have a set of goals that increases flow [4, 11, 15, 19, 27, 30]. The design of the game was iterated through pilot tests to make sure it is easy enough that the player does not get frustrated with it. The game builds upon previous work by implementing the goals as tasks that can be completed using gestural interventions. With this gestural implementation, the boundaries of how gestures were

implemented in the past were pushed, and the gestures were incorporated in a way that it would affect difficulty and flow. From this related literature, we anticipated:

- **H2:** *Presenting gestures by name will have significantly less flow than presenting them as symbols.*

- **H3:** *Game difficulty will be negatively associated with flow.*

Based on prior work [35], a gesture game was modified into a zombie game that can be toggled between two types of gesture presentation: name-based and symbols. The game setting for the study was based in a small town, where the primary goal of play was to kill as many zombies possible within a specified time period while trying to survive the entire duration of the time period. These two goals were chosen because they fit well into the zombie theme and provide goals for the player thus improving flow [4, 11, 30]. From the differences between these two tasks from one gesture presentation to the next, it was anticipated:

- **H4:** *Presenting gestures by name will have less survival time than presenting them as symbols.*

- **H5:** *Presenting gestures by name will have a lower kill count than presenting them as symbols.*

We will go into the additional game details about the touch-based gesture drawing game that was designed to test these hypotheses a little later. The first user study was performed using this game.

3.2   Experimental Design

Because flow is affected by the difficulty-to-skill level ratio [4, 11, 30], we used a within-subjects design to implement this study. This allowed for a direct flow comparison for each intervention as the general skill level of the participant remained constant for each individual comparison across both interventions. A between-subjects study would have introduced skill level biasing and thus would have made it exceedingly difficult to compare flow results from participant to participant. To eliminate ambiguity of the gesture stroke direction, a dot was placed at the gesture starting point on the gesture reference sheet that was used for training the participants as can be seen in Figure 3.4. This was done in order to help reduce difficulty biasing introduced by natural perceived directional differences between participants. Even with the dot in place on the gesture reference sheet, many participants still attempted to draw the gestures in the direction that they were most familiar.

The study was counterbalanced by alternating the two conditions of the independent variable of gesture presentation style. Participants with an even ID number started with the game iteration that has the symbol names written over the zombies' heads. Participants with an odd ID number started with the game iteration that has the symbol drawn over the zombies' heads. This was done to mitigate bias caused by ordering effects.

A trial run of the game to get familiar with its gameplay and the gestures themselves was written into the study to help mitigate the learning effects from the participant going from one intervention to the other.

3.3 A Touch-based Gesture Game of Killing Zombies

Our game features a small town that is overrun with zombies. The player has the ability to use gestures to kill zombies while they try to survive within the town. From a birds-eye view, the player can see their own character, who has a hard hat and a light blue ring around him/her to differentiate the player from the non-playable characters (NPCs). The player moves their character by touching anywhere within this ring and dragging in any direction. As the player does this, a line is drawn in red and the character follows this path. If the player draws a path to the edge of the screen, then ceases movement, but keeps touching the screen, the field of view will rotate to allow the player to navigate farther than what is viewable on a single screen. There is a margin line that is displayed in gray to indicate the area of the screen where this effect occurs. Gray was chosen as it is a neutral color and previous studies that we performed showed that the color did not matter in user perspective of the game. For marketability, versions of the game after this study was complete include options to change the border color to any color in the RGBA spectrum. In addition, the camera rotation style and speed can be adjusted via the game options screen. During the study, these features were locked at a single setting that was determined from the average of best setting from earlier studies.

One hundred zombies (NPCs) are spawned in random valid locations within the town. In other words, the zombies cannot spawn in the same location as another zombie, nor can they spawn inside other objects such as cars, buildings, etcetera. In order to guarantee that the zombies spawn in valid locations, a spherecast is performed using a sphere with a slightly larger radius than the hitboxes of the zombies. This spherecast is calculated by casting from the same level as the camera down to the navmesh (the allowable moving space for the zombies and player character, i.e. streets, sidewalks, etc.). If the sphere intersects with any objects, a new random location is chosen for the zombie. This continues until a valid location is found. Their spawn height is set just above the
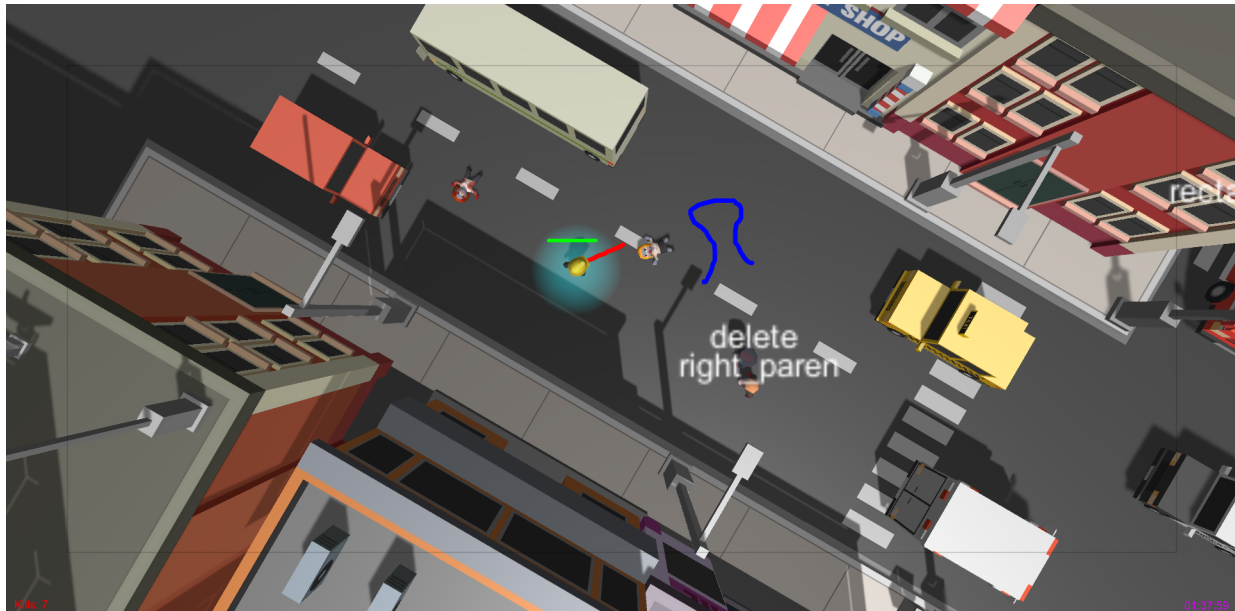
Figure 3.1: The above figure shows a couple of dead zombies, the player's character, the path being followed in red, a gesture drawn in blue, and a couple of zombies with the names of the symbols above their heads.

street level to guarantee they are not spawned below the surface of the navmesh.

Each zombie emits a moaning sound at random intervals. The zombies wander aimlessly around the town (upon the navmesh) until they are within visual distance of the player's character, in which case they begin to chase the player's character. If they come into contact with the player's character, they start to deal damage to the character, as indicated by the health bar above the player character's head (See Figure 3.1). In addition, a biting sound is played to give an audible sense that the zombie is biting the player's character. A magenta timer and a red kill counter are positioned in the lower corners of the screen as elements of the heads up display (HUD). These features can be seen in Figure 3.1 and Figure 3.2. When the player's health reaches zero, the timer stops, the character dies, and becomes a zombie. A leader board feature was added as a way to port to a fully

Figure 3.2: This figure shows a few dead zombies, the character and his ring, a gesture drawn in blue, and a few zombies with the green symbols above their heads.

marketable game later. This feature was only used during the study as a way of knowing when the player's character has died as it will display upon death. The associated scoring feature was disabled during the user study to help keep the users objective about their opinion of how well they performed as is asked in the GEQ.

As each zombie is spawned, it is assigned a symbol that will kill it. The zombie will either have the symbol drawn in green above its head, or the symbol's name in white written above its head, depending on the game iteration that is being played at the time. The player is able to draw non-path lines (unistroke gestures) in blue when they touch and drag outside of the player character's ring. If a gesture is drawn that matches the symbol assigned to a zombie, the zombie will die when it comes into contact with the gesture. To accomplish this, the player can draw the gesture directly on the zombie, instantly killing it, or they can draw the gesture on the ground in the path

of the character to provoke the zombie into walking into the gesture, at which point the zombie will die. Figure 3.2 shows the iteration of the game with the gesture symbols and Figure 3.1 shows the iteration with the gesture names.

Only one gesture can be drawn at a time. Beginning a new gesture will remove the previously drawn gesture. Likewise, a gesture will only work once. If two zombies are approaching with the same symbol, the gesture will only kill one of them and a new gesture will need to be drawn to kill the second zombie. There is a kill counter on the lower left side of the screen which increments by one each time a zombie is killed.

There are seven health pickups that are placed in the town at designated locations at the start of the game. Collecting one of these will replenish health that has been lost by the player's character. Every fifth zombie that is killed will generate a health pickup in a random location. Early pilot tests showed that spawning a health pickup every third zombie death was too frequent to make the game challenging, so it was increased to every fifth zombie death prior to the study. Double tapping the screen anywhere outside the character's ring will display a map of the city. The map displays the location of the player's character, the path that has been drawn, and the location of all of the health pickups. A single tap will close the map and go back into the game.

The gesture interpreter used in this game is an easy to train and easy to deploy gesture system called Jackknife, originally developed by Taranta et al. [36]. Training is done in real-time rather than scanning pre-drawn images into the computer. This was done during the game design phase so participants were replicating the gestures we entered, and not creating their own. There are ten gestures used in this game and they are shown in Figure 3.4. These gestures were chosen because they have a variety of complexities from simple to complex with some in the middle. They were implemented as a gameplay mechanic because it imposes continual use rather than intermittent usage such as a menu system would impose.

3.4   Technical Contributions

Various forms of this game were used in multiple user studies testing out different features. Some of these studies were run during simultaneous development of the game for this study. There are a few contributions we made to modifying the original game that was created to adjust it for this study. The most prominent change we made to the game is the zombie skins. From the original game, we re-textured the Simple Town characters into Simple Zombies.

We also added music and sound effects. We got a free license to using some creepy music for user studies as well as zombie moaning sounds and bite sounds. When the player takes damage from a zombie, the bite sound is played. Moans occur at random intervals from each zombie, so the distance in the game affects how loud the moan is. The creepy music plays in the background the entire time the game is running.

The healthbar and health tracking system were added, such as gaining health and taking damage. The healthbar uses the same Ink Object in Unity that the gesture symbols use to be drawn. We created the health pickups using Unity's base cube assets, adding the glow effect and constant angled rotation. Everything associated with the health pickups was also added for this study, such as spawning, collecting, etcetera.

Adjustments were made to the zombie AIs, to cause them to run toward the player's character when they see the character. Upon colliding with the character, the character takes damage that is represented by the healthbar getting smaller. The kill counter and survival timer were also added for the purposes of adding goals to the study to increase flow. The survival time clock stops ticking when all of the zombies are killed.

The ink for drawn gestures turned black and remained in the game in our previous studies. For this study, we changed it to be removed when a zombie is killed. Additional raycasting was utilized

Figure 3.3: HP EliteBook 2760p.

to prevent the player from being able to path their character under objects such as umbrellas and
bridges. This prevents them from being able to release their path in strange locations that would
get them stuck.

3.5   Technical Specifications and Apparatus

The game was built in Unity3D version 2019.1.11f1 and was deployed on a HP EliteBook 2760p.
This particular hardware device was chosen for this study due to its ability to run Unity, its touch
screen capabilities, and its availability within the Interactive Systems and User Experience Labo-
ratory at the University of Central Florida. With several backups of the device, it was simple to
deploy the game on multiple machines and have them ready and loaded prior to each session with
participants so that they could quickly be swapped if any study-stopping errors occurred during the

user study phase of the study. The device is shown in Figure 3.3.

3.6   User Study Procedure

The first user study took place using a touch-based game world where the player fights off zombies by drawing the appropriate gesture that matches a particular zombie's symbol. The participant was given a demographics survey to fill out. Upon completion of the demographics survey, they were shown a list of the symbols and gestures in the game and their names. Figure 3.4 shows the exact list they were provided in the study. They were given ample time to get familiar with the list. Once they were ready to proceed, they were instructed on how the game works and were given a trial run of the game to get familiar with it. During the trial run, they were allowed to reference the symbol and gesture sheet as they were likely still getting familiar with the more complex symbols and gestures.

The participant then played one of two iterations of the game, trying to kill as many zombies as possible and have their character survive the entire time limit of three minutes of game clock time as determined by our pilot testing. After playing the first iteration of the game, the participant filled out a Game Experience Questionnaire (GEQ) [16] and NASA Task Load Index (NASA-TLX) [13]. After filling out these surveys, the participant repeated playing the game, but this time they played the iteration of the game that they did not play the first time. They then filled out the GEQ and NASA-TLX surveys again, in relation to this intervention. This method of questionnaire distribution is supported by the results of Hassenzahl et al. in [14] as it reduces recall biasing. The GEQ was chosen because it is a validated questionnaire that obtains the data we are looking for, flow and difficulty. The NASA-TLX is also validated and obtains the cognitive load of a task.

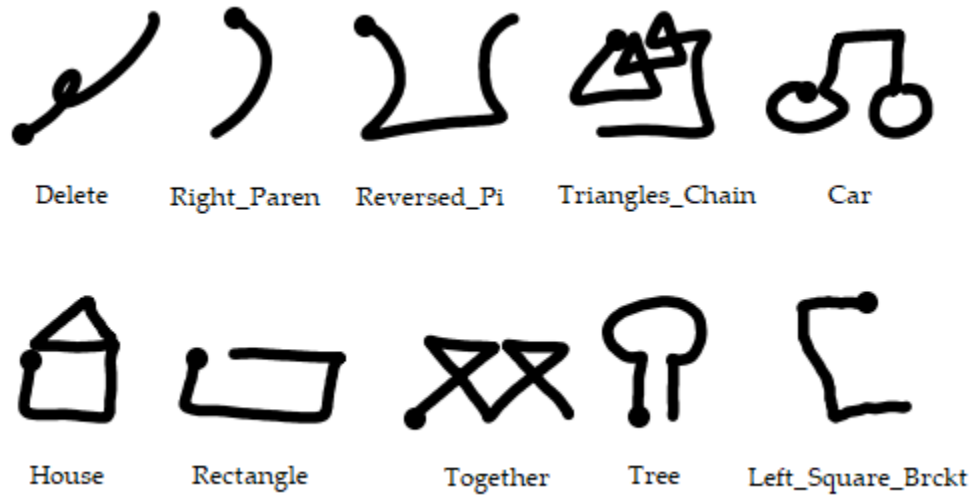After both game interventions and associated surveys were completed, the participant was asked,

Figure 3.4: The subset of touch-based gestures presented to users is shown above using symbols (top) and names (bottom).

"Which of the two styles of the game do you like better?" and "Why?" This kind of questioning was used as a way of helping us to confirm the results of the GEQ survey. In addition, it helped us to understand the effect of difficulty on flow, whether that effect be an increase or a decrease. The participant was given a chance to ask any questions they may have had. These questions were answered, then they were thanked and given ten US dollars cash for their time. The most common question we received from participants was if and when our game would become a game on an app store so that it can be downloaded and played later.

3.7 Findings

We begin with our touch-based gesture analysis approach and findings.
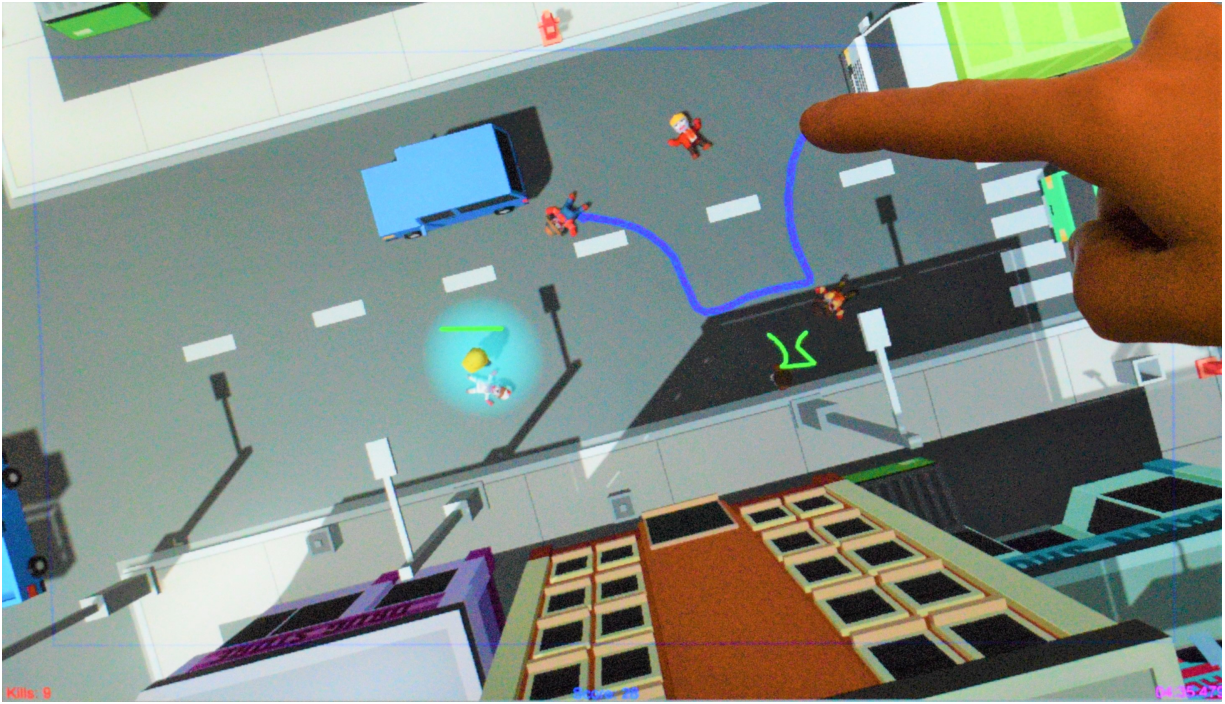
Figure 3.5: The player is shown drawing a gesture that is the same as the shown zombie in an attempt to kill the zombie.

3.7.1 Dependent Measures

The NASA-TLX uses a 21-pt Likert scale per question to rank tasks. We coded these results as raw TLX scores by following the provided instructions. The GEQ uses a 5-pt Likert scale per question and the results were coded by following its accompanied instructions. This produces 12 categories from the 17 questions that were asked. The two categories that we were interested in were the level of challenge (difficulty) and flow. In addition to the GEQ, we used a counter in-game to keep track of the number of kills the player obtained. The game also kept track of the time played, and the timer stopped if the player died in-game. When the player finished playing with each gesture presentation, we wrote down the number of kills and the time that they survived. If the participant

survived the entire time up to the limit, their survival time was written as that maximum time and their kill count at that point in time was written down.

3.7.2 *Data Analysis Approach*

We first assessed the reliability of our measures using Cronbach's alpha and found our measures to be reliable (Difficulty: $\alpha = .822$, Flow: $\alpha = .844$). We then screened our measures for normality using a Shapiro-Wilk test. Since difficulty and flow were not normality distributed, we used the appropriate non-parametric Wilcoxon-Signed Ranks Test [21]. A Spearman's Rank Correlation Coefficient Test was used across both interventions to determine if difficulty has a direct correlation to flow. From our performance data we tested our kill counts and survival time for significance across both interventions. Since they too were not normally distributed, we used the Wilcoxon Signed Ranks Test for both.

In addition to our statistical analyses, we performed a post hoc qualitative analysis on the open-ended questions asked to participants during the debriefing session. Specifically, participants were asked which gesture they preferred and why. We coded this qualitative data based on their preference and the underlying reason for their preference.

The responses to why the participants preferred one method of gestural presentation over the other were analyzed and coded to categories based off their meaning. There were five common themes that emerged from this data and we used them as our categories; easier, more fun, more challenging, more clear, and less restrictive.

3.7.3   Recruitment and Participants

An a priori power analysis was performed using G*Power to determine the sample size for this study. An effect size of 0.5, α = 0.05, power = 0.8, was used and the sample size required to detect significance was determined to be 34 participants. Participants were primarily recruited via a sign-up sheet posted on www.signupgenius.com. Additional participants were recruited by word of mouth through impromptu verbal solicitation. After reviewing the consent form and obtaining proper consent, the participant was given a chance to have any questions answered before proceeding into the study.

We had a reasonably representative population of participants. There were 21 males and 13 females whose ages ranged from 18 to 60+ and whose education levels varied from high school diploma to post doctorate. We had 32 participants who are right-handed when writing and 2 that are left-handed, with no ambidextrous writers. When it came to using touch-based devices, we had 25 participants that primarily use their right hand, 4 that use their left, and 5 that were ambidextrous. Likewise, we had 6 participants that never use pen-based devices, 14 that rarely use them, 12 that sometimes use them, and 2 that frequently use them. Four of our participants never play video games, 4 rarely play, 9 sometimes play, 12 frequently play, and 5 always play. When it comes to experience using a touch-based device, we had one participant that never uses them, one that sometimes uses them, 13 that frequently use them, and 19 that always use them. Nineteen of our participants correct their vision with either glasses or contacts, while fifteen of them have no correction.

Table 3.1: The standard statistics of our dependent variables, difficulty and flow. Each value is out of a possible range from 0-5.

	2D Difficulty and Flow						
DV	*Min*	*Max*	*Range*	*Mean*	*Median*	*Mode*	*StDev*
Difficulty:							
Symbols	0.0	3.0	3.0	1.16	1.0	1.2	0.76
Names	0.6	3.2	2.6	1.62	1.6	1.8	0.68
Overall	0.0	3.2	3.2	1.39	1.3	1.8	0.75
Flow:							
Symbols	0.4	4.0	3.6	2.45	2.3	4.0	1.11
Names	0.4	4.0	3.6	2.52	2.4	4.0	1.12
Overall	0.4	4.0	3.6	2.49	2.3	4.0	1.11

3.7.4   Hypothesis Testing Results

Table 3.1 summarizes the descriptive statistics of our two conditions in relation to difficulty and flow. Both the minimum difficulty value and the maximum difficulty value are higher when it comes to gestures presented as names. As per the mean, median, and mode, we can see that gestures presented as names have generally higher difficulty than gestures presented as symbols but they are generally pretty low, all occurring under 50% of the overall possible maximum value. The mean difficulty value increases from 1.16 for symbols to 1.62 for names, a 39.6% increase. There is not much change in flow from the symbols presentation to the names presentation. They have the same minimum and maximum values, resulting in the same range. Their means and medians have minor differences, the mean only increases by 2.8%. With the mode being the max value for the ranking, it shows that most participants reached maximum flow regardless of gesture presentation style.

Table 3.2 summarizes the descriptive statistics of our two conditions in relation to kill count and survival time. Both the minimum kills and the maximum kills obtained occurred within our pre-

36

Table 3.2: The performance of our participants in terms of zombie kills and survival time. This was out of 100 total possible kills and a max survival time of 7:40. Time is shown in the format of m:ss.

	2D Performance						
Task	*Min*	*Max*	*Range*	*Mean*	*Median*	*Mode*	*StDev*
Kills:							
Symbols	7	81	74	54.0	57.5	60	22.3
Names	5	87	82	47.7	50.0	50	24.3
Overall	5	87	82	50.9	56.5	50	23.4
Time:							
Symbols	1:54	7:40	5:46	6:49	7:40	7:40	1:51
Names	1:16	7:40	6:24	6:55	7:40	7:40	1:49
Overall	1:16	7:40	6:24	6:52	7:40	7:40	1:49

sentation of names with no participant killing all of the zombies. Players are more consistent in their kill count when presented with symbols than they are when presented with names. The mean number of kills resides fairly close to half the total number of zombies in both interventions with symbols residing slightly more than names, causing an overall mean to be almost exactly half. The most common occurrence of kills is higher in the symbols presentation than it is in the names presentation. When it comes to survival time, the times weigh heavily toward the maximum time as seen by the mean, median, and mode. Most participants survived the entire allotted time, in fact so many people survived the full time that it caused the median to be the maximum time. This tells us that the task of killing the zombies was difficult enough that nobody was able to complete it, however the task of surviving the entire duration was easy enough that most people were able to complete the task.

For H1, we assessed the main effect of gesture presentation on difficulty. We found a significant effect ($Z = -3.910$, $p < .001$), where gestures presented by names were considered significantly more difficult than when they were represented by symbols. For H2, we assessed the main effect

Table 3.3: This table shows which hypotheses are supported and which ones are rejected along with the values used for determination.

Hypothesis Results

Hypotheses	Values	Determination
H1	(Z = -3.910, p < .001***)	Supported
H2	(Z = -0.343, p = .732)	Unsupported
H3	(r = 0.411, p < .01**)	Unsupported
H4	(Z = -0.415, p = .678)	Unsupported
H5	(Z = -1.936, p = .053)	Unsupported

(**Key:** * < .05, ** < .01, *** < .001)

of gesture presentation on flow. We found no significant effect (Z = -.343, p = .732), where gesture presentation does not significantly have an effect on flow. For H3, we assessed the correlation between difficulty and flow. We found a significant effect (r = .411, p < .01), where flow increases as difficulty increases. This effect can be seen in Figure 3.6. For H4, we assessed the main effect of gesture presentation on survival time. We found no significant effect (Z = -.415, p = .678), where gesture presentation does not significantly have an effect on survival time. For H5, we assessed the main effect of gesture presentation on the players' number of kills. We found no significant effect (Z = -1.936, p = .053), where gesture presentation does not significantly have an effect on the players' number of kills. Table 3.3 summarizes these findings.

3.7.5   Post Hoc Analyses and Qualitative Results

Given that some of our hypothesized relationships were unsupported and/or were significant in the opposite direction than we anticipated, we proceeded to conduct a post hoc analyses to provide further insights based on our data. There were a few interesting patterns that emerged from the extra data we collected.

38
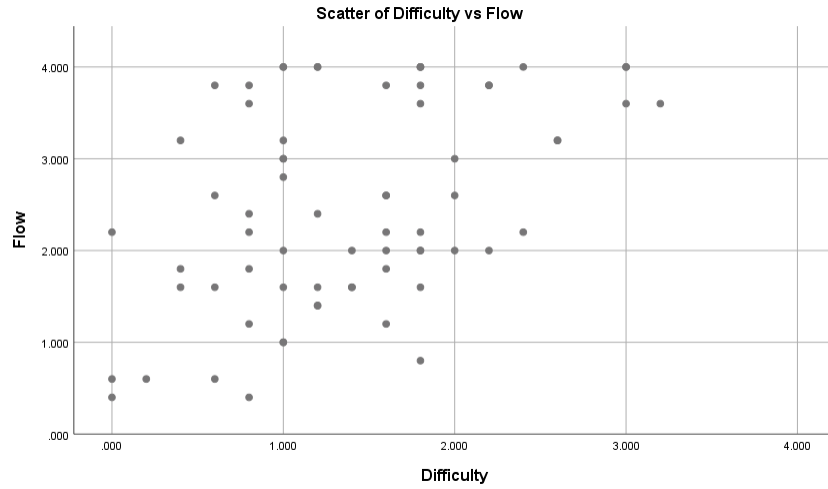
Figure 3.6: This figure shows difficulty plotted vs flow.

In lieu of our within-subjects study design, we saw significant learning effects in regards to the game's difficulty in both comparisons, symbols first vs names first ($F_{1,16}$ = 75.629, p < .001) and symbols second vs names second ($F_{1,16}$ = 129.167, p < .001). With this learning effect data, we obtained the means and standard deviations of the orderings; symbols first: M = 1.20000, SD = 0.689202, symbols second: M = 1.76471, SD = 0.704398, names first: M = 1.12941, SD = 0.848355, names second: M = 1.48235, SD = 0.640542.

Of our 34 participants when asked which gestural presentation they preferred, we had 3 (9%) participants with no preference, 18 (53%) that preferred the symbols, and 13 (38%) that preferred the names. The reasons they gave for their preference ranged from preferring how easy the symbols are to recognize to liking the added challenge of the names. We categorized the participants' comments and these results are shown in Table 3.4. When participants preferred symbols over names, the predominant reason (N=16) was they found this gesture presentation more intuitive and easier to use. This was consistent with our hypothesis testing results, where we found that gesture

39

Table 3.4: This table shows the gesture presentation style preference of our participants and the reasons given as to why they preferred that style. This table excludes the three participants that had no preference.

	Preference	
Reason	*Symbols*	*Names*
Easier	16	1
More Fun	2	0
More Challenging	0	7
More Clear	0	3
Less Restrictive	0	2
Totals	18	13

names had significantly increased difficulty. We present some of the categorized comments the participants made.

3.7.5.1 Easier

The easier symbol-based gesture presentation made participants feel more relaxed:

"The symbols was easier to play, more relaxed to play."

And, because participants did not have to recall the gesture from the name, this took less memory:

"The names requires you to remember, which is harder. Symbols is less memory, it is easy."

3.7.5.2 More Fun

Some of the participants really enjoyed the way the symbols were presented as a more intuitive way of representing what needs to be done to kill the zombies:

"The symbols felt more engaging how the solution is presented."

where they could rapidly kill zombies without much thinking, but would have enjoyed the names if given more time to play the game:

"I liked the challenge of the names if given more time to get familiar with the gestures. The symbols gave more of a sense of achievement because you are able to accomplish more."

3.7.5.3   More Challenging

Several participants preferred names over symbols, the predominant reason (N=7) was they enjoyed the added challenge imposed by the memory aspect of recalling the gesture:

"I liked the challenge of it. Was more demanding."

This falls in line with our hypothesis testing results, where we found that as the difficulty increased, flow also increased. Although we had some participants claim to not have a preference, it was clear by their comments that they had a preference:

"The symbols was easier, but I liked the the added challenge of names."

3.7.5.4   More Clear

A few of the participants had preferences that were not based on difficulty of task rather visibility of the symbols. Due to the size of the touch screen, the symbols were reduced in size compared to the sample handout that they were allowed to use during their trial run. Because of this, a few participants (N=3) preferred the names because they could read it whereas they had difficulty interpreting the symbols in their small size in-game:

"The symbols were cluttered and hard to distinguish compared to the printout."

3.7.5.5   Less Restrictive

Similarly, some participants felt that the symbols made it seem like they had to be copied exactly, whereas the names were more interpretive and allowed freedom of variance in performing the gestures:

"The symbols seemed more restrictive, like an expectation."

This was an interesting find since the same gesture recognizer was used for both interventions and the player had the freedom of variance all along in both cases. Overall our game had good flow as many of our participants enjoyed it to the point that they requested that we add more levels and put the game into the mobile device application store.

There was an interesting trend of performing the gestures incorrectly, in which case we did not say anything so as not to bias the results. We did not collect data for when this occurred, but we did notice that several participants would attempt to draw the gestures in a different direction from which they were shown. In addition, there was a tendency in several participants to get confused between which direction a right parenthesis and a left square bracket faces when given the names. This trend stands out because these two symbols are specifically the only two symbols from our gesture set that are of common usage. They are both characters on the keyboard, so we anticipated that participants would be more familiar with these two symbols than any of the other less commonly used symbols.

3.8   Discussion

We discuss our results and interpret their meaning.

3.8.1   *Gesture Presentation, Difficulty, and Flow in Touch-based Video Games*

As expected from the work of Srivastava et al. [33], we found that gestures presented in the form of names are more difficult than gestures presented as symbols since names uses recollection over recognition as with the symbols. This falls in line with Dewan [8] and Grady et al. [10] that pictures are easier to recall than words. As gesture presentations, we have confirmed that this holds true. Not only is it easier to recall pictures over words, but as a gesture presentation, it has a direct effect on the difficulty of the game. This implies that we should keep in mind how our gestural presentation can make a game more difficult. If we do not take this into consideration when designing games, we can end up making games too difficult for our players and cause them undue frustration, which would reduce flow, or we can intentionally make it more challenging, to increase flow [4, 11, 30], which some participants liked.

The most likely reason we did not see any significance in the main effect of gesture presentation on flow is that the two interventions were too similar to each other, thus having similar flow. This is a likely reason as the data in Table 3.1 shows that flow is only marginally different between the two interventions. Another possible reason for this hypothesis discrepancy is that we did not have enough gesture presentation styles to truly gauge the differences in flow that they cause. An increase in the amount of gestures or the complexity of gestures would make things more complex and likely increase difficulty [42]. Another possibility would be to use multistroke gestures [1]. If the difficulty increases enough by one of these methods, we should be able to get the game to become difficult enough to get results on the other side of the curvilinear relationship mentioned

43

by Schoenau-Fog [30]. Further studies will need to be done to see if any of these possible scenarios have an effect on flow. If there is an effect, we would need to first increase the level of gestural difficulty then perform research into finding its optimal difficulty. These possible results can greatly help us to design better games by helping us to increase difficulty without overloading the players with gestural challenge.

Despite the trend we saw in early pilot studies, we determined the opposite of our hypothesis, when correlating difficulty with flow. Our results imply a linear relationship between difficulty and flow, but the work of Buncher [4] implies a curvilinear relationship. The linearity of our results can be seen in Figure 3.6. The reason for this relational difference would imply that the difficulty differences between our two interventions is not a large enough difference to be on opposite sides of the curve. In other words, our participants' skill level was high enough that our game was too easy in both scenarios [4, 11, 30], causing our entire group of participants to lie on the upward slope of the curvilinear relationship between difficulty and flow. This relationship is supported by our survival time and difficulty results as seen in Table 3.2. Most of our participants survived the entire duration of the allotted time, implying that this task was an easy task and our difficulty results trended toward the lower half of the overall realm of possibilities, implying the game was easy in general. It is possible that this overall low difficulty level may be the reason we did not see significant effects with flow. To gain further insight into this relationship, we calculated skill level and compared the results. After we calculated the skill level of each participant per intervention, and we used that to determine their difficulty-to-skill level ratio, we were able to affirm this relationship as reported in the related literature [4, 11, 30]. Since our study population was entirely made of adults, additional insight may be deemed from doing the same study with children to which the skill level may be a better match with the game difficulty. The results of such a study can also be compared to this study's results to determine if there is any difference in the effects based off childhood vs adulthood.

3.8.2   Gestural Presentation and Performance

Even though the names intervention turned out to be generally more difficult, some players had a higher skill level with that intervention. Additionally, as seen from our qualitative data, some participants felt the symbols were not presented to them as clearly as the text, so that would affect their kill count between both interventions. This may have caused an opposite effect from what Dewan [8] reported as the symbols were so muddled that the names became easier to recognize and recall than to interpret the symbol itself. This was likely caused by the variance in the samples when recorded through the Jackknife gesture recognizer [35, 36] as there was intentional variance with the samples due to recording samples from several people with varying speed and accuracy. One way to make these symbols more clear in the future would be to have the line renderer use a thinner line to draw the symbols and/or increase their size in the game. These possible adjustments would increase the negative space on the symbols, allowing for more distinct lines, increasing their clarity. An additional option to reduce the possibility of getting muddled symbols is to remove the random sampling of gestures and only display the best (most clear) sample of each type of symbol. The gesture accuracy of the Jackknife gesture recognizer [36] and participants was not analyzed as it was part of prior work by Taranta et al. [35] in a similar setting. While conducting the study, it did not seem like any gestures that were drawn were failed to be recognized by the recognizer, and no participants mentioned having any difficulty performing the gestures. As such, their silence does not necessarily mean that there were no issues, additional research would need to be performed to determine if there may have been accuracy issues that might have influenced our results.

By a direct mean comparison of symbols order and of names order, we can conclude that the learning effect we saw is greater when going from names to symbols than it is when going from symbols to names. This is caused by the memory aspect of each intervention. When participants were performing the names intervention first, they were associating the names with each symbol

45

as they played the game. When they switched to the symbols intervention, they already knew the symbols, but now they did not need to associate the name anymore, making it an already learned task. The opposite was true for the participants that did symbols first, as they played the game for the symbols intervention, they did not have to associate names with the symbols, so they forgot about the names that were shown to them on the gesture sheet during the training session. When they moved to the names intervention, they already knew the symbols, so that was the learning effect, but they still had learning to do as they were then learning which symbols go with which names.

Another factor that may have resulted in the tilt of balance between interventions is the randomness of game behavior. The zombies were spawned into random locations and would choose random paths that would sometimes cause them to cluster in areas of the city. If a large cluster was found while performing one intervention, but not the other, it would greatly skew the kill count and survival time between interventions if the player got swarmed in one and not the other. We noticed an interesting trend during the study where many participants would encounter a horde and start killing zombies as fast as they could while forgetting about paying attention to their character. There were a few times when a zombie wandered in from off the screen and began attacking the player's character without the participant noticing.

3.8.3   Implications for Design

Gestures are a good mechanic for gameplay and should be explored further to better help gesture-based game designers. Our participants really enjoyed our game but could benefit further from a harder version of the game. As seen by Buncher [4] the difficulty of a game affects its flow, even if a change in difficulty is caused by a change in gesture presentation. When presenting the gestures as symbols, the game can become very simple and easy, and thus may benefit from a

larger gesture set. When presented as names, the game becomes more difficult than just presenting them as symbols, but not such a large difference to change the flow of the game. When designing games to use gestures as a game mechanic, we recommend doing pilot studies to determine if the difficulty of the game needs to be greater or lessened. If the game requires more difficulty, increasing gesture complexity [5] or increasing the size of the gesture set are both evident ways of increasing a game's difficulty. But, switching to gestures presented as names is an unconventional approach that will slightly increase the difficulty by adding an extra challenge to the game. It may be pertinent to use a larger gesture set in combination with presenting the gestures as names, as this should cause an exponential increase in difficulty since each gesture has increased cues from symbols to names [33].

CHAPTER 4: VIRTUAL REALITY GESTURE EXPERIMENT

From the first study, we found that both interventions were too easy overall for our participants. This drove us to design the second study in virtual reality. By going three dimensional, we were interested to see if the task difficulty increased to a level that makes a difference from the first study. We were also interested in determining if the results we obtained from the first study holds true for different game platform modalities and not just touchscreen games.

4.1   Zombie Conversion to Virtual Reality

The touch-based version of the game was modified to a virtual reality modality. The VR version of the game is exactly the same as the touch-based version when it comes to the 3D models, spawning, and sounds. The other aspects of the game had to be modified in order to make the VR version as similar to the touch-based version as possible.

The symbols and names were only modified slightly with the exception that different gestures were used to provide more three dimensional gestures rather than 2D gestures in 3D space. The names and symbols remain above the zombies heads, but the names are set to always face the player. This makes the symbols readable from any angle and allows the player more freedom in killing the zombies without having to run around to the correct side to read the gesture name. The symbols were difficult to tell what they are in 3D space if they constantly face the player, so they were set to face a static direction so the player can walk around the symbol and get a sense of its 3D features.

The touch-based healthbar floated above the player character's head as a way of remaining on the screen. This was modified in the VR version to always remain on the screen in the format of an element on the HUD. The kill counter and timer were also kept as elements of the HUD but moved
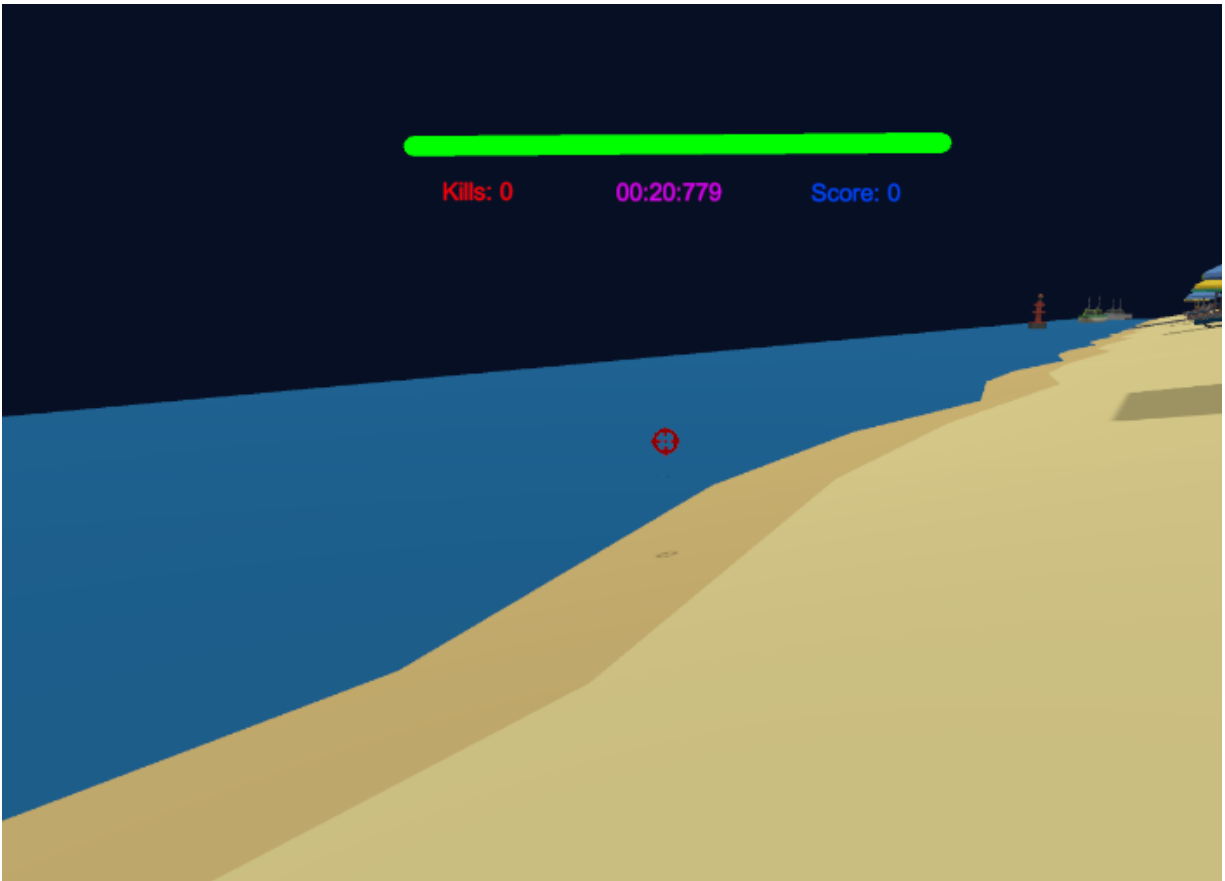
Figure 4.1: HUD for VR: Healthbar, timer, killcount, and score (not used in this study).

from the bottom of the screen to the top of the screen just underneath the healthbar so as to free up the viewable area for the player. This can be seen in Figure 4.1.

In order to maintain the freedom of gesture placement that the touch-based game offers, the VR version of the game was designed to enlarge any drawn gestures and project them to the location at which the player is looking. A targeting reticle is located on the nearest object to which the player is looking. This was done by tracing a ray down the viewpoint of the camera and obtaining the position of the first intersection of an object. An example of a gesture being projected to a

Figure 4.2: The wheel gesture projected to the targeting reticle. (This figure is next in sequence with Figure 4.9)

zombie and killing that zombie can be seen in Figure 4.2. The targeting reticle off in the distance in-game can be seen in Figure 4.1 and a closeup of the implemented targeting reticle can be seen in Figure 4.3.

The map icons were recreated for the 3D version as the 2D map enlarged the player and health items to show their location on the map. In the VR version, the player character object is not a visible object as it was in the 2D version. Additionally zooming out and enlarging the objects to view the map might cause disorientation in VR, so icons were created to add to the map and

Figure 4.3: Targeting reticle for VR.

Figure 4.4: Player map icon for VR.

opening the map attaches it to the controller. The icons that were used can be seen in Figures 4.4 and 4.5.

4.2   Technical Specifications and Apparatus

The virtual reality zombie game was built in Unity3D version 2019.1.11f1 and deployed on an HP Z VR Backpack G1 Workstation along with an HTC Vive acting as the interaction device for the participants. The computer used to run the VR equipment can be seen in Figure 4.6. This study utilized the Vive's head mounted display (HMD) unit, two Vive controllers, and two base stations (lighthouses) as can be seen in Figure 4.7.

Figure 4.5: HP map icon for VR.

Figure 4.6: HP Z VR Backpack G1 Workstation.

4.3 Experimental Design

In order to allow for a more direct comparison between touch-based gestures and VR gestures, the design of the VR study and game were kept as similar as possible while still incorporating the lessons learned from the first study. As seen from the first study, the training session was not sufficient enough to mitigate learning effects from one intervention to the other, so the VR study was designed with this lesson in mind to be a between-subjects study instead of the earlier within-subjects design that the touch-based study utilized. This caused our required number of

Figure 4.7: HTC Vive HMD, 2 Controllers, and 2 Lighthouses.

participants to jump up to 128 for this second study.

We also found with the first study that our game may have been too easy for the majority of players, causing our mean scores for kills and time to be on the high end, with the median and mode of the time being the maximum time limit. Due to the skew caused by this issue, we decided to remove the time limit on this second study and instead have the participant survive as long as possible while trying to kill as many zombies as possible.

Our first study also had an issue with the game clock being driven by the framerate of the hardware and therefore a conversion was needed for the time limit. The VR game was modified to fix this issue by having the clock within the game run based off delta time.
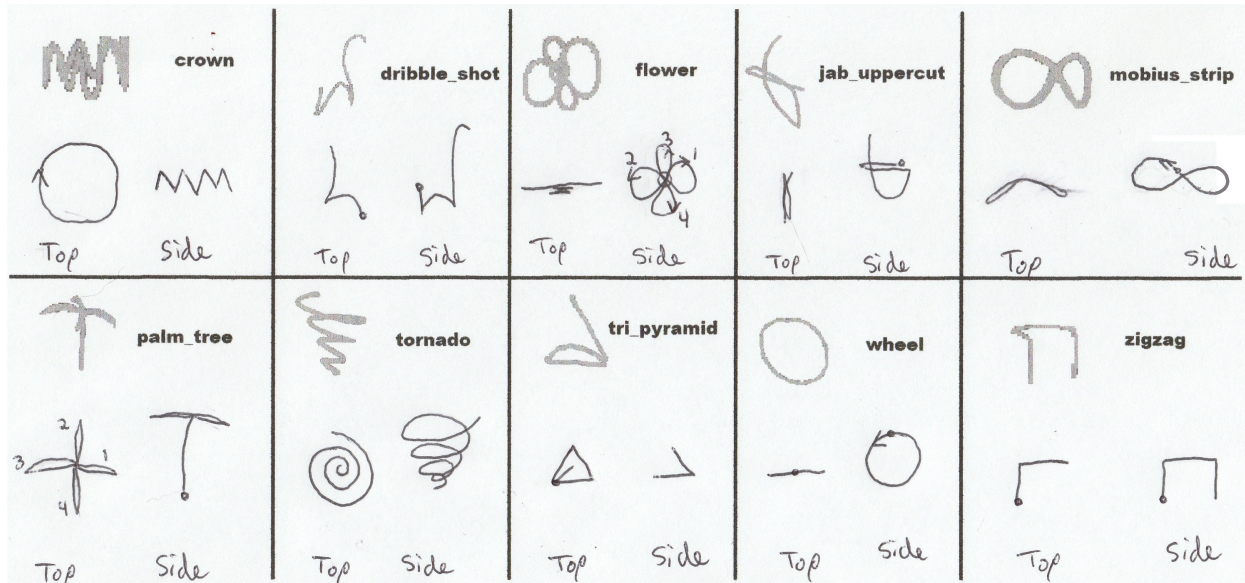
Figure 4.8: The subset of VR gestures presented to users is shown above using symbols as seen in the game (top left), names (top right), symbols top view with directional indicators (bottom left), and symbols side view with directional indicators (bottom right).

The gestures were changed from the 2D version of the game so that the gestures in the VR version would be in 3D space and not just 2D gestures on a 2D plane in 3D space. The gestures for this VR study can be seen in Figure 4.8

The study was counterbalanced in that participants alternated between playing game iterations of gesture names and symbols. The training session was kept in to help the participants understand how the game is played and to familiarize with the gestures before performing the study.

4.4   Hypotheses

As seen with the first study and corroborated by Srivastava et al.'s work [33], we anticipated the results would be the same in this study when it came to how gesture presentation affected difficulty.

Therefore, we hypothesized:

- **VRH1:** *Presenting gestures by name will be significantly more difficult than presenting them as symbols.*

The VR version of the game is quite a bit harder than the 2D version as there are many more factors to it. Due to the added physical difficulty of going from finger movement to full hand/wrist/arm movement for drawing gestures in VR, we anticipated a more significant effect in the difference in difficulty between gesture symbols versus names in VR over that seen in the 2D version of the game. This added physical difficulty coupled with the added symbol complexity moving from 2D space to 3D space (more straight lines, curves, corners, and points [5]), we anticipated that the overall difficulty level would be pushed over the curvilinear hump [11, 18], such that it would reside greater than the skill level of the participants when presented as names. Because of this, we anticipated the upset in balance would cause the level of flow to change from symbols to names. Thus we hypothesize:

- **VRH2:** *Presenting gestures by name will have significantly less flow than presenting them as symbols.*

- **VRH3:** *Game difficulty will be negatively associated with flow.*

In the 2D version of the game, the player is able to see all $360°$ around their character. This makes it easy to see zombies sneaking up on the character with the exception that the player's hand tends to cover part of the screen as they are gesturing and this can sometimes hide an approaching zombie. With the VR version, you get a similar effect as the hand in the way when gesturing as the ink object can obscure zombies in the background behind it. An added difficulty comes in that is a bit more like real life in that you cannot see behind you, or even a good section to your side is

not visible. Because of this, zombies can sneak up on you and unless you pay attention to your healthbar, they can quickly take you down without you noticing. With all the cognitive processes going on with all of the added difficulty aspects, we anticipated that the participants of the VR study would have a more difficult time surviving the zombie apocalypse. Thus, we hypothesized:

- **VRH4:** *Presenting gestures by name will have less survival time than presenting them as symbols.*

Another added challenge comes into play when dealing with the targeting system. In order to maintain the ability of targeting zombies off in the distance and drawing the gestures on top of them, we added a targeting reticle that stays directly in front of the player's view and positions itself on the closest object along that path. Because of this, it can be difficult to draw a gesture and maintain your view on a moving target in the distance. This can cause gestures to miss their target and when rushed by a swarm of zombies every second counts so you do not want to be re-drawing gestures. From this added difficulty, we anticipated:

- **VRH5:** *Presenting gestures by name will have a lower kill count than presenting them as symbols.*

Additional information about how the game was changed from the touch-based gesture drawing game to the virtual reality version can be found in Chapter 5: Testbed.

4.5 User Study Procedure

The second user study took place using a virtual reality game world where the player fights off zombies by drawing the appropriate gesture that matches a particular zombie's symbol. The participant was given a demographics survey to fill out. Upon completion of the demographics survey,

56

they were shown a list of the symbols and gestures in the game and their names. Figure 4.8 shows the exact list they were provided in the study. They were given ample time to get familiar with the list. Once they were ready to proceed, they were instructed on how the game works and were given a trial run of the game to get familiar with it. During the trial run, they were allowed to reference the symbol and gesture sheet as they were likely still getting familiar with the more complex symbols and gestures.

The participant then played one of two iterations of the game, trying to kill as many zombies as possible and have their character survive for as long as possible, only stopping early if they managed to kill all 100 zombies. After playing the assigned iteration of the game, the participant filled out a Game Experience Questionnaire (GEQ) [16] and NASA-TLX [13]. The NASA-TLX and GEQ were chosen because they were the measurement instruments of the first study and we wanted to keep the two studies as similar as possible.

After playing the game intervention, the participant was given a chance to ask any questions they may have had. These questions were answered, then they were thanked and given ten US dollars cash for their time.

4.6   Findings

We continue with our virtual reality findings, followed by the comparative results between the two modalities of touch-based gestures and gestures in virtual reality.

Figure 4.9: The player is shown drawing a gesture that is the same as the shown zombie in an attempt to kill the zombie.

4.6.1 *Dependent Measures*

The NASA-TLX uses a 21-pt Likert scale per question to rank tasks. We coded these results as raw TLX scores by following the provided instructions. The GEQ uses a 5-pt Likert scale per question and the results were coded by following its accompanied instructions. This produces 12 categories from the 17 questions that were asked. The two categories that we were interested in were the level of challenge (difficulty) and flow. In addition to the GEQ, we used a counter in-game to keep

track of the number of kills the player obtained. The game also kept track of the time played, and the timer stopped if the player died in-game. When the player finished playing with each gesture presentation, we wrote down the number of kills and the time that they survived.

4.6.2   *Data Analysis Approach*

We screened our measures for normality using a Shapiro-Wilk test. Since flow was not normally distributed, we used the appropriate Mann-Whitney U Test to test for significance [21]. The flow data for the symbols and the names interventions had roughly the same shape when plotted on a histogram, so we were able to check across medians for the Mann-Whitney U Test.

Our difficulty data was normally distributed, passing five out of the six assumptions for an Independent T-Test. We performed a Levene's Test to determine if the data passed the sixth assumption and it did. Therefore we performed the appropriate Independent T-Test on the difficulty data. When plotting flow versus difficulty, the plot did not come out linear as seen in Figure 4.10, thus a Spearman's Rank Correlation Coefficient Test was used across both interventions to determine if difficulty has a direct correlation to flow.

From our performance data we tested our kill counts and survival time for significance across both interventions. Since they too were not normally distributed, we used the Mann-Whitney U Test for both. Our kill count data had roughly the same shape between interventions but our survival time did not. This meant we were able to use the Mann-Whitney U Test on the medians of our kill count data, but we could only do the test on the means of the survival time data.
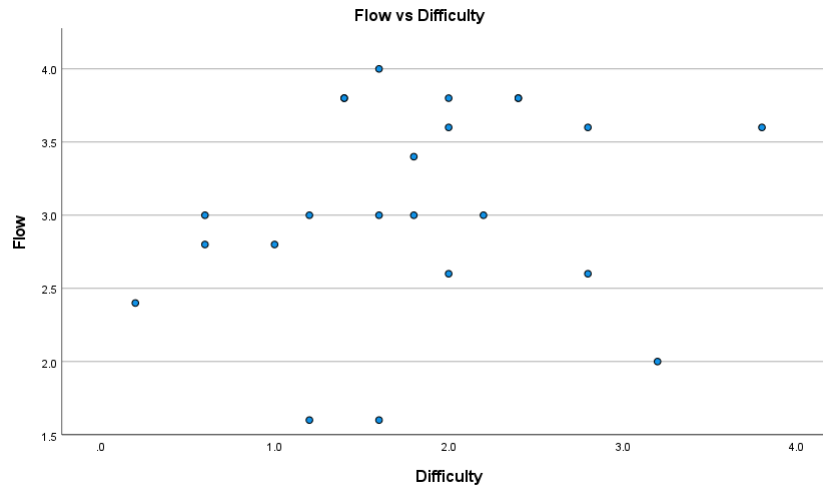
Figure 4.10: This figure shows difficulty plotted vs flow.

4.6.3   Recruitment and Participants

An a priori power analysis was performed using G*Power to determine the sample size needed to get significance for this second study. An effect size of 0.5, $\alpha = 0.05$, power $= 0.8$, was used and the sample size was determined to be 64 participants per intervention for a total of 128 participants. Participants were primarily recruited via an email notification with a sign-up sheet hosted on cal-endly.com. The recruitment email informed potential participants that if they suffer from any sort of balance issues such as vertigo or motion sickness, or if they suffer from seizures, they would not be permitted to participate in the study. Not permitting balance issues helped to mitigate possible added difficulty caused by losing balance from performing wide arm strokes while standing in VR. After reviewing the consent form and obtaining proper consent, the participant was given a chance to have any questions answered before proceeding into the study.

Unfortunately due to the isolation effects of the COVID-19 pandemic, we were unable to obtain

the necessary number of participants within the time that had been allotted. We were able to get a total of 23 participants, all between the ages of 18 and 29. There were 16 participants that identified as males and 7 that identified as females. One participant had their master's degree, one had a bachelor's degree, and the rest had only completed some college at the point which they participated in the study. We had one participant that writes ambidextrously, three that were left handed, and 19 that were right handed. When it comes to touch-based devices, we had 17 participants that use their right hand for touching, 3 that use their left, and 3 that approach it ambidextrously. As for experience, we had 3 participants that rarely play video games, 5 that sometimes play, 10 that frequently play, and 5 that always play. There were 4 participants that have never used a pen-based device, 8 that rarely use them, 4 that sometimes use them, 6 that frequently use them, and one that always uses them.Touch-based devices have much more use with 11 participants that frequently use them and 12 that always use them. Virtual and augmented reality experience had a more even spread across experiences with 6 participants who never use VR or AR devices, 7 that rarely use them, 6 that sometimes use them, and 4 that frequently use them. We had 15 participants that had corrected vision and 8 that did not.

4.6.4   Hypothesis Testing Results

Table 4.1 summarizes the descriptive statistics of our two conditions in relation to difficulty and flow. The minimum difficulty value is higher when it comes to gestures presented as names, but the maximum difficulty value was higher for symbols. As per the mean, median, and mode, we can see that gestures presented as symbols have generally higher difficulty than gestures presented as names but they are generally pretty low, all occurring under 50% of the overall possible maximum value. The mean difficulty value decreases from 1.87 for symbols to 1.74 for names, a 6.95% decrease which is a very minor difference of change. There is not much change in flow from the symbols presentation to the names presentation. They have the similar minimum and maximum

61

Table 4.1: The standard statistics of our dependent variables, difficulty and flow. Each value is out of a possible range from 0-5.

	VR Difficulty and Flow						
DV	*Min*	*Max*	*Range*	*Mean*	*Median*	*Mode*	*StDev*
Difficulty:							
Symbols	0.2	3.8	3.6	1.87	1.8	2.4	0.93
Names	0.6	3.2	2.6	1.74	1.6	2.0	0.82
Overall	0.2	3.8	3.6	1.81	1.8	1.6	0.86
Flow:							
Symbols	1.6	3.8	2.2	3.00	3.2	3.8	0.81
Names	2.0	4.0	2.0	3.15	3.0	3.0	0.60
Overall	1.6	4.0	2.4	3.07	3.0	3.8	0.71

values, resulting in the similar ranges. Their means and medians have minor differences, the mean only increases by 5%.

Table 4.2 summarizes the descriptive statistics of our two conditions in relation to kill count and survival time. Both the minimum kills and the maximum kills obtained occurred within our presentation of symbols with a couple of participants killing all 100 of the zombies. On average, players killed more zombies when presented with symbols than those that were presented with names. The mean number of kills resides fairly close to one-third of the total number of zombies for symbols and close to one-quarter for names. The most common occurrence of kills is higher in the symbols presentation than it is in the names presentation as multiple participants reached the maximum kill count when presented with symbols and no two participants obtained the same kill count when presented with names. Even though the most time survived was obtained from a participant presented with the names intervention, the larger mean survival time was with participants presented with symbols. Interestingly enough, the lowest survival time came from a participant presented with the names intervention.

Table 4.2: The performance of our participants in terms of zombie kills and survival time. This was out of 100 total possible kills and no set max survival time. Time is shown in the format of mm:ss. The N/As represent "Not Applicable" due to no duplicate values within the respective categories.

	VR Performance						
Task	*Min*	*Max*	*Range*	*Mean*	*Median*	*Mode*	*StDev*
Kills:							
Symbols	2	100	98	31.8	19.0	100	34.7
Names	3	81	78	22.3	20.0	N/A	22.3
Overall	2	100	98	27.2	20	100	29.2
Time:							
Symbols	2:41	38:15	36:34	17:46	13:11	N/A	14:31
Names	2:12	45:04	42:52	12:57	08:28	N/A	13:12
Overall	2:12	45:04	42:52	15:28	09:06	N/A	13:48

For VRH1, we assessed the main effect of gesture presentation on difficulty. We found no significant difference between the symbols (M = 1.867, SD = .9277) and names (M = 1.745, SD = .8153) conditions; (t(21) = .331, p = .744), where gesture presentation does not significantly have an effect on difficulty in virtual reality. For VRH2, we assessed the main effect of gesture presentation on flow. We found no significant effect (U = 73, p = .663), where gesture presentation does not significantly have an effect on flow in virtual reality. For VRH3, we assessed the correlation between difficulty and flow. We found no significant effect (r = .230, p = .292), where flow and difficulty do not have a correlation in virtual reality. For VRH4, we assessed the main effect of gesture presentation on survival time. We found no significant effect (U = 54, p = .460), where gesture presentation does not significantly have an effect on survival time. For VRH5, we assessed the main effect of gesture presentation on the players' number of kills. We found no significant effect (U = 58, p = .622), where gesture presentation does not significantly have an effect on the players' number of kills. Table 4.3 summarizes these findings.

Table 4.3: This table shows all VR hypotheses being rejected along with the values used for determination.

VR Hypothesis Results

Hypotheses	Values	Determination
VRH1	(t(21) = .331, p = .744)	Unsupported
VRH2	(U = 73, p = .663)	Unsupported
VRH3	(r = .230, p = .292)	Unsupported
VRH4	(U = 54, p = .460)	Unsupported
VRH5	(U = 58, p = .622)	Unsupported

(**Key:** * < .05, ** < .01, *** < .001)

4.7   Discussion

We designed our study to take in the lessons learned from the 2D study and went with a between-subjects design. Due to the COVID-19 pandemic and its isolation procedures, it took many months before we were able to run the VR study and once we were able, many participants signed up for the study and did not show for it. Because of these issues, we ran out of time to complete the study and did not get our full needed population of users for the study. From the populace that we did see, we did not obtain significance with any of our results. Due to the small sample size, we cannot confirm whether our lack of significance holds true in general, or if it was simply due to the fact we did not have enough participants. To fully understand how gesture presentation affects flow and difficulty in VR, future work will need to be performed.

The most obvious avenue to explore for future work would be to further the virtual reality gesture study. This makes sense since we obtained less than half of our total required number of participants to see significant results as determined by our A priori power analysis. Currently our data for the study is all over the place and it may very well converge toward or away from our hypotheses given enough participants. The only way to determine such factors would be to continue the study

and collect more participants.

CHAPTER 5: TOUCH- AND VIRTUAL REALITY-BASED GESTURE COMPARISON

We now compare our results between touch-based gestures and gestures in virtual reality to better understand if our dependent variables share similar results across both modalities or if they differ from each other.

Ideally we would like to perform a statistical analysis between the two-dimensional study results and the virtual reality study results using a comparative procedure such as Pearson's or Spearman's Rho. Unfortunately, there is not enough VR data to be able to run a proper comparison through IBM's SPSS software. More than half of the two-dimensional data would need to be thrown out to get the sample sizes similar enough to run the comparison. Because of this, we can only perform and report on a direct comparison using statistics that do not require similar sample sizes such as means and standard deviations.

Though we cannot tell from our sample size if the result is significant, we can see from Table 5.1 that the virtual reality game had a higher minimum flow when compared to its two dimensional counterpart. They both had the same maximum flow, but VR had a higher mean, median, and lower standard deviation from the 2D version. This means that on average our participants experienced more flow with the VR game than with the 2D version. This also makes sense with the way VR tends to be more immersive since it blocks a user's peripheral vision from the outside world and instead provides in-game peripheral imagery.

Difficulty follows a similar pattern as can be seen in Table 5.2. The difficulty level for VR had a higher minimum, maximum, mean, median, and standard deviation. This shows that the data is trending toward VR being more difficult than its 2D counterpart. This also makes sense on both

Table 5.1: The standard statistics of flow in both studies. Each value is out of a possible range from 0-5.

	Flow Comparison						
IV	*Min*	*Max*	*Range*	*Mean*	*Median*	*Mode*	*StDev*
Symbols:							
2D	0.4	4.0	3.6	2.45	2.3	4.0	1.11
VR	1.6	3.8	2.2	3.00	3.2	3.8	0.81
Overall	0.4	4.0	3.6	2.60	2.7	3.6	1.06
Names:							
2D	0.4	4.0	3.6	2.52	2.4	4.0	1.12
VR	2.0	4.0	2.0	3.15	3.0	3.0	0.60
Overall	0.4	4.0	3.6	2.68	2.6	4.0	1.04
2D Overall	0.4	4.0	3.6	2.49	2.3	4.0	1.11
VR Overall	1.6	4.0	2.4	3.07	3.0	3.8	0.71
Overall	0.4	4.0	3.6	2.64	2.6	4.0	1.05

a physical and cognitive level as the VR has an additional dimension within which to move about and about which to think. Basically the gestures have to be drawn in an additional direction and the player has to think about this direction as well as physically move within it.

Kill count between the two versions of the game had more scattered results. As can be seen in Table 5.3, the 2D version of the game resulted in a higher minimum kill count, but a lower maximum kill count. This is likely due to the time limit imposed on the 2D study. Many of our participants ran out the clock and would have kept killing zombies in the first study if there was no time limit. There is no telling how high the participants might have gone given infinite time. The mean and median kill counts for VR are roughly half what they are for the 2D version, but their standard deviations are fairly similar and quite large as far as deviations go. This implies that it is likely the VR version of the game is harder than the 2D version of the game.

It is even more difficult to compare times directly as the 2D study had an imposed time limit of

Table 5.2: The standard statistics of difficulty in both studies. Each value is out of a possible range from 0-5.

	Difficulty Comparison						
IV	*Min*	*Max*	*Range*	*Mean*	*Median*	*Mode*	*StDev*
Symbols:							
2D	0.0	3.0	3.0	1.16	1.0	1.2	0.76
VR	0.2	3.8	3.6	1.87	1.8	2.4	0.93
Overall	0.0	3.8	3.8	1.35	1.2	1.2	0.86
Names:							
2D	0.6	3.2	2.6	1.62	1.6	1.8	0.68
VR	0.6	3.2	2.6	1.75	1.6	2.0	0.82
Overall	0.6	3.2	2.6	1.65	1.6	1.6	0.71
2D Overall	0.0	3.2	3.2	1.39	1.3	1.0	0.75
VR Overall	0.2	3.8	3.6	1.81	1.8	1.6	0.75
Overall	0.0	3.8	3.8	1.50	1.6	1.8	0.80

Table 5.3: The standard statistics of kill count in both studies. Each value is out of a possible range from 0-100.

	Kill Count Comparison						
IV	*Min*	*Max*	*Range*	*Mean*	*Median*	*Mode*	*StDev*
Symbols:							
2D	7	81	74	54.03	57.5	60	22.34
VR	2	100	98	31.75	19	100	34.74
Overall	2	100	98	48.22	56	60	27.55
Names:							
2D	5	87	82	47.71	50	50	24.32
VR	3	81	78	22.27	20	N/A	22.26
Overall	3	87	84	41.49	43	50	26.05
2D Overall	5	87	82	50.87	56.5	50	23.39
VR Overall	2	100	98	27.22	20	100	29.19
Overall	2	100	98	44.89	50	7	26.88

Table 5.4: The standard statistics of survival time in both studies in the form of mm:ss.s.

IV	Min	Max	Range	Mean	Median	Mode	StDev
Survival Time Comparison							
Symbols:							
2D	1:54	7:40	5:46	6:49.0	7:40.0	7:40	1:50.7
VR	2:41	39:15	36:34	17:45.8	13:11.5	N/A	14:30.9
Overall	1:54	39:15	37:21	9:40.3	7:40.0	7:40	8:48.6
Names:							
2D	1:16	7:40	6:24	6:55.0	7:40.0	7:40	1:49.4
VR	2:12	45:04	42:52	12:57.2	8:28.0	N/A	13:12.1
Overall	1:16	45:04	43:48	8:23.5	7:40.0	7:40	6:59.9
2D Overall	1:16	7:40	6:24	6:52.0	7:40.0	7:40	1:49.3
VR Overall	2:12	45:04	42:52	15:27.8	9:06.0	N/A	13:48.4
Overall	1:16	45:04	43:48	9:02.3	7:40.0	7:40	7:56.9

7:40 and the VR study had no time limit. Due to this lack of restriction in the second study, the survival times for participants in the VR study are quite a bit higher than those from the 2D study, as can be seen in Table 5.4.

No videos of either study were recorded. Videos of both studies could have been insightful for comparing the two studies, for looking into gesture accuracy. With such videos it would be possible to determine if it was more or less difficult for participants to physically perform 2D gestures or 3D gestures.

CHAPTER 6: CONCLUSIONS AND FUTURE WORK

We designed two zombie survival games to perform two user studies determining the effects of gesture presentation on flow and difficulty when implemented into a game as an element of gameplay. We ran one of our studies as a within-subjects study using a 2D touchscreen game, and the other study as a between-subjects study using a 3D VR version of the same game. We pushed the boundaries of gesture research by testing our gesture presentations as elements of gameplay whereas previous research has been focused on gestures being used as menu navigation techniques.

Game designers should keep in mind that the presentation of their gestures when used as gameplay elements can affect other gameplay elements such as the game's difficulty. Presenting gestures to players in the form of gesture names causes a higher game difficulty than presenting them as symbols of their respective shapes, at least for 2D touch-based gesture games. We suggest further research with a larger user study sample size to determine if this effect holds true in VR games. We did not see significance when it comes to gesture presentation effects on flow. We did however find that as difficulty increased in our 2D environment, so did flow. Previous literature would suggest that this is a curvilinear relationship and we recommend further research using a similar 2D study with less imposed restrictions to obtain a better determination of this relationship. There was no significance found in 2D or VR for the effects of gesture presentation style on the number of zombie kills or survival time.

Our two-dimensional study had a wide range of study population demographics, however both of our studies had a common weight around younger males that have completed some college. It would be interesting to determine from a future study if the same results are obtained with a more even distribution of participants. Another avenue of this could be to do a few studies with different ranges of ages such as an elderly population and a child population. The results could then be

compared to the results of the two studies we performed to determine if age plays a factor in how gesture presentation affects difficulty and flow.

Our in-game clock to which we used as our timer for participants in the 2D study ran off the device's frame rate. Each frame of the game would increase a count on the timer. This was helpful in keeping participants from knowing the true amount of time they spent in the game. However, this required us to make conversions into real time for the data. After determining that our touchscreen game was easy for the majority of participants, we removed the time limit for the second study and obtained longer survival times. The first study could benefit from being re-run with a real time timer as well as the removal of the time limit.

Other gesture presentations might provide additional balance with the skill-to-difficulty ratio and thus cause an affect on flow. One such presentation to compare with might be to use pictorial references that are not the same as the symbol of the gestures. As a direct example, using our zombie game, it could be modified to have each doctor zombie spawn where the delete gesture kills them, and each police zombie is killed by the triangles_chain. The zombies would not have names or symbols above their head because the zombies themselves would be the association needed for the recall of what gesture to use. A few other references might be to use audible sounds that each zombie would make, or use floating pictures above the zombies' heads. These are only a few of the probably many gesture presentations that could be utilized in future studies to determine how they affect difficulty and flow.

We would be excited to see how an expanding gesture set study might result in comparison. As the player continues to kill zombies, and more gestures get added, the difficulty of the game should increase as more gestures to remember should increase the cognitive load required to play the game. The study could be designed to have an infinite number of zombies spawning in the game and add a gesture every time a specified number of zombies is killed, such as five or ten. Each time

the player reaches one of these milestones or tiers, the game gets harder for them. When the player dies, the tier they reached could be noted and compared with the other participants' reached tiers. This would provide a wide variety of difficulties to compare with and determine flow. If playing the names intervention, a tutorial of the new gesture would need to be displayed so the participant will know how to perform the gesture based off the new name they see. In theory, this should allow for better balancing the skill-to-difficulty ratio and provide the curvilinear flow relationship we expected to see from our literature review.

Our two studies primarily focused on one genre of game, survival horror. It would be interesting to know if our findings hold true for other genres such as racing games or strategy puzzle games. There are a great number of varieties of games that exist and many of them employ gestures as gameplay elements. A comparative study across various genres of games could show whether or not these findings are the same in all gesture-based games, or if some variety changes how flow and difficulty are affected by the gesture presentation.

With the easy way the gestures are implemented, the 2D game can be easily altered to help teach the alphabet, other alphabets, numbers, or any other kind of writing such as cursive. In addition, it can be altered in a way to allow for mathematical teaching. Instead of displaying the names of the symbols, equations can be shown such as *3 x 4 = ?*. Then the gesture it is expecting for input would be the answer to the equation (12). Research can be performed on this variant to determine if it aides in learning the times tables or any other mathematical concept. Similarly, the VR version can be utilized to teach physical moves such as karate or dance or some other movement based action.

Multiplayer gameplay using gestures is another natural avenue of research that warrants exploration. Such a study might help determine if there is a greater effect on difficulty and flow. This future work may be split into two avenues, versus and cooperative play. Both of these studies

focused on unistroke gestures. A possible future study could be to utilize multistroke gestures and determine if they have a greater or lesser impact on difficulty and flow.

APPENDIX A: 2D STUDY SUPPORTING DOCUMENTS

Demographics Survey   Participant ID #: _____

Age: _____

Gender: _____

Education Level: _____

Are you prone to one of the following triggerable ailments?

     Vertigo   Seizures   Motion Sickness   None

Which is your Primary Hand for Writing:

     Right-Handed   Left-Handed   Ambidextrous

Which is your Primary Hand for Touching (e.g. a smart phone):

     Right-Handed   Left-Handed   Ambidextrous

How often do you play video games?

Never   Rarely   Sometimes   Frequently   Always

How often do you use **Pen**-**based** electronic interfaces?

Never   Rarely   Sometimes   Frequently   Always

How often do you use **Touch**-**based** electronic interfaces?

Never   Rarely   Sometimes   Frequently   Always

Do you wear **glasses** or **contacts**?   Yes   No

Game Experience Questionnaire

Participant ID #	Task

Please indicate how you felt while playing the game for each of the items, on the following scale:

0 = not at all
1 = slightly
2 = moderately
3 = fairly
4 = extremely

1 I felt content

2 I felt skillful

3 I was interested in the game's story

4 I thought it was fun

5 I was fully occupied with the game

6 I felt happy

7 It gave me a bad mood

8 I thought about other things

9 I found it tiresome

10 I felt competent

11 I thought it was hard

12 It was aesthetically pleasing

13 I forgot everything around me

14 I felt good

15 I was good at it

16 I felt bored

17 I felt successful

18 I felt imaginative

19 I felt that I could explore things

20 I enjoyed it

21 I was fast at reaching the game's targets

22 I felt annoyed

23 I felt pressured

24 I felt irritable

25 I lost track of time

26 I felt challenged

27 I found it impressive

28 I was deeply concentrated in the game

29 I felt frustrated

30 It felt like a rich experience

31 I lost connection with the outside world

32 I felt time pressure

33 I had to put a lot of effort into it

Please indicate how you felt after you finished playing the game for each of the items, on the following scale:

0 = not at all
1 = slightly
2 = moderately
3 = fairly
4 = extremely

1 I felt revived

2 I felt bad

3 I found it hard to get back to reality

4 I felt guilty

5 It felt like a victory

6 I found it a waste of time

7 I felt energized

8 I felt satisfied

9 I felt disoriented

10 I felt exhausted

11 I felt that I could have done more useful things

12 I felt powerful

13 I felt weary

14 I felt regret

15 I felt ashamed

16 I felt proud

17 I had a sense that I had returned from a journey

NASA Task Load Index

Participant ID #	Task

Instructions: For each of the following questions, mark the box that best represents the level of your answer.

Mental Demand: How mentally demanding was the task?

Very Low   Very High

Physical Demand: How physically demanding was the task?

Very Low   Very High

Temporal Demand: How hurried or rushed was the task?

Very Low   Very High

Performance: How successful were you in accomplishing what you were asked to do?

Failure   Perfect

Effort: How hard did you have to work to accomplish your level of performance?

Very Low   Very High

Frustration: How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low   Very High

APPENDIX B: 3D STUDY SUPPORTING DOCUMENTS

Demographics Survey   Participant ID #: _____

Age: _____

Gender: _____

Education Level: _____

Are you prone to one of the following triggerable ailments?

    VR Sickness   Vertigo   Seizures   Motion Sickness   None

Which is your Primary Hand for Writing/Drawing:

    Right-Handed   Left-Handed   Ambidextrous

Which is your Primary Hand for Touching (e.g. a smart phone):

    Right-Handed   Left-Handed   Ambidextrous

How often do you play video games?

Never   Rarely   Sometimes   Frequently   Always

How often do you use **Pen**-**based** electronic interfaces?

Never   Rarely   Sometimes   Frequently   Always

How often do you use **Touch**-**based** electronic interfaces?

Never   Rarely   Sometimes   Frequently   Always

How often do you play VR/AR games?

Never   Rarely   Sometimes   Frequently   Always

Do you wear **glasses** or **contacts**?   Yes   No

Game Experience Questionnaire

Participant ID #	Intervention

Please indicate how you felt while playing the game for each of the items, on the following scale:

0 = not at all
1 = slightly
2 = moderately
3 = fairly
4 = extremely

1 I felt content

2 I felt skillful

3 I was interested in the game's story

4 I thought it was fun

5 I was fully occupied with the game

6 I felt happy

7 It gave me a bad mood

8 I thought about other things

9 I found it tiresome

10 I felt competent

11 I thought it was hard

12 It was aesthetically pleasing

13 I forgot everything around me

14 I felt good

15 I was good at it

16 I felt bored

17 I felt successful

18 I felt imaginative

19 I felt that I could explore things

20 I enjoyed it

21 I was fast at reaching the game's targets

22 I felt annoyed

23 I felt pressured

24 I felt irritable

25 I lost track of time

26 I felt challenged

27 I found it impressive

28 I was deeply concentrated in the game

29 I felt frustrated

30 It felt like a rich experience

31 I lost connection with the outside world

32 I felt time pressure

33 I had to put a lot of effort into it

Please indicate how you felt after you finished playing the game for each of the items, on the following scale:

0 = not at all
1 = slightly
2 = moderately
3 = fairly
4 = extremely

1 I felt revived

2 I felt bad

3 I found it hard to get back to reality

4 I felt guilty

5 It felt like a victory

6 I found it a waste of time

7 I felt energized

8 I felt satisfied

9 I felt disoriented

10 I felt exhausted

11 I felt that I could have done more useful things

12 I felt powerful

13 I felt weary

14 I felt regret

15 I felt ashamed

16 I felt proud

17 I had a sense that I had returned from a journey

NASA Task Load Index

Participant ID #	Intervention

Instructions: For each of the following questions, mark the box that best represents the level of your answer.

Mental Demand: How mentally demanding was the task?

Very Low   Very High

Physical Demand: How physically demanding was the task?

Very Low   Very High

Temporal Demand: How hurried or rushed was the task?

Very Low   Very High

Performance: How successful were you in accomplishing what you were asked to do?

Failure   Perfect

Effort: How hard did you have to work to accomplish your level of performance?

Very Low   Very High

Frustration: How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low   Very High

APPENDIX C: IRB APPROVAL LETTERS

UNIVERSITY OF CENTRAL FLORIDA

<u>APPROVAL</u>

October 31, 2019

Dear Jack Oakley:

On 10/31/2019, the IRB reviewed the following submission:

Type of Review:	Initial Study
Title:	Determining the Effects of Cognitive Memory on Flow in Video Games
Investigator:	Jack Oakley
IRB ID:	STUDY00001072
Funding:	None
Grant ID:	None
IND, IDE, or HDE:	None
Documents Reviewed:	• ConsentForm, Category: Consent Form; • demographics.doc, Category: Survey / Questionnaire; • Email Notification, Category: Recruitment Materials; • Game Experience Questionnaire, Category: Survey / Questionnaire; • Gestures, Category: Test Instruments; • NASA Task Load Index, Category: Survey / Questionnaire; • Pictures of Game, Category: Test Instruments; • Protocol, Category: IRB Protocol;

The IRB approved the protocol from 10/31/2019.

In conducting this protocol, you are required to follow the requirements listed in the Investigator Manual (HRP-103), which can be found by navigating to the IRB Library within the IRB system.

If you have any questions, please contact the UCF IRB at 407-823-2901 or irb@ucf.edu. Please include your project title and IRB number in all correspondence with this office.

Sincerely,

Adrienne Showman
Designated Reviewer

UNIVERSITY OF CENTRAL FLORIDA

APPROVAL

June 8, 2020

Dear Jack Oakley:

On 6/8/2020, the IRB reviewed the following submission:

Type of Review:	Initial Study
Title:	Determining the Effects of Cognitive Memory on Flow in Virtual Reality Video Games
Investigator:	Jack Oakley
IRB ID:	STUDY00001807
Funding:	None
Grant ID:	None
IND, IDE, or HDE:	None
Documents Reviewed:	• Consent, Category: Consent Form; • DemographicsSurvey, Category: Survey / Questionnaire; • EmailNotificationV2.docx, Category: Recruitment Materials; • GEQ.docx, Category: Survey / Questionnaire; • NASA_TLX.docx, Category: Survey / Questionnaire; • Protocol, Category: IRB Protocol; • screenshots.docx, Category: Other; • Script.docx, Category: Interview / Focus Questions; • symbols.jpg, Category: Test Instruments; • VideoOfNamesIntervention, Category: Other; • VideoOfSymbolsIntervention, Category: Other;

The IRB approved the protocol from 6/8/2020.

Due to current COVID-19 restrictions, in-person research is not permitted to begin until you receive further correspondence from the Office of Research stating that the restrictions have been lifted.

In conducting this protocol, you are required to follow the requirements listed in the Investigator Manual (HRP-103), which can be found by navigating to the IRB Library within the IRB system. Guidance on submitting Modifications and a Continuing Review or Administrative Check-in are detailed in the manual. When

you have completed your research, please submit a Study Closure request so that IRB records will be accurate.

If you have any questions, please contact the UCF IRB at 407-823-2901 or irb@ucf.edu. Please include your project title and IRB number in all correspondence with this office.

Sincerely,

Adrienne Showman
Designated Reviewer

LIST OF REFERENCES

[1] Lisa Anthony and Jacob O Wobbrock. 2010. A lightweight multistroke recognizer for user interface prototypes. In *Proceedings of Graphics Interface 2010*. 245–252. `DOI:http://dx.doi.org/10.5555/1839214.1839258`

[2] Benjamin B. Bederson. 2004. Interfaces for staying in the flow. In *Ubiquity (Vol. 2004, Issue September)*. ACM, New York, NY, USA, 1. `DOI:http://dx.doi.org/10.1145/1029383.1074069`

[3] Pierre Benz. 2010. *Gesture-Based Interaction for Games on Multi-touch Devices*. Ph.D. Dissertation. University of Cape Town. `https://openexhibits.org/wp-content/uploads/papers/pierreReport.pdf`

[4] Michael W. Buncher. 2013. *The Effects of Video Game Difficulty Selection on Flow Experience*. Ph.D. Dissertation. Cleveland State University, Cleveland, OH. `https://etd.ohiolink.edu/`

[5] Xiang Cao and Shumin Zhai. 2007. Modeling Human Performance of Pen Stroke Gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. Association for Computing Machinery, New York, NY, USA, 1495–1504. `DOI:http://dx.doi.org/10.1145/1240624.1240850`

[6] Mihaly Csikszentmihalyi. 1990. Flow: The Psychology of Optimal Experience. *Journal of Leisure Research* 24, 1 (1990), 93–94. `DOI:http://dx.doi.org/10.1080/00222216.1992.11969876`

[7] Alastair H. Cummings. 2006. The Evolution of Game Controllers and Control Schemes and their Effect on their games. `https://citeseerx.ist.psu.edu/10.1.1.531.7191`

[8] Pauline Dewan. 2015. Words Versus Pictures: Leveraging the Research on Visual Communication. In *Partnership: the Canadian Journal of Library and Information Practice and Research (Vol. 10, Issue 1)*. Partnership, 1–10. `DOI:http://dx.doi.org/10.21083`

[9] David Duke and Michael Harrison. 2006. *A Theory of Presentations*. Vol. 873. 271–290. `DOI:http://dx.doi.org/10.1007/3-540-58555-9_100`

[10] Cheryl L. Grady, Anthony R. McIntosh, M. Natasha Rajah, and Fergus I. M. Craik. 1998. Neural correlates of the episodic encoding of pictures and words. In *Proceedings of the national Academy of Sciences of the United States of America (PNAS 95)*. PNAS, 2703–2708. `DOI:http://dx.doi.org/10.1073`

[11] Erik M. Gregory. 2008. Understanding Video Gaming's Engagement: Flow and Its Application to Interactive Media. Media Psychology Review. Vol. 1(1). (2008). `http://mprcenter.org/review/gregory-video-game-engagement/`

[12] Craig Harris. 2004. Nintendo DS Specs. (2004). `https://www.ign.com/articles/2004/09/21/nintendo-ds-specs` [Online; accessed 16-April-2020].

[13] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*. Elsevier Science Publishers B.V., North-Holland, 139–183. `DOI:http://dx.doi.org/10.1016/S0166-4115(08)62386-9`

[14] Marc Hassenzahl and Nina Sandweg. 2004. From mental effort to perceived usability: transforming experiences into summary assessments. In *CHI '04 Extended Abstracts on Human Factors in Computing (CHI '04)*. ACM, New York, NY, USA, 1283–1286. `DOI:http://dx.doi.org/10.1145/985921.986044`

[15] H. Hern. 2002. Fun factor for game developers. (2002). `https://www.gamedev.net/articles/game-design/game-design-and-theory/fun-factor-for-game-developers-r1828/`.

[16] W.A. IJsselsteijn, Y.A.W. de Kort, and K. Poels. 2013. *The Game Experience Questionnaire*. Technische Universiteit Eindhoven. `https://research.tue.nl/en/publications/the-game-experience-questionnaire`

[17] Hyo-Min Joo, Hyo-Sun Kim, Hye-Ryeong Kim, and Kwang-Hee Han. 2010. In *Korean Journal of the Science of Emotion and Sensibility (Vol. 13, Issue 1)*. KoreaScience, 129–146. `DOI:http://dx.doi.org/10.1518/hfes.45.4.615.27084`

[18] Martin Klasen, René Weber, Tilo T. J. Kircher, Krystyna A. Mathiak, and Klaus Mathiak. 2011. Neural contributions to flow experience during video game playing. *Social Cognitive and Affective Neuroscience* 7, 4 (05 2011), 485–495. `DOI:http://dx.doi.org/10.1093/scan/nsr021`

[19] Christoph Klimmt, Christopher Blake, Dorothée Hefner, Peter Vorderer, and Christian Roth. 2009. Player Performance, Satisfaction, and Video Game Enjoyment. In *Proceedings of the 8th International Conference on Entertainment Computing (ICEC '09)*. Springer-Verlag Berlin, Heidelberg, Germany, 1–12. `DOI:http://dx.doi.org/10.1007/978-3-642-04052-8_1`

[20] Patrick Langdon, Umesh Persad, and P. John Clarkson. 2010. Developing a model of cognitive interaction for analytical inclusive design evaluation. In *Interacting with Computers (Vol. 22 Issue 6)*. OUP, 510–529. `DOI:http://dx.doi.org/10.1016/j.intcom.2010.08.008`

[21] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. 2017. *Research Methods in Human-Computer Interaction* (2 ed.). Morgan Kaufman, 50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States.

[22] Johnny Chung Lee. 2008. Hacking the Nintendo Wii Remote. In *IEEE Pervasive Computing (Vol. 7, Issue 3)*. IEEE, Bethesda, MD, USA, 39–45. DOI:http://dx.doi.org/10.1109/MPRV.2008.53

[23] Sven Mayer, Valentin Schwind, Huy Viet Le, Dominik Weber, Jonas Vogelsang, Johannes Wolf, and Niels Henze. 2019. Effect of Orientation on Unistroke Touch Gestures. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–9. DOI:http://dx.doi.org/10.1145/3290605.3300928

[24] Merriam-Webster Online. 2020. Merriam-Webster Online Dictionary. (2020). http://www.merriam-webster.com

[25] Meredith Ringel Morris, Jacob O. Wobbrock, and Andrew D. Wilson. 2010. Understanding users' preferences for surface gestures. In *Proceedings of Graphics Interface 2010 (CHI '10)*. ACM, New York, NY, USA, 261–268. DOI:http://dx.doi.org/10.5555/1839214.1839260

[26] Miguel A. Nacenta, Yemliha Kamber, Yizhou Qiang, and Per Ola Kristensson. 2013. Memorability of Pre-designed & User-defined Gesture Sets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1099–1108. DOI:http://dx.doi.org/10.1145/2470654.2466142

[27] Dylan Nordberg. 2017. Let's Talk About Difficulty vs. Fun in Video Games. (2017). https://ruminationonthelake.wordpress.com/2017/11/08/lets-talk-about-difficulty-vs-fun-in-video-games/.

[28] Luke Plunkett. 2010. Report: Here Are Kinect's Technical Specs. (June 2010). `https://kotaku.com/report-here-are-kinects-technical-specs-5576002`.

[29] Dean Rubine. 1991. Specifying Gestures by Example. In *Computer Graphics (Vol. 25, Issue 4)*. SIGGRAPH, New York, NY, 329–337. `DOI:http://dx.doi.org/10.1145/122718.122753`

[30] Henrik Schoenau-Fog. 2011. The Player Engagement Process – An Exploration of Continuation Desire in Digital Games. In *Proceedings of the 2011 DiGRA International Conference: Think Design Play (DiGRA '11)*. `DOI:http://dx.doi.org/11307.06025`

[31] David Sharek and Eric Wiebe. 2014. Measuring Video Game Engagement Through the Cognitive and Affective Dimensions. In *Simulation and Gaming (Vol. 45, Issue 4-5)*. NASAGA, 569–592. `DOI:http://dx.doi.org/10.1177/1046878114554176`

[32] Fernando Winckler Simor, Manoela Rogofski Brum, Jaison Dairon Ebertz Schmidt, Rafael Rieder, and Ana Carolina Bertoletti De Marchi. 2016. Usability Evaluation Methods for Gesture-Based Games: A Systematic Review. In *JMIR Serious Games (PMC5069401)*. PMC, Bethesda, MD, USA. `DOI:http://dx.doi.org/10.2196/games.5860`

[33] Nisheeth Srivastava and Edward Vul. 2017. A Simple Model of Recognition and Recall Memory. In *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*. Curran Associates Inc., Red Hook, NY, USA, 292–300. `DOI:http://dx.doi.org/10.5555/3294771.3294799`

[34] John Sweller. 2010. Element interactivity and intrinsic, extraneous and germane cognitive load. *Educational Psychology Review* (2010). `DOI:http://dx.doi.org/10.1007/s10648-010-9128-5`

[35] Eugene M. Taranta, Corey R. Pittman, Jack P. Oakley, Mykola Maslych, Mehran Maghoumi, and Joseph J. LaViola. 2020. Moving Toward an Ecologically Valid Data Collection Protocol for 2D Gestures In Video Games. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–11. `DOI:http://dx.doi.org/10.1145/3313831.3376417`

[36] Eugene M. Taranta II, Amirreza Samiei, Mehran Maghoumi, Pooya Khaloo, Corey R. Pittman, and Joseph J. LaViola Jr. 2017. Jackknife: A Reliable Recognizer with Few Samples and Many Modalities. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 5850–5861. `DOI: http://dx.doi.org/10.1145/3025453.3026002`

[37] Eugene M. Taranta II, Andrés N. Vargas, and Joseph J. LaViola Jr. 2016. Streamlined and accurate gesture recognition with Penny Pincher. In *Computers and Graphics (Vol. 55, Issue C)*. Pergamon Press, Inc., Elmsford, NY, USA, 130–142. `DOI:http://dx.doi.org/10.1016/j.cag.2015.10.011`

[38] Huawei Tu, Xiangshi Ren, and Shumin Zhai. 2012. A Comparative Evaluation of Finger and Pen Stroke Gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1287–1296. `DOI:http://dx.doi.org/10.1145/2207676.2208584`

[39] WePC. 2019. 2019 Video Game Industry Statistics, Trends Data. (November 2019). `https://www.wepc.com/news/video-game-statistics/`.

[40] Jessica Witvoet. 2013. *Does cognitive load influence performance in a game-based learning task?* Ph.D. Dissertation. University of Twente. `https://essay.utwente.nl/64203/`

[41] Jacob O. Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A. Myers. 2005. Maximizing the guessability of symbolic input. In *CHI '05 Extended Abstracts on Human Fac-*

tors in Computing Systems (CHI '05). ACM, New York, NY, USA, 1869–1872. `DOI:` `http://dx.doi.org/10.1145/1056808.1057043`

[42] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-Defined Gestures for Surface Computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09).* ACM, New York, NY, USA, 1083–1092. `DOI:` `http://dx.doi.org/10.1145/1518701.1518866`

[43] Jacob O. Wobbrock, Andrew D. Wilson, and Yang Li. 2007. Gestures without libraries, toolkits or training: a \$1 recognizer for user interface prototypes. In *Proceedings of the 20th annual ACM symposium on User interface software and technology (CHI '07).* ACM, New York, NY, USA, 159–168. `DOI:http://dx.doi.org/10.1145/1294211.1294238`