Graduate Thesis
15 Credits, Elementary

Visual Attention in Level Design for a 3D Adventure Platform Game
- Analyzing Visual Cues in a 3D Environment

Felix Helgesson
Emil Petersson

Thesis: Bachelor 180 credits
Main field: Computer Science
Program: Game Development
Date of final seminar: 2018-05-31

Supervisor: Arezoo Sarkheyli-Hägele
Examiner: Lars Holmberg
Abstract—Studying the field of visual attention in the context of games can help level designers to make the players feel more immersed and increase the general enjoyment of a game. With the help of visual cues, the designers are able to lead the players through the levels without breaking the aesthetics of the game, thus preventing to break the immersion. This study is presenting a method for analyzing visual cues in a 3D adventure platform game. Gathering data with the help of today’s eye-tracking technology. The results can ultimately be used as a guideline for future work in the field.

Index Terms—Visual Cues, Visual Attention, Eye Tracking, Video Game, Level Design.

I. INTRODUCTION

LEVEL DESIGN is an important part of any video game. R. Kremers describes that level design can: “ [...] make or break a game [...]” (Kremers, 2009). Without a well constructed level, a player can become frustrated and may eventually stop playing. A good level keeps the player interested and involved in the game. Csikszentmihalyi (1990), describes a state which he is calling flow. This state will occur when a player’s skill is matched with the challenge of the game or level. During this state, the player will: Get a sense of active control, lose self-awareness and distortion of the experience of time will occur. Achieving mentioned flow is important to level and game designers. When players are in this “zone”, they feel the most immersed (Kremers, 2009). For players to experience this, the designers must find the right balance in the levels and one challenge is to not make the player feel lost. One way to guide the player is by written hints and tips throughout the game. This may lead players to think that they do not have their own choice to explore the game. However, if you leave the player in the level without any direction or goal, the player can feel lost, anxious and ultimately may stop playing (Baron, 2018). The challenge is therefore, to let the player explore freely, without the feeling of captivity or observation but also to keep the player on track. A way for level designers is to lead the player indirect through the levels is with cues. These cues are described by Moriarty and Sayre (1992) as: “ [...] a signal of something or a reminder of something. It brings to mind something from past knowledge or previous experience that provides a framework of meaning that can be used to interpret the sign.”.

From this, making a object or an area stand out in a game, with a different color, texture or lighting, the player may distinguish this from the environment and use their previous experience to help them navigate in the game. With these cues placed throughout the levels, El-Nasr and Yan (2006), describe that it can decrease frustration and increase engagement. The input from a player is relying heavily on the visual feedback from the game. Therefore, the research will be using cues that are based on visual aspects. This study will define these cues as visual cues.

To direct the focus to these visual cues, the players use their visual attention. In the article “Visual attention: The past 25 years” from 2011, Carrasco (2011), describes visual attention as:

“Each time we open our eyes we are confronted with an overwhelming amount of information. Despite this, we experience a seemingly effortless understanding of our visual world. This requires separating the wheat from the chaff, selecting relevant information out of irrelevant noise. Attention is the key to this process; it is the mechanism that turns looking into seeing.”

There is a large amount of research on the field of visual attention in behavioral and neurophysiological purposes (Carrasco, 2011). Though there is research on visual attention in games (S. Almeida, Mealha & Veloso, 2010, S. Almeida, Veloso, Roque & Mealha, 2011, S. Almeida, Mealha & Veloso, 2016, El-Nasr & Yan, 2006, Lewallen, 2013), there is a gap in gathering and presenting the actual data on visual cues in games. By presenting the data to the field it can prove the importance of visual cues in level design and give support to the designers to see what visual cues are best to use in a specific scenario.

This study will be conducted on a game called Nome, a 3D action-adventure game; a genre that recently had an increase in popularity since the end of the 2000’s (Griffiths, 2018). The game is developed by the authors and other students from Malmö University. The genres are defined as:

- “Adventure: adventure games focus on puzzle solving, and item collection and inventory management.”
- “Platformer - Platform games can be both 2D and 3D games in which jumping or climbing onto platforms on various elevations is a major focus of the game [...] a platform game often features a mascot character jumping (or swinging or bouncing) through challenging platform environments. Shooting and fighting may also be involved.” (Rogers, 2014).

We believe that visual cues are important to adventure platform games due to the fact that they are strongly relying on puzzles and navigation through the levels, which makes visual cues a good practice in how to keep the flow in a game level.

This study presents a method of how to extract data from visual attention with the help of a Tobii Eye-Tracker 4C. With this device you can get data on where the eyes fixates and also the movements in between them (S. Almeida et al., 2016). With this technology we can log how the players direct their attention and discover if some visual cues are better suited for different purposes. In addition interviews will be conducted to complement the data and strengthen the results. They will help us to get a better understanding of how the testers interpret the visual cues. The results will lead to a better understanding on what types of visual cues are best suited in 3D action-adventure level design.

A. Research Questions

By collecting data from tests and interviews, this thesis will try to answer the following questions:

- Which types of visual cues are more prominent in a 3D adventure platform game?

By analyzing if there is a more prominent visual cue with the help from eye-tracker technology, the paper will also try to answer the following question:
• How can eye-tracking data be used when analyzing visual cues in a 3D adventure platform game?

B. Limitations

The research we present is spanned over several different fields, such as: visual attention, level design, game design, game studies. It also touches the subject of perception and cognitive science. Cognitive science is a wide research field with subfields such as: philosophy, psychology, artificial intelligence, neuroscience, etc. (Thagard, 2011). However, this thesis is focusing on presenting data on how effective visual cues are claiming players attention and how well the cues are performing due to their reason in the game. The research in this paper is not asking the questions about why or how a player perceives the visual cue in the way they did it. It would have extended too far over the field of cognitive science, which this study does not have the resources for.

Limitations on the visual cues are also present. The types of cues are limited to lights, texture and a combination of the two. This limitation was made because of the amount of data that is acquired from the tests. To analyze more types of cues would have taken to much of our resources. The game genre chosen for this thesis was chosen because of the importance of navigation and flow in various puzzles. Even if this can be important to many types of games we believe that research conducted on this genre, ultimately contributes more to the field. Testers was limited to 10 persons; the testers were hand-picked as a group with varying experience in this game genre. The amount of persons testing provided the research with sufficient data, though it could be scaled up.

II. Previous Work

A. Level design

Kremers (2009), summarizes level design and level designers as: “[…] level designers are responsible for the implementation of the game and the levels are where the game takes place.”

Rogers (2014), makes an analogy to level design by comparing it to a theme park: “I found that theme parks are designed to move guests from one adventure to the next in the most effective way possible, much like a well-designed game level.” He continues in the same chapter that in a theme park, landmarks are a way to lead the guest between different locations (Rogers, 2014). In this study visual cues and landmarks share the same purpose; to grab the player’s attention to a object or an area.

Though level design is important in game development, there is an under-representation of scientific research in the field. In 2010, D. Miliam and M.S El Nasr presented a model for evaluating and understanding good level design. They presented a set of level design patterns for guiding the player through the levels. The patterns was based on interviews with peoples from the game industry and six play sessions. One pattern they found was the Path Target pattern which describes visual targets in a game level that can be translated as visual cues. The following is their definition:

“Orients and directs player movement or attention towards visible targets in the level. This behavior reinforces vertical or horizontal scanning of an area to apprehend a target. This could be a visible landmark to attract movement or a means to orient in the direction of important entities using a targeting device, for instance a camera or weapon.”(Milam & El Nasr, 2010).

This definition was verified by the various developers in the study presented the authors. Proving that the use of visual cues is a practice in game developing. Making research regarding visual attention and visual cues an important asset in the future of level design.

B. Eye-Tracking in video games

Eye tracking is the process of measuring either the point of gaze or the motion of an eye relative to the head. Early analysis of eye movements date back to 1879 when Louis Émile Javal conducted a study where he suggested that the human eye moved by a series of jerks (Richardson & Spivey, 2004). That study was done by naked eye observation and the measurement of eye movements did not become valid until hardware devices, that could record eye movements, appeared. Since then great advances have been made and the technology available to researchers today is far more advanced and accurate. It is used in several areas of research such as advertising, television, the web and video games, for both interaction and analytic purposes (‘Eye tracking for research’, 2018).

Tobi is at the forefront when its comes to interactivity for video games and the eye-tracker 4C supports, at the time, more than one hundred games (‘100+ Games Powered with Eye Tracking’, 2018). There are other alternatives to the Tobii eye-tracker 4C, one of them is a head mounted tracker created by S. Jalaliniya and Malmö University (IoTaP, 2017). However, during tests with this device, it showed that the eye-tracker needed to be perfectly still for the data to be valid. Which makes the method of data-collection problematic because it would take too much time to control the movement of the tester.

Studies have been done exploring the use as an evaluation tool for video game levels and usability. M. Seif El-Nasr and Su Yan conducted early research on visual attention in 3D with the help of eye-tracking. They presented a analysis method for analyzing eye-tracking data for two types of games with support of two different search patterns (El-Nasr & Yan, 2006). More on their findings will be explained in 2.3.

A more recent study by S. Almeida et al. (2016), presents a method to measure player interactions in game environments, the constructing of their method started in 2009 (S. d. J. Almeida, 2009). In its first iteration the intention was to understand how players interacted with different game scenarios and three groups attended the study (inexperienced, casual and hardcore players). In another study from S. Almeida et al. (2010), this method was put to use to analyze the differences between hardcore and inexperienced players’ interaction behaviours within the game Call of Duty: Modern Warfare (this game is used in all studies surrounding the method).
C. Visual Attention

Visual attention is often described with two visual search mechanisms, bottom-up and top-down or stimulus-driven and goal-directed (Egeth & Yantis, 1997; Connor, Eggeth & Yantis, 2004). The first of the two is handling the raw sensory input, making us quickly direct our visual attention. In contrary, the top-down mechanism is used for directing our visual attention, depending on the cognitive strategies (Connor et al., 2004). The article describes that the bottom-up mechanism is used for quickly spotting sudden movement or change in color without consciously recognizing it. The top-down mechanism is for searching more thoroughly for things that we need. For example, if we are hungry we search for a shift in color in a green environment that could be a fruit (Connor et al., 2004). The two visual search patterns were further researched in an article from Magy Seif El-Nasr and Su Yan in 2006. As described earlier, they presented an early method for implementing an eye-tracker for visualizing the data from the play-tests. Their goal was to document if visual search patterns were present during playing games and which pattern was the most dominant in 3D video games. Their hypothesis was that the top-down search pattern was going to be the most dominant, due to the fact that games are a highly goal-oriented medium (El-Nasr & Yan, 2006). They presented evidence that showed that both search patterns were widely used. One example from where the top-down search pattern was prominent, was during a play-test with the game “Legacy of Kain: Blood Omen II”. In one segment of the game, the player was trapped and had to find a way out. The players attention and focus was on the doors, but none of them were unlocked. The designer had made a bright wall, where the exit was, but because the top-down search pattern is used during a specific goal, the players did not notice the bright wall. However during a segment when the players did not have a specific goal in mind, the players clearly noticed the visual cue when it appeared. In this case it was an object that quickly turned red (El-Nasr & Yan, 2006). In an article from 2012, G. Healey and J. Enns discusses a field within visual attention called preattentive processing. This field is covering the low level visual perception in humans, how we direct our attention to objects with unique visual properties, for example; colors that stand out from the environment, change in form, flickering, etc. (Healey & Enns, 2012).

D. Summary of previous studies

The research on visual attention in games has been done since El-Nasr and Yan presented their findings from the the 2006 study of visual attention in 3D games. The authors of the articles present the visualizations of the results, in the forms of heat maps, diagrams and screenshots from the game (El-Nasr & Yan, 2006) S. d. J. Almeida, 2009; S. Almeida et al., 2010). In the research on preattentive processing, all of the research was conducted on 2D imagery, like pictures and film (Healey & Enns, 2012), with no research on games or other interactive media. However, we think that the principles can be used in a 3D environment, since the methods that direct our attention can be used as visual cues in the game in this research.

Figure 1. Figure from the article by Healey and Enns (2012), showing a change in color of an object in a environment of similar visual properties.

III. Study Context

A. The Game

The research will as mentioned be conducted on a game called Nome, developed by the authors and four other students in 2017. It is a 3D adventure platform game where the player will have to traverse platforms and solve environmental puzzles in order to progress through the level. Due to the fact that the game mechanics already have been developed, made it easier to begin designing the levels. The second reason for choosing Nome is its game play. It doesn’t require fast reflexes or quick decision making and there are no dangers that the players need to avoid. This makes it more accessible for a wider number of players since they won’t rely on fast reflexes or skill.

B. Visual cues

In order to make a distinction between our constructed visual cues in the level we defined two sub-categories, light and texture cues. The cues were categorized out of the visual features in the compilation of the preattentive processing features by Healey and Enns (2012). These visual features are based on luminance, hue, direction and the visual cues created for this study are made to use as many of these as possible. The design of the visual cues was also determined by Unity, the engine that the game is developed on (unity3d.com, 2018). The game engine has a built in function for texturing and coloring objects, which makes visual cues easy to create and implement. The engine also has a lighting system with different light sources. As spotlights, ambient lights and directional lights.

These categories are done to make a more efficient analysis since it gives us the possibility to determine which visual cue is most effective for it’s given purpose. All intended visual cues in our level will follow these definitions, and they also need to be placed in connection to a puzzle or an area of interest. The two categories are defined as following:

- Light cue: a visual cue that uses an in-game light source or particle effect to gather the player attention. An example for this would be a light that turns on after the player performs an action signaling that something happened.
Texture cue: a visual cue that uses a texture pattern or color that makes it stand out more in the environment. An example for this is to give objects that the player can interact with a different color/texture than other similar objects.

<table>
<thead>
<tr>
<th>Name</th>
<th>Cue</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trees</td>
<td>Texture</td>
<td>1.1</td>
</tr>
<tr>
<td>Red light window</td>
<td>Light</td>
<td>1.2</td>
</tr>
<tr>
<td>Generator Light</td>
<td>Light</td>
<td>1.3</td>
</tr>
<tr>
<td>Generator Lever</td>
<td>Texture</td>
<td>1.4</td>
</tr>
<tr>
<td>Door</td>
<td>Texture</td>
<td>1.5</td>
</tr>
<tr>
<td>Door Lights</td>
<td>Light</td>
<td>1.6</td>
</tr>
<tr>
<td>Doors on drawers</td>
<td>Texture</td>
<td>2.1</td>
</tr>
<tr>
<td>Mug on Floor</td>
<td>Texture</td>
<td>2.2</td>
</tr>
<tr>
<td>Plates on Bench</td>
<td>Texture</td>
<td>2.3</td>
</tr>
<tr>
<td>Plates and Mugs on Table</td>
<td>Texture</td>
<td>2.4</td>
</tr>
<tr>
<td>Computer screen</td>
<td>Texture</td>
<td>3.1</td>
</tr>
<tr>
<td>Sparks under Desk</td>
<td>Texture/Light</td>
<td>3.2</td>
</tr>
<tr>
<td>Sparks from Powercell</td>
<td>Texture/Light</td>
<td>3.3</td>
</tr>
<tr>
<td>Key 1</td>
<td>Light</td>
<td>3.4</td>
</tr>
<tr>
<td>Drawers</td>
<td>Texture</td>
<td>3.5</td>
</tr>
<tr>
<td>Computer Screen Change</td>
<td>Texture</td>
<td>3.6</td>
</tr>
<tr>
<td>Key 2</td>
<td>Light</td>
<td>3.7</td>
</tr>
</tbody>
</table>

Table 1: Table of visual cues that will be measured (new section divided by ———— means new part of the level).

The reason for choosing these cues are because of their real world implementation and to not break the visual aesthetics of the game. By interacting with everyday objects in-game, immersion will be enhanced (Peniche, 2016), which may help players stay in the state of flow.

C. The Level

The level created for this research is of linear design, which Kremers describe as “[...]where the gameplay events follow a strict line for the player to follow.” (Kremers, 2009). By using a linear design it becomes less cumbersome to analyse the data we gathered from the tests.

The level in this study is divided into three sections (outdoor, kitchen and garage) where each one contains one or two puzzles to solve. They become more complex as the player progress and the visual cues are placed in relation to a puzzle or hinting on where the player needs to go, or what action the player needs to perform. What follows is a walk-through of the level explaining the puzzles and where the visual cues were placed.

Outdoor: As the game starts the player is presented with the view displayed in figure 2.

Main objective is to get inside the house and the player needs to locate the main entrance. Next to the entrance is a power-generator that needs to be activated for the door to function. Next the player needs to push one of the chairs in front of the door for it to open. The player will then enter the kitchen section.

The main visual cues in this scene is the red light in the window (1.2), which stands out from the rest of the environment. This in combination with the sun’s position, the yellow leaf trees (1.1) and the skyscrapers are used to create a focal point in the level. The use of a focal point in a level can be for aesthetics or functional gameplay, A. Galuzin describes this: “[...]as to draw attention for the player to examine a location, an item, navigate towards an objective and to help the player orient themselves in the environment.” (Galuzin, 2016). The intention is to draw the player’s attention since the entrance to the house is located underneath where a door can be clearly seen (1.5).

Next visual cue is the light on the power-generator (1.3) and its yellow lever (1.4) see figure 3. Intention of these are to first get the players attention and trigger the player to trying to reach the lever. As it turns on, the house will have power and a light will appear in front of the door (1.6), this is another visual cue that changes depending on player or AI interactions. Color of the light can change between yellow (see figure 4), red (see figure 5) and green (see figure 6) giving clear indications to the player about the doors status.
Kitchen: In this section there’s one puzzle for the player to solve which is how to open the next door. In order to do this the player has to scale the kitchen counter and then jump to the closest chair and onto the table. Plates and cups are placed on the table and the player has to at least push one off the edge, this will activate the vacuum cleaner which will patrol around the house opening doors as it goes. Following the vacuum cleaner the player can enter the hallway and garage. Visual cues in this section are the yellow texture of the cups and plates(2.3 and 2.4) see figure[7]. The ones on the table and kitchen bench are stacked on top of each other making them easy for the player to push over. A mug will be located at the base of the table as the player enters suggesting to the player that the other ones on the table and bench also can be pushed down. The second visual cue is the kitchen drawers(2.1), they are pulled out in a stair-like fashion making it easy for the player to climb up. Texturing of the drawer’s knobs is yellow making them stand out more in the scene.

Garage: The garage is the final section and it contains two puzzles with the goal to collect two keys (the player will get notified about this), players will be notified about this goal as they enter the section(see figure[8]). If the player collects them both the game is won. The first key is located under a machine and it has to be turned on in order to get the key (see figure[9]). First step of the puzzle is to locate the missing power-cell. Once found the player needs to drag it to its socket, this will turn on a computer located on top of the desk. To reach it the player must climb up on the workbench by dragging out a drawer and climb up the workbench. On top of the workbench is a roomba(a sweeper) that will charge towards the player attempting to push him down. By placing a patch of ice the player can make it slip off the bench or avoid it with well timed jumps. Between the bench and the desk is a tipped over drawer-unit and the player can platform his way on to the desk by jumping on it. On the desk the player can activate the computer which turns on the machine that blocks the key. The other puzzle involves the roomba that needs to slip off the table in order to get the second key. This is done by placing ice patches as it charges at the player. Once it’s down the vacuum cleaner will try to suck it up resulting in it breaking and spitting out the other key.
Visual cues first visible to the player as he enters the section is the empty power-cell socket(3.2) below the desk and the computer on top. The power-cells all have a bright yellow lightning symbol and sparks emits from the empty slot and the missing power-cell(3.3), hinting at the player that they belong together. The computer displays a message saying “no power” in bright red colors (3.1). This texture changes once the power-cell is connected to “press to activate cutter”(see figure 10), in bright green color (3.6). Handles of the workbench drawers that the player can interact with have a bright yellow texture (3.5), compared to the black texture of the other drawers. Both keys have a bright glow making them hard for the player to miss (3.4 and 3.7).

Figure 9. Location of key 1(3.4).

Figure 10. On top of the desk with texture cue 3.6.

IV. METHOD

What follows is a description of our research method, what data was collected, how it was collected and how it was analyzed.

A. Design Science

Our research process will be based on the Design Science (DS) research methodology (Peffers, Tuunanen, Rothenberger & Chatterjee, 2007). This methodology has a process that includes six steps, that are in turn use to evaluate IT artefacts with the purpose of solving a specific amount of identified problems (Von Alan, March, Park & Ram, 2004). This process involves, designing these artifacts, evaluate its design, and communicating the results to appropriate audiences (Peffers et al., 2007).

The six steps in more details of DS process are:

1) Problem Identification and Motivation: In the first step a specific research problem should be defined along with justification of its solution (Peffers et al., 2007). The justification accomplishes motivating the researcher as well as the audience to pursue the problem, and understanding the researchers reasoning associated with the problem (Peffers et al., 2007).

2) Objectives of the Solution: The second step involves rationally acquire the objectives of a solution from the defined problem and knowledge of what is feasible(Peffers et al., 2007). The objectives can be either quantitative or qualitative.

3) Design and Development: In this step the artifact is developed, this includes determining its desired functionality, architecture and finally creating the actual artifact(Peffers et al., 2007). Artefacts are categorize and defined by (Von Alan et al., 2004) as: constructs (vocabulary and symbols), models (abstractions and representations), methods (algorithms and practices), and instantiations (implemented and prototype systems).

4) Demonstration: In this step the artefacts ability to solve one or more instances of the problem is demonstrated, using an appropriate activity, e.g., case study, experiment or simulation (Peffers et al., 2007).

5) Evaluation: At this point relevant metrics and analysis techniques are used to compare the solution objectives (step 2) to actual observed results from the artefacts in the demonstration (step 4) (Peffers et al., 2007).

6) Communication: In the final step, the problem, its motivation, as well as the artefact itself and its effectiveness, is communicated to a relevant audience, such as researchers or practicing professionals (Peffers et al., 2007).

B. Applying the method in our study

1) Problem Identification and Motivation: Getting the player to understand where to go or which gameobject is intractable without telling the player explicitly is can be a challenge in level design. In this study we will use a set of defined visual cues in our game called Nome to analyze the players visual attention and how they can assist the player when orienting a level and solving puzzles. By using interviews and eye-tracking technology we will be able to determine how effective each visual cue was at getting the players attention. Our findings will benefit other games in the genre of 3D adventure platformers.

2) Objectives of the Solution: We conducted a literature research on the subject of level design, visual cues, eye tracking and visual attention in the context of video games. This gave us a better understanding about how we could approach a solution to our problem. A more detailed description on how the research was conducted can be found in section 4.3.

3) Design and Development: In this step we defined two categories for visual cues and constructed the level. Placement of our selected visual cues was determined.
by obstacles in the level and each cue needed a specific purpose in assisting the player to overcome that obstacle. Also in this step we familiarized us with the eye tracking device and came up with a solution to use it during a gameplay session.

4) Demonstration: The demonstration of the artifact was done by presenting it to testers. Each session was then followed by an interview with the goal of obtaining information about which visual cues he/she noticed in the game and if they ever helped the tester to progress. See section 4.5 for a detailed description of our data collection approach.

5) Evaluation: The metrics from the play sessions lays the groundwork for our evaluation, giving us an insight into which visual cues was the most prominent. The data from the interviews is backed by the gaze data and giving us the possibility to analyze the duration of the initial fixation and how many fixations it received.

6) Communication: In the last step, we documented the solution as in this study to share the acquired knowledge.

C. Literature Research

In order to get a better understanding on how to conduct the research, we had to examine earlier studies regarding the areas of level design and visual attention in the context of video games. We also research how previous work used eye tracking technology as an analysis tool for video games to get an idea how we could implement it in our study. We used Google scholars database (google.com, 2018), to find relevant information using keywords, “visual cues”, “level design”, “visual attention” and “eye tracking” in combination with “games” and “game analysis”. Books and websites also granted us valid information. We had to use the snowball sampling methodology (Explorable.com, 2009) to find more related work since our searches in the Google scholar were not as plentiful as we hoped.

D. Eye tracking

As presented before, the research will be conducted with the help of a Tobii eye tracker 4C device.

What follows is an overview of the different eye movements the human eye can perform, how an eye-tracker records this data and which one of the types our study will be measuring. It will not go into depth on how the algorithms for gaze data filtration works since its not relevant to this study.

By using eye-tracking, researchers can register the five different movements the human eye is capable of performing: (1) ‘smooth pursuits’, (2) ‘vergence movements’, (3) ‘vestibular movements’, (4) ‘saccades’ and (5) ‘fixations’:

(1): Smooth pursuit is the result of a mechanic that in the present of a moving object, respond with a corresponding course of movement. This allows vision to stay fixed on the object (Hall, 2015). This movement can’t be triggered voluntarily, in absence of a moving target.

(2): Vergence movements are as they move in opposite directions disjunctive. If a person is looking from an object at a greater distance to one placed closer, the eyes will rotate toward the nose (converge). Per contra, moving from a object placed closed by to one further away, the eyes will instead rotate towards the ears (diverge).

(3): Vestibular movements, also known as the vestibular-ocular reflex. These movements focuses the retinal image while the head is in movement. This is feasible through the counter-rotation of the eyes at the same velocity the head moves in the opposite direction (Wong, 2008).

(4): Saccades are defined as voluntary, rapid and reflexive eye movements. These are used to reposition the fovea (an area of the retina responsible for sharp vision) to a new location in the visual field.

(5): Fixations are responsible for the ability to fix eye gaze on a specific object in the visual field. Fixations are controlled by two neuronal mechanisms: (1) the voluntary fixation mechanism, which allows humans to voluntarily find the object on which they want to fix their vision; (2) the involuntary fixation mechanism, which holds the eye on the object once it has been found (Hall, 2015).

For this study, fixations will be the eye movement to analyze since they can reveal useful information about attention. Tobii website states “[...an increase in the time taken to make a first fixation on a target suggests a decrease in the salience or visual attractive power of that feature. An increase in average fixation duration on a target or area could signal greater effort required to make sense of something or could suggest that what is looked at is more engaging.” ‘Types of eye movement’, 2018. This is important to this study since the research is focus on what draws the players attention in a game level. What follows is an explanation how the eye-tracker gathers data and how it then gets processed.

As an eye-tracker records eye movements, it creates gaze points which are the instantaneous spatial locations of the visual axis landing on the stimulus. They have an (x, y) coordinate and a timestamp that corresponds to its measurement, in our case the screen. This is the raw data gathered and the Tobii eye-tracker 4C used in this study will report the gaze points 11 ms apart. However gaze points themselves are not useful for research purposes, but they are used to determine fixations. Fixations are constructed by using algorithms that translates the gaze points into an associated sequence of fixations. Fixations are real in the sense that they are meaningful episodes of looking that our visual system generates. These episodes have clear-cut dynamic characteristics that gaze point to fixation conversion algorithm or fixation filter, is designed to model. The goal is to reconstruct these significant eye movements as truly as possible. Visualisation of recorded gaze data will be done with gaze plots which Tobii Pro Lab is capable of. Gaze plots are described in S Almeida et. al (2016) as:

“A ‘Gaze Plot’ summarizes eye behavior and displays fixations and scan paths. Furthermore, it indicates the sequence and order of an individual’s eye movements. A ‘Gaze Plot’ uses circles and lines to represent data. Circles are used to represent fixations. The larger the size of the circle, the longer the duration of the fixation. The lines that connect these circles (fixations) represent scan paths (saccadic movements), (rapid) movements that occur between fixations. Normally, gaze plot
representations will include in their circles numbers that indicate the chronological order in which the eye movements and fixations occurred.

V. DATA COLLECTION

This section will present the method of the data collection. How the tests was conducted and how the data was gathered. The data was collected from a laptop with an additional screen connected to it. The Tobii eye-tracker 4C was placed under the second screen and was calibrated for every new tester to get the most accurate results. A game controller was placed at the second screen for the tester to control the game.

![Figure 11. (1) Laptop, (2) Second screen, (3) Tobii eye-tracker 4C, (4) Game controller.](image)

Each tester will play through the level while the following data is collected:

1. Gaze data from the Tobii eye-tracker 4C using Tobii Pro Studio with timestamp.
2. Gameplay footage from the rendered gaze plots.
3. Video export from the Tobii Pro Lab software.

The Tobii Pro Lab software is designed to be used for image analysis due to its supported file formats. In order to use it for video games we had come up with a own solution. We used a dual monitor solution where we collect two video exports. One from the game-session and one from the gaze data recording. The two video exports will be recorded at the same time on two different screens.

The two recordings are then be merged into one video file with the free video editing software, DaVinci Resolve (Blackmagicsdesign.com, 2018). The gaze plots video will be put on top of the game session video with a overlay blend mode. This will make the gaze plots visible on top of the game video export and makes it easier to analyze. To synchronize the videos, the tester looked in the top corners from left to right, counted to one second and the made the avatar in the game jump. This made it easy to see where the two recordings synchronized with each other. Sessions are overseen by a supervisor taking notes or assisting if any problems should occur and will take approximately 10-15 minutes to complete.

The tester will then be interviewed and asked a series of question, the first ones are designed to get an idea of their gaming habits and an estimation about how much time they spend on video games. The remaining questions are specified for each section of the level and aims to get data on which ones of the visual cues they noticed. The opening question for each of the section is about what drew their attention as soon as they entered it. This is designed to get a understanding of the testers first impressions of the section and what stood out to them. The next question(s) revolve around the puzzles in the section how the player solved these. These questions will be designed to get the tester to explain what they saw surrounding the puzzles and if anything in particular assisted them in solving them. The final questions will ask the tester if they can specify an object or visual element they found most prominent in the section and if he/she ever felt lost in the section and if the case why. The data gained from these are self explanatory. During each interview the interviewer will note each visual cue mentioned by the player, as well as the once not mentioned. This will be the main data collected from the interview along with the data on how many of the testers felt lost.

There is a possibility that some of the visual cues might be forgotten by the tester due its design or placement in the level. To counter this we could have used a multichoice questionnaire where the tester could select the visual cues they remembered from each section. The reason for not using a questioner is the risk that the options presented might help the player to remember each and every visual cues making it rather difficult to determine which one was more prominent. By instead doing an interview directly after a play session the tester should remember the once most salient in each section, since each session only lasts for 10-15 minutes. In other words, the chosen method of using an interview was seen as more preferable as it gave us the opportunity to gain a better understanding of the collected data without potentially tainting the testers’ answers if e.g. a questionnaire was chosen.

A manual video analysis will then be done on the gameplay footage recording with gaze plot overlay to gather data on initial fixation and number of fixations each visual cue receives. Due to the large amount of data collected by the eye-tracker some of the fixations had to be discarded. The once not mentioned were the fixations on an visual cue that the tester interacted with or if the visual cue took up the whole screen space. The timestamp on the video will then be compared with an exported files timestamp from Tobii Pro Lab to determine the initial fixations duration in milliseconds. The Tobii Lab software also gives data on what type of eye movement it recorded ensuring that it indeed was a fixation and not a saccade.

A. Data Analysis

The data collected from our testing sessions will be evaluated by a statistical analysis. Visual cues for each level section will be evaluated based of the answers the testers gave during the interview and the gaze data collected with the eye-tracker. When analyzing the answers from the interview questions we will look at how many of the testers mentioned a visual cue and if they felt lost in a particular section of the level. If a visual cue is mentioned it will be interpreted that the tester acknowledge the visual cue. The number of mentions will then be compared to the gaze data in terms of which visual cues the
testers fixated on, the average number of times the visual cue received a fixation and the duration of the initial fixation to see if there is any relationship between the interview answers and the gaze data. We will also analyze the reasons for why some testers felt lost in the level section and if this has anything to do with the visual cues.

VI. RESULTS

In this section of the paper, we will present the results from the data collection. It will be visualized with the help of tables and diagrams to get a better understanding of the data collected.

All participants in the study had experience playing 3D video games with an average playtime at 9.2 hours per week. The tests were done in a lab environment and eight of the tests were completed with no issues. Two tests needed to be remade due to bugs in the game. The two testers who ran into these problems had to restart the game and continue from the point in the level before the bug. Other minor problems did occur during a couple of the other tests as well, but those did not interfere with the session and could be sorted out after. The results will be presented in the order of the section of the levels and the visual cues will be presented with the following metrics:

**metric i.** The number of testers that confirmed that the noticed the visual cue in the interview.

**metric ii.** The number of testers that fixated on the visual cue according to the gaze data.

**metric iii.** The average number of times a visual cue received a fixation among the ii testers.

**metric iv.** The average duration of the initial fixation the visual cue received from all the testers.

A. Outdoor section

<table>
<thead>
<tr>
<th>The cues for the outdoor section:</th>
<th>Name</th>
<th>Cue</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trees</td>
<td>Texture</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>Red light window</td>
<td>Light</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>Generator Light</td>
<td>Light</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>Generator Lever</td>
<td>Texture</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td>Door</td>
<td>Texture</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>Door Lights</td>
<td>Light</td>
<td>1.6</td>
<td></td>
</tr>
</tbody>
</table>

1.1: None of the testers recalled the yellow trees surrounding the main house and only half of the testers fixated on them and within that group the average number of times the visual cue received a fixation was 2.6 times. The average initial fixation was 370 ms.

1.2: The red light in the window gathered more of the players attention with a total of six testers noticing it, eight of the testers fixated on it and in that group the visual cue received an average of 3.6 fixations. The average initial fixation was 339 ms. 1.3 and 1.4: The generator were recalled by six of the tester due to its light and three testers also recalled
its yellow lever. The eye tracking data shows that everyone in the test group fixated on them with an average of 3.2 times each. Visual cue 1.3 average initial fixation was 678 ms and visual cue 1.4 average initial fixation was 440 ms.

1.5 and 1.6: Only one of the testers mentioned the door before it was illuminated but once illuminated seven of the testers recalled it. The eye tracking data shows that both visual cues received fixations by all the participants but the door with the light turned off received more fixations with an average of 3.6 times while the door with the light on received an average of 3.5 times. Visual cue 1.5 average initial fixation was 368 ms and visual cue 1.6 average initial fixation was 259 ms.

1) Summary of outdoor section: When asked if the tester understood what to do in this section seven said no and they had to be assisted during the test (see figure 15). However this was mainly due to the design of the first puzzle which the testers found not to be particularly intuitive or clear. The testers who played the first section any issues were also critical over the puzzle and said it was just by chance they solved it. The metrics from this section shows that visual cue 1.3 received the highest average in initial fixation time among the testers, but visual cue 1.6 received more mentions in the interview. These cues along with 1.4 and 1.5 all received fixations from all the testers which needs to be taken into consideration since they all needed to be interacted with in some way for the tester to progress. Cue 1.2 and 1.5 received highest average in fixation counts.

B. Kitchen section

<table>
<thead>
<tr>
<th>Name</th>
<th>Cue</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knobs on drawers</td>
<td>Texture</td>
<td>2.1</td>
</tr>
<tr>
<td>Mug on floor</td>
<td>Texture</td>
<td>2.2</td>
</tr>
<tr>
<td>Plates on Bench</td>
<td>Texture</td>
<td>2.3</td>
</tr>
<tr>
<td>Plates and Mugs on Table</td>
<td>Texture</td>
<td>2.4</td>
</tr>
</tbody>
</table>

2.1: The knobs on the drawers were not mentioned by any testers even though seven of them fixated on them with an average of 1.8 times. The average initial fixation was 348 ms.

2.2: The mug on the floor was mentioned four times by the testers and it received fixations six times with an average of 3 times. The average initial fixation was 673 ms.

2.3: The plates on the kitchen bench were not mentioned by any tester in the interview but they did receive fixations from six of the testers with an average of 2.3 times each. The average initial fixation was 688 ms.
2.4: Four testers mentioned the plates and cups on the table and nine of the tester fixated on them with an average of 3.8 fixations. The average initial fixation was 526 ms.

Figure 19.

1) **Summary of kitchen section:** Every tester in this section understood that they needed to get the door open and three of the testers didn’t figure out what to do and had to be assisted to progress. Seven of the testers did eventually solve the puzzle, but the majority of them did not understand how they solved it (see figure 19). The cues measured in this section were all of texture type. Cue 2.4 got the highest number of mentions in the interview and the highest number of average fixations of 3.7 times. Cue 2.1 was never mentioned, but three of the testers did mention that the reason for approaching the drawers was due to them being positioned in a stair like formation. Cue 2.3 was never mentioned as well, but both of them did receive fixations seven and six respectively.

### C. Garage section

<table>
<thead>
<tr>
<th>Name</th>
<th>Cue</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer screen</td>
<td>Texture/Light</td>
<td>3.1</td>
</tr>
<tr>
<td>Sparks under Desk</td>
<td>Texture/Light</td>
<td>3.2</td>
</tr>
<tr>
<td>Sparks from Power-cell</td>
<td>Texture/Light</td>
<td>3.3</td>
</tr>
<tr>
<td>Key 1</td>
<td>Light</td>
<td>3.4</td>
</tr>
<tr>
<td>Drawers</td>
<td>Texture</td>
<td>3.5</td>
</tr>
<tr>
<td>Computer Screen Change</td>
<td>Texture</td>
<td>3.6</td>
</tr>
<tr>
<td>Key 2</td>
<td>Light</td>
<td>3.7</td>
</tr>
</tbody>
</table>

Figure 20. Result for metrics i and ii for garage section.

Figure 21. Result for metrics iii for garage section.

Figure 22. Result for metrics iv for garage section.

3.1 and 3.6: Only one of the testers recalled seeing the red computer screen on the desk, but the gaze data shows that seven testers fixated on it an average of 2 times each. Average initial fixation is 483 ms.

3.2: Five testers mentioned this cue in the interview and nine of the testers fixated on in according to the gaze data with an average of 5 times each. Average initial fixations is 400 ms.

3.3: The powercell was mentioned seven times during the interview with nine tester fixating on it according to the gaze data with an average of 4 fixations each. Average initial fixation was 448 ms.

3.4: The first key were mentioned by 5 of the testers and the gaze data shows
that it received fixations from all testers with an average of 6 times each. Average initial fixation was 722 ms. 3.5: The drawers with yellow handles only got mentioned two times but received fixations by all testers with an average of 4 times each. Average initial fixation was 364 ms. 3.7: The second key got mentioned 4 times in the interview and nine testers fixated on them at an average of two times each. Average initial fixation was 371 ms.

The results for the number of testers that actually fixated on the visual cue are more vague, especially when compared to the once recalled by the tester. Many of the visual cues received fixations from all testers, but not nearly as many was recalled by the testers. Reason for this is most certain due to the design of the object its associated with or its placement in the level since all that received fixations from all testers had to be interacted with in order to progress. A possible solution to counter this would be to construct a level were the player could progress in more than one way and use textured cues surrounding one obstacle and light cues for the other. Another solution could be to develop two levels sharing the same structure but with different visual cues in each.

Reasons for the prominence of lighting cues may be due to their preattentive processing. Healey and Enns, (2012), is presenting a collection of preattentive visual features that the visual cues inherit. In the game, the lighting cues are using visual features equal to hue, luminance and lighting direction that was compiled in the article. The texture cues are as well using hue, but also using direction, density and number. Evidence show that some of these visual features can be more salient than others. The visual system favours color over shape and luminance over hue (Healey & Enns, 2012). This could also be an explanation of why lighting cues are better at grabbing the players attention.

B. Discussing the method:

The design science methodology proved itself to be a good base for the research in the paper. The method was shaped to fit the research setting due to the fact that the game was not completely developed from the ground up. Gameplay mechanics such as moving, jumping, dragging was already iterated and tested from a previous course. The actual artifact that needed to be iterated in this research was the level in which would contain the tests. During the tests small bugs and glitches hindered the tests to proceed and was fixed after each tests. The alterations of the levels did not influence the data from the research because it did not affect the visual cues.

C. Discussing the data collection

The data collected from each section suggest that the gaze data doesn’t exactly correlate with our results from the interviews in regards to if a tester saw a visual cue as explained previously. This raises the question on the value of the gaze data used in this thesis and if same or better data could be obtained by doing a more extended interview? No in the sense that there are no other technologies that can measure eye movements other than eye tracking. But if a more extended interview was used instead along with a questionnaire valuable data could still be obtained. The majority of interview questions in this thesis where all designed to get data on how a tester solved the different sections in the level and if there were anything in the level that stood out to them, subsequently giving us information about what visual cues they noticed. If this interview were to be extended or complemented with a questionnaire some more data could be gathered, though
some problems might have occurred with that approach as well explained in section (data collection). One of the drawbacks by using an only interview approach is that the data obtained can only really tell us about what the tester saw in terms of what they remember from each session. That in itself gives the information about what visual cues worked better and which don’t, but it doesn’t say anything about if the tester actually saw the ones not mentioned. So the reason why a visual cue was not mentioned by the tester gets harder to determine, is it because of its placement so the tester misses it completely or is it because of its design making it less salient compared to other visual cues? At the same time, this complementary process is seen the other way around. The interview answers add depth to the data gained by the eye-tracker by more specifically pinpointing which cues are strongest, when the testers explicitly mention these. In other word, the eye-tracker on its own can be faulty without an added personalized perspective by the tester. It is impossible to know which cues are actually helping the tester or merely distracting the player. An eye-tracker can be a good complement to this since it can tell if the tester fixated on the visual cue or not, making it easier to determine what needs to be tweaked for better results.

D. Future Work

By continuing on the research presented in this paper, more valuable results can be discovered. For example:

- 1. Comparing visual cues with different tests. By switching places of the texture and lighting cues, more evidence of the cues purpose could be presented. In this study, lighting cues drew the most attention during gameplay and the players navigated the most towards them. By switching a light cue to a texture cue, you could see if the players would still notice it and if it shared the same purpose of making players navigate towards it as well.

- 2: Compare a level containing visual cues with a level without visual cues. By comparing these two level variations, evidence of the importance of visual cues could be found. Would the player know where to go or is visual cues such an important feature?

- 3: Perform more tests. By increasing the quantity of the tests, more data would be collected and would strengthen the results. For this thesis the amount of tests were sufficient, because enough data were collected to find a pattern of which visual cue was more prominent than the other.

- 4: Automation. In this thesis, the data collection was the most time consuming. By analyzing the artifacts produced by the tests, a significant amount of time were spent on exporting and watching the video recordings, write down the time stamp, export the data from the eye-tracker and find the time stamp in the data export. By automate this process it would have decreased the amount of time spent collecting the data and opened up for more resources analyzing and comparing the visual cues. A simple program that had the option to search through a data sheet with search parameters, such as time stamp would have been sufficient.

- 5: Combine the video export from the tests with the gaze visualization initially. The Tobii 4C eye-tracker was proven to give a substantial amount of data to this thesis. However, it revealed itself to be more adapted to 2D imagery. To record a game in this software became problematic due to the fact that games are an interactive non-predestined media. With the help of a Tobii Pro Glasses- device, a recording of what the players are looking at are recorded in real-time, with a camera that is attached to the front of the glasses (‘Products’, 2018). The results of an already combined video file would have helped us to direct our resources on the analysis of the artifacts, instead of creating them.

VIII. Conclusion

This study presents how eye-tracker data can be used to determine which visual cues out of two defined categories are the most prominent in a 3D adventure platform game. We constructed a level where we designed a series of puzzles and placed a series of visual cues through the level with the intention to assist the player in solving these puzzles. We then performed user-based testing where ten participants all with 3D gaming experience got to play through the level. During each session we recorded the players gaze using the eye-tracker in addition to the gameplay footage. Each session was followed by an interview were the tester got to explain how he/she solved the various puzzles and if something in the level assisted them in doing so. We then noted if the tester mentioned any of the visual cues we placed in the level. We then merge the gaze data video file with the gameplay footage and performed a video analysis. During this we noted each and every fixation the cues received from the player and measured the initial fixation time. We could then draw the conclusion that light cues were more prominent receiving a higher number of mentions in interviews along with longer initial fixation times and a higher number of fixation on average. However the way this study analyzed the eye-tracking data might not have been effective in regards to some metrics since they don’t correlate with some of the data from the interview. The video export created during data collection gave us the capability to measure the number of times each visual cue got a fixation and their duration’s, but when comparing the visual cues the testers recalled to the once they fixated on its limitations becomes clear. Since the cues that the testers recalls are in a majority of cases far lower than the ones the data shows them fixating on. This can be a result of the cues placement in relation to an obstacle as explained when addressing the first research question, but it could also be dependent on the design of the object the visual cue is associated with and the gaze data alone will not be able to answer this.

The research presented in this study is well suited for future work, due to extensive data still available to discover and a variety of more test cases available to conduct.
REFERENCES


Interview Questions

Q = question.
R = reason for question, desired information.

Q: On average, how many hours do you play video games per week?
R: In combination with experience with adventure puzzle games, determine which group they will be placed in.

Q: Are there any specific genre you prefer?
R: Get information about preference, do the player prefer mobile games, strategy games or action games for example.

Outdoor:
Q: Did you understand where to go at the start of the game?
R: To see if the visual cues in the outdoor environment leads the player to the house.

Q: As you approach the house, where there any objects that drew your attention in particular?
R: Give a sense to if the layer notice the generator and the door, or if anything else was more distracting.

Q: At the main door, where there any objects that drew your attention?
R: To verify if the player picked up on any of the cues outside the door.

Q: Can you define one visual element or object that made you understand how to proceed to the next part?
R: Information if there is a more valuable visual cue.

Q: Did you ever feel lost?
  - Why?
  - When did you find out where to go?

Kitchen:
Q: As you entered the building (kitchen) were there anything that drew your attention?
R: To verify if the player picked up on any of the cues in this room.

Q: Can you define one visual element or object that made you understand how to proceed to the next part?
R: Information if there is a more valuable visual cue.

Q: Did you ever feel lost?
  - Why?
  - When did you find out where to go?

Garage:
Q: Once you entered the garage, was it something that drew your attention?
R: To verify if the player picked up on any of the cues in this room.

Q: Can you define one visual element or object that made you understand how to proceed to the next part?
R: Information if there is a more valuable visual cue.

Q: Did you ever feel lost?
   - Why?
   - When did you find out where to go?