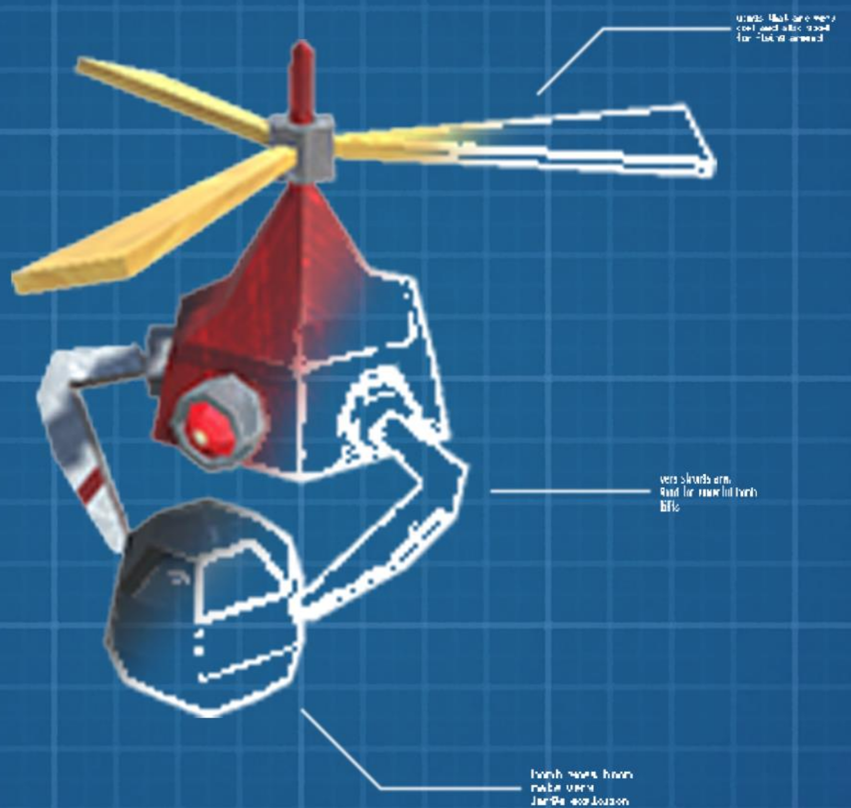
Investigating Differences in Direct and Indirect Scaffolding when Explaining Enemy Behaviour



Master Thesis by:

Kasper Halkjaer Jensen
Kasper Urban Kajgaard

Aalborg University, Denmark

May 2016

Abstract

This master thesis investigates the effect of two different approaches of scaffolding when introducing enemies in a game. Scaffolding is an approach to teaching a learner. There are different methods for implementing scaffolding and the results vary according to the method. We investigate this to be able to provide some guidelines for game developers to follow when implementing various teaching methods in their games. Throughout the report, we use two types of scaffolding which we establish as direct and indirect scaffolding. Detailed in this master thesis will be the design process of the enemies needed for the investigation as well as the design and implementation of a game. Using this game, an online user test was conducted in order to gather data to explore the effect of the direct and indirect scaffolding methods on users' playing experience and performance. The gathered data showed that the types of scaffolding used in the experiment has an impact on player performances, both in terms of completion times, and damage taken during levels. Answers from the post-test questionnaire support the claim that too much scaffolding has the potential to make the experience uninteresting for the player.

Introduction

Good game design is often graded on how well the game communicates its content to the player. Veteran game developers risk falling into the trap of assuming players know what is happening at all times based on the use of clichés established in the gaming industry, such as wooden crates containing useful pickups and barrels exploding when hit. These are conventions that experienced players will be able to recognize from other games they have played, and as such they do not really have to figure out what is going on - they will just know, based on their experience with other games. The risk for game developers is that if they fail to take into account new players who are not accustomed to these clichés and how games work in general, they may not be able to pick up on the same hints given by the game that established players can. The use of clichés can be seen as way of scaffolding and as a tool to teach the player something within the game.

Scaffolding, a way that teachers use to provide information to a learner, is widely used in game design, and a growing amount of studies confirm its relevance in designing good game experiences. Past research (Kasper Halkjær, 2015) indicates that the type of scaffolding and the approach to using them is a delicate balance. If a game developer adds a large amount of scaffolding to their game the players will find it easier, but they also tend to find the experience less interesting overall. At least that is true for scaffolding in certain areas of a game. Therefore, it is interesting to investigate the effect that different scaffolding methods have on players during a game play experience to see if the same theory applies to other areas of a game's design. This project investigates two different types of scaffolding, using an indirect and a direct approach, and compare them, in an attempt to discover their effect on game experience and how they affect difficulty and enjoyment. The scaffolding in our test is applied to the presentation of enemies in a game, including their appearance and the information that the players are provided about them. To complete this investigation, a game was designed for the purpose of the experiment, including several variations of the game utilizing different types of scaffolding, as well as a control version utilizing none. In an attempt to maximize the reach of the game, it was uploaded on the internet, and shared on various social media outlets, and data was gathered from players testing it. This also means that the game was designed in a way where the test can be completed without the help of a test facilitator.

The report is split up into seven chapters, not counting this introduction or anything preceding it. The first chapter following this is the Related Works section, which goes into the essential knowledge that we have gathered in the process of creating this project, and is useful to understand in order to follow along with our work. The Related Works section is split into a Game Design section, and a more general section on Scaffolding. After that, there is a dedicated chapter on Scaffolding, which defines the terminology that will be used throughout the report. Then we start getting into the development details with the Enemy Design chapter, which will go over the methods we used to decide on the designs of the enemies in the game, and describe the design of every enemy individually. With an understanding of how the enemies work and look, we move on to the Game Design chapter, detailing how the game was built around those enemies, what choices were made in how to teach the game to the player, the choice of controls and more. With all the design work finished, the Experiment chapter goes into detail with how the game was adapted in order to best possibly facilitate the test that we wish to carry out, and what data we wanted to collect from the test. Said results are analysed in the Results chapter, and we then have the final chapter, Discussion and Conclusion where we look at what conclusions can be drawn from our results, what the results mean, and what it would be interesting to look at within this field in the future.

Contents

1 Related Works

In this chapter we summarize relevant works along with other related studies in the field which have inspired us in shaping this project. Each study is briefly described and its relevance accounted for. This chapter is split up into two subsections concerning game design and theory, and theory on scaffolding respectively.

1.1 Game design and theory

Martin J. Osborne & Ariel Rubinstein, 1994

There are many ways to define what a game is. Martin J. Osborne and Ariel Rubinstein (1994) define it as;

> "A game is a description of strategic interaction that includes the constraints on the actions that the players can take and the players' interests, but does not specify the actions that the players do take."

In essence, a game can be defined by a task's capability to make a user think strategically in order to solve it given a set of rules to follow while doing so. It is stressed that in order for the users to become players they will have to be able to make decisions of their own within the system. Not just follow instructions given by the game. This definition provided by Martin J. Osborne and Ariel Rubinstein fits perfectly into our vision for this project. Having the players being able to make decisions by themselves despite the implemented scaffolding methods in the game is important to truly be able to test their effect. Thus we intend to design the game in a fashion that leaves the player with multiple options for solving the tasks given by the game.

Katie Salen & Eric Zimmerman, 2004

There are a lot of criteria that have to be considered for a system to truly be categorized as a game. According to Salen and Zimmerman (2004) there are almost as many definitions as to what games are as there are theoreticians within this field. In their work they have gathered the definitions of games from recognized authors of game theories and plotted them in a table (Table 1). As Salen and Zimmerman address, each of the authors have defined games for particular reasons and within specific contexts and not all of them operates in the field of game design, which can help explaining the dissimilar definitions. Despite mixed opinions there are some similarities in their different definitions e.g. a game is objective oriented and should have a goal. This knowledge can be used when figuring out where to put the focus during the design phase

Elements of a game definition	Parlett	Abt	Huizinga	Caillois	Suits	Crawford	Costikyan	Avedon / Sutton-Smith
Proceeds according to rules that limit players	■	■	■	■	■	■		■
Conflict or contest	■					■		■
Goal-oriented / Outcome-oriented	■	■			■		■	■
Activity, process or event		■			■			■
Involves decisision-making		■				■	■	
Not serious and absorbing			■					
Never associated with material gain			■	■				
Artificial / Safe / Outside ordinary life			■	■		■		
Creates special social groups			■					
Voluntary				■	■			■
Uncertain				■				
Make-believe / Representational				■		■		
Inefficient					■			
System of Parts / Resources and tokens						■	■	
A form of art							■	

Table 1: An adaptation of the table made by Salen and Zimmerman. Besides rules and to some degree goals, there are disagreements among the theoreticians within the field what defines a game.

Mike Lopez, 2006

In his article, Gameplay Design Fundamentals, Mike Lopez (2006) touches upon gameplay progression and its relevance for good game design. He states that the act of advancing forward towards a goal along with realizing a pattern of advance are essential parts of an enjoyable experience for players. Mike Lopez further proposes five key elements of gameplay progression:

1. Game Mechanics
2. Experience Duration
3. Ancillary Rewards
4. Practical Rewards (gameplay relevant)
5. Difficulty

Gameplay progression is in essence the art of awareness towards balancing and structuring the distribution of the content of a game in such a way that the players neither become overwhelmed with information and impressions nor become bored with too little new content or challenges. Since the game that we are to create is supposed to serve the purpose of being a platform for testing direct and indirect scaffolding and nothing else, we do not plan to make more than enough content necessary to test our hypothesis. In spite of that, it is still of importance for us to create a gaming experience that resembles the one of playing a real game. Hence when designing the game these five key elements of gameplay progression will be kept in mind.

Scott Rogers, 2014

Scott Rogers (2014) describes that good enemy design begins with determining the function of the enemy. Attributes contained within the function are; size, behaviour, speed, movement, attacks, aggression, health. From these attributes a huge variety of different enemies can be created and only when it has been decided how the enemy will function the visual appearance of the enemy can be designed. This is important both for saving man hours avoiding having to redo a design and for making a visual design for the enemy that as precisely as possible communicates the enemy's nature to the player. We want to avoid that unsatisfying or bad enemy design is a negative factor for the test participants affecting their experience and the data. Thus this knowledge is of great use for the enemy design.

1.2 Scaffolding

James Paul Gee, 2003

James Paul Gee (2003) argues that some of the learning principles incorporated in many good games surpasses those which are used by schools, workplaces and academic researchers and that much can be learned by these institutions about how these principles are incorporated in games to enhance learning. One of the learning principles James Paul Gee highlights is:

> "Good games give information 'on demand' and 'just in time,' ... "

Unlike in good games, school teachings often happen out of context and apart from the learners' actual goal or purposes, which in turn makes it harder for people to learn and remember. Another interesting principle he points out is the fact that a game's capability at being challenging but do-able serves as a good motivator for human beings since we like to be pleasantly frustrated, and he argues that;

> "Good games operate at the outer and growing edge of a player's competence, remaining challenging, but do-able…"

In a well-designed game, players will be confronted with a similar type of problem till they have acquired a taken-for-granted mastery in that type of problem. The players are now ready for a new type of problem in which their newly honed skills might be re-evaluated or used in unison with newly acquired skills or knowledge. This will in turn result in another taken-for-granted mastery which will yet again make way for more challenging problems. As James Paul Gee states, this is the basis cycle for producing expertise in any area. We believe that having these principles in mind and incorporating them into the game for the experiment is beneficial for testing the effect of the scaffolding.

Wood et al., 1976

With focus on the nature of the tutorial process, Wood et al. executed an experiment with the purpose of determining the importance of the instructional relationship between tutor and learner. They propose that the process of scaffolding:

> "...consists essentially of the adult "controlling" those elements of the task that are initially beyond the learner's capacity, thus permitting him to concentrate upon and complete only those elements that are within his range of competence."

It is also proposed that when done correctly, the scaffolding would lure the learner into doing actions that would produce recognizable solutions for them. In essence, for the scaffolding process to be successful the learners should not be directed in what actions to take but rather be tutored into taking a direction that could lead to the correct action in order for the learners to discover the solution by themselves. When taking the first steps into deciding how to design the scaffolding in the game, this can prove being valuable knowledge in the field.

Weppel et al., 2012

Weppel et al. (2012) conducted an experiment purposed with discovering the optimal balance of learner support necessary to win the game in the experiment while still keeping the players engaged. Based on Cates and Bruce's Model of Scaffolding (Figure 1), Weppel et al. tested the four quadrants of scaffolding spanning from intrusive to non-intrusive support on the horizontal axis and prescriptive to non-prescriptive support on the vertical axis. Testing on 21 students, Weppel et al. found that there were no incidents of test persons reporting that they received too much help solving tasks and that the test persons in the intrusive/prescriptive quadrant showed better understanding of the game functionalities. Furthermore, the prescriptive scaffolding resulted in the highest satisfaction rating among the four groups. Worth having in mind is that the game used for this experiment was a learning tool purposed with teaching programming and artificial intelligence on an introductory level. The system's objective to teach could have occluded the game-like mechanics, which could be perceived as being forced into the system. In such case, the players would not gain knowledge on how to solve a problem by playing, but would need to be directly taught what to do, to solve a problem. Nevertheless, the results from Weppel et al. are interesting and useful for designing the direct scaffolding for our experiment.

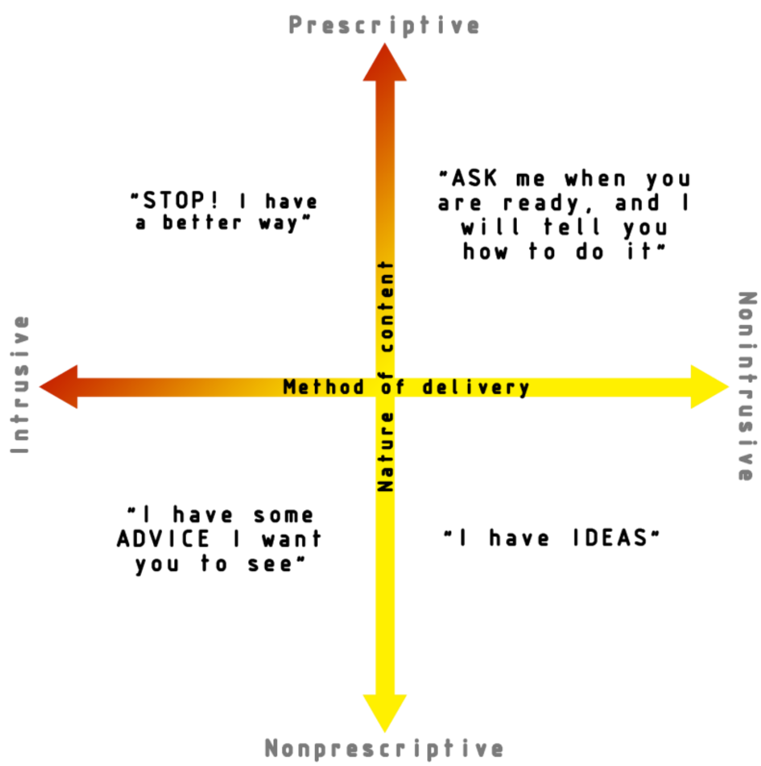


Figure 1: An adaptation of Cates and Bruce's Model of Scaffolding. The model divides a learner support space into four quadrants made up of the "nature of content" and the "method of delivery".

Kasper Halkjær, 2015

Investigating the effect scaffolding has on player experience, Kasper Halkjær (2015) executed an experiment in which two different methods of teaching problem solving in a game were tested. In one version the test persons were hinted how to solve the puzzles by clever level design, while in the other version the test persons were informed by text boxes that hinted how to solve the puzzles. Both versions were tested against a control version which had neither of the hints provided in the other versions. Kasper Halkjær discovered that the test persons found the puzzles with inlaid hints to be significantly more interesting than the control puzzles, but also that the test persons had a slightly higher feeling of satisfaction when completing the control puzzles. In conclusion he suggests that finding the best level of scaffolding is a question of balancing the amount of hints given by the game, to avoid that the feeling of accomplishment while playing should diminish. Since we want to investigate the difference in effect of direct and indirect scaffolding we are inclined to try and avoid this balance to emphasize the initial difference

Sarit Barzilai & Ina Blau, 2013

Sarit Barzilai and Ina Blau (2013) examined how adding an external conceptual scaffold to a business simulation game impacted the learner's perceived learning and ability to solve formal financial-mathematical word problems following that game. Their experiment had three experimental conditions; play only, study and play, play and study. In the play only version the scaffolding was not applied, while in the two other versions it was applied respectively before and after the game. After the participants had completed a condition they were presented with a task to solve formal financial-mathematical problems. The scaffolding in their experiment was a study activity that would teach the participant how to succeed in the game. Sarit Barzilai and Ina Blau discovered that test participants in the study and play condition performed significantly better at solving the post-game formal problems. They also noted that these participant's perceived learning from the game was reduced, though the scaffold did not have a negative influence on flow and enjoyment. This finding is interesting since we propose a similar approach for a condition in our experiment in which we expect the opposite result. We do not have a set of post-game formal problems for the participants to solve though, and this could prove to be an impacting factor on the results.

2 Scaffolding

Scaffolding comes in many shapes and sizes, as is also evident from all the previous work done in that particular field. Throughout the course of this report, we mainly talk about two different types of scaffolding, which we refer to as Direct and Indirect. In this section we wish to elaborate briefly on these two types in order to establish the terminology we will be using for them later on.

Figure 2 shows a sample scenario from a previous experiment conducted by Kasper Halkjær (2015) in which a player is tasked with completing a level. To do so the player has to push a ball unto a button located on the left hand side of the room. The test participants either got no help in completing the task, or were provided with one of two types of scaffolding; direct or indirect.

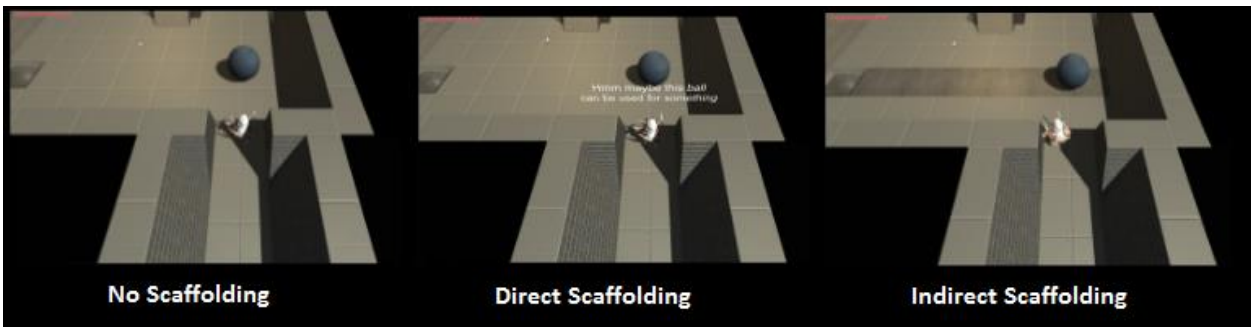


Figure 2: The leftmost picture shows the task with no form of help. In the middle picture the player is directly told via a text prompt how to solve the task. In the rightmost picture the player is indirectly guided how to solve the task by marking on the floor.

Direct scaffolding is when the player is explicitly told what to do. This includes audio clips where a narrator instructs the player what to do, as well as when there is a text on screen explaining what the player has to do. This can include both text found inside the game world, but also popups in the user interface. In Figure 2 the direct scaffolding appears in the form of a text prompt giving the player a hint as to the purpose of the ball.

Indirect scaffolding is when information that allows the player to figure out their objective is presented to them, but they are not explicitly told what they have to do. Examples of this can include a visual queue the player recognizes, from something they have done previously, being repeated in an unfamiliar scenario, or simply visuals of a level leading the attention of the player towards certain objects. In the example in Figure 2 the indirect scaffolding is provided in the form of differently coloured tiles on the ground, creating a line, leading the player's attention from the ball in the direction of the button they have to push it to.

During this project we use scaffolding to increase the readability of enemies in a game. How the scaffolding has been implemented will be elaborated on in later sections concerning enemy design (Chapter 3) and experiment design (Chapter 5).

3 Enemy design

Based on Scott Rogers' (2014) theories regarding enemy design for electronic games, and Nathan Savants' (2016) thorough investigation of enemies in The Legend of Zelda: A Link to the Past (Nintendo, 1991), a small survey preliminary to the enemy design phase for the game was conducted. The purpose of the experiment was to support the final design of the appearance of the enemies to be used in the game.

Though the purpose of the survey was to determine the visual appearance of the enemies, a robotic theme was chosen beforehand to be used in the game. This decision was made to avoid spending unnecessary man hour on making realistic and sophisticated animations for organic enemies. By implementing a robotic theme for the enemies many of the more time consuming animations like walk cycles could be circumvented and alternative types of movement, like wheels, belts and hovering, were used instead.

Prior to the experiment, six types of enemy behaviours as defined by Scott Rogers were chosen to be used in the game partly due to their diversity and partly because of their suitability for the purpose.

The chosen enemy behaviours were:

- Chaser
- Shooter
- Flyer
- Blocker
- Burrower
- Teleporter

Some enemy types were consolidated into one enemy. The flyer and bomber were combined to one enemy with the behaviour of both, and the same is true for the guard and blocker.

A set of small concise interviews were carried out on interviewees one at a time. Firstly, the interviewer explained that the enemies would all be robotic, which should be kept in mind. Then each enemy's behaviour was presented one at a time and the interviewees were asked about their opinions regarding how they think each type of enemy behaviour should be depicted and represented in a game.

Not every interviewee had an opinion about the appearance of each of the enemy behaviours and no such were forced. Nevertheless, a clear pattern of preferences regarding the enemies' appearance quickly became visible and it was deemed unnecessary to continue the survey past the fourteen participants. From the gathered data the visual design for the enemies could confidently be decided with little risk of making their appearance not match their behaviour.

3.1 Final Enemy Designs

The names and behaviours of all enemies were described to the interviewees in the initial interviews, and the designs were created to reflect the expectations participants had based on those.

For each enemy a set of sound effects were designed to help the players recognize each enemy and distinguish between them. Each enemy has a unique sound effect for; movement, attacking, and taking damage. For some enemies the sound effects will telegraph to the player what is about to happen e.g. when the teleporting enemy teleports from off-screen, or when the flying enemy is about to drop a bomb.

Chaser

The behaviour of the Chaser (Figure 3) is to follow the player at all times. It does damage if the player comes in contact with its hands, and it is vulnerable to all the different attacks the player has access to. The final version of the Chaser enemy contains many of the design elements suggested to us in the preliminary design survey, including a radar / satellite dish to make it look like it can track the player, wheels as the preferred method of travel for an enemy that is moderately quick, as well as hands to make it obvious that it poses a threat if you allow it to reach you. This enemy worked well from the first implementation, and no further changes to its behaviour were made.

Figure 3: The Chaser enemy as it appears in the game. The radar indicates that it is seeking and the outstretched arms indicates that it is reaching out to grab the player character.

Shooter

The Shooter (Figure 4) is generally static, and fires shots at the enemy in volleys of three at a time. It will move away from the player if they get too close. It is fairly small compared to some of the other opponents, based on our design test interviewees mentioning that it should look weak, as a justification for using ranged weapons presumably. It also has visible guns to make it clear that it is going to shoot at the player before it does so. Based on play testing, the shooters behaviour was changed so that when it collides with the walls of the arena it dodges away. This made players unable to corner it and kill it off very quickly, making the encounter more interesting. It was then further tweaked to be able to dodge bullets from the player, as shooting it proved a way better strategy for taking it down compared to approaching it with the sword.

Figure 4: The Shooter as it appears in the game. The guns on each side of its torso are made to take up a third of the Shooters frame to emphasize its nature as a ranged attacker.

Flyer

The behaviour that was described to interviewees regarding the Flyer enemy (Figure 5) was fairly limited compared to how the final implementation ended up. The initial description was just that it would 'fly around' and drop bombs on the player. Based on the fact that the enemy was flying, suggestions to its appearance included options such as wings or helicopter blades. We ended up going with the latter, as it made it simpler for us to animate it. Another suggestion reflected in the final design was to have the bomb it drops visible to the player, to give an indication of what it is going to do. The final behavioural pattern of the flyer is to move towards the player until it is close enough to drop a bomb. After doing so it will idle for a few seconds, and then fly away from the player for a short period. It then repeats the whole process.

Figure 5: The Flyer as it appears in the game. The propeller tells the player that this enemy might be airborne. Furthermore, the arms holding an object beneath its body indicates it has something it might drop on you.

Blocker

The Blocker (Figure 6) is unable to move away from its starting location, meaning that players are able to tackle it at their own pace, since they can just move away from it if they need a break from the action. It always turns to face the player, and cannot be damaged from the front due to the large shield. It does however have a heavy melee attack with a long recharge animation, which gives the players a chance to get behind it and hit it in its weak point. Based on interviewee feedback, this enemy is the largest by far, towering over the player. It has a large square shield, and a huge sword in the other hand, resembling a knight a little bit based on its medieval weaponry. The weak point at its back is a cabinet with exposed electronics.

Figure 6: The Blocker as it appears in the game. Its large frame telegraphs that it most likely is heavy and slow in its movements. The shield indicates that it will be difficult to deal damage to.

Burrower

The Burrower (Figure 7) was simply described as being an enemy that will chase the enemy while underground, then dig up and attack the player, and then burrow underground again a short while after. Most interviewees suggested that this enemy have a drill, and the final design reflects this. When it is above ground after attacking, it runs around confusedly, kind of resembling a mole that has been pulled out from its tunnels. The visual appearance of the enemy is reminiscent of a mole because of this as well. After playtesting the enemy, we found that players were able to kill it very quickly by focusing it down after its attack, and thus we made it dig underground regardless of how long it had been above ground whenever it is attacked. This meant that the player had to allow it to attack more than just a couple of times and made it a much more interesting encounter.

Figure 7: The Burrower as it appears in the game. The large drill at its front indicates that it could be capable of drilling or digging. The periscope-like eye works as a way to show where the Burrower is while it is submerged.

Teleporter

The Teleporter (Figure 8) is probably the most complicated enemy of them all. It is presented as a turret, and is thus generally stationary, but it will teleport to a random location in the arena at a fixed interval. Its attack is a continuous laser beam shooting from its eye, and it will turn towards the player trying to hit them with it. Whenever it teleports it will always end up facing the player. When the teleporter takes damage it will also teleport away from the player, meaning that the player cannot rapidly deal with it once they get close. Based on the design interviews, this enemy looks a bit cleaner and more futuristic than the other enemies, to make sense of the fairly ‘high tech’ nature of how it works. It was supposed to shoot projectiles like the Shooter, but this was changed to a laser beam attack on suggestions as well, to keep in line with the rest of its appearance. After playtesting, we also made it teleport away when it was about to be hit with a projectile, to make sure the blaster was not the dominant weapon against most of the enemies.

Figure 8: The Teleporter as it appears in the game. Its lamppost shape makes it appear static like a turret. The clean and elegant design along with the glowing light hoop around its body indicate that it is high tech and might have some tricks up its sleeve.

4 Game design

This chapter describes the design process of the game made for the experiment step by step. First, the thoughts and reasoning behind the decisions made during the designing process are accounted for. Secondly, the core gameplay and mechanics are presented.

To be able to make a fair experiment for comparison between indirect and direct scaffolding within a game, the game was designed to resemble an actual game experience. To achieve this, the game was designed by following game design rules and theories from professional game designers as well as recognized researchers in the field.

Even though using an already existing game for the purpose had been up for consideration, that option was abandoned early on in the process. The need to be able to control every mechanic and visual aspect within the game was deemed as critical factors, and they far surpassed the convenience it would be to save man hours on not having to create a new product for the experiment.

4.1 Game Genre

It is up for discussion whether a certain genre of video games would be better suited for this kind of experiment than other genres would, and we cannot argue against the fact that a different choice of genre might have provided different or more saying results. That the game had to be suitable for the inclusion of enemies was the only criteria the genre had to be based upon. Besides that, the choice of game genre was based on achieving as believable a game experience as possible rather than having it based on personal preference. Several other reasons for this choice include reducing the risk that poor game genre choice would influence the results negatively, and that certain features were found complementing both the goal of the experiment as well as the resources available. Thus the genre of the game was one of the last aspects of the game design which could be decided. Finally, when every aspect of the game had been accounted for, the final product ended up being a game that can be described as a third person tactical action game.

That the game is played in third person perspective was a decision taken to give the player a better overview of the play area as well as a better sense of distance between player character and game objects. To enhance the relationship between player and player character the player character was designed to be likeable and human-like despite being a robot.

4.2 Game Progress

As noted by Mike Lopez (2006) the feel of progression is a vital component in proper game design. Having limited resources, as a way to save man hours, an early decision was made to not have a storyline even though a story is a natural way to make players feel progression. As an alternative, level completion would be a solid choice for obtaining a feel of progression in the game. It would be less time consuming implementation wise and would not compromise the fact that the game had to feature enemies. Additionally, having the game being divided into levels made other progression features (e.g. experience duration, ancillary rewards, difficulty) easier to implement and control.

The experience duration is naturally controlled by the duration of each level and is closely related to the difficulty of each task in these levels. At this point, to avoid overcomplicating the game by introducing separate tasks and puzzles it was decided that the objective of the game would be to defeat the enemies in a stage. On the completion of a stage the players will know that they have progressed and that they are now closer to the conclusion of the game.

To keep the game interesting the task of each level should increase in complexity between levels. This was kept in mind when designing both the enemies and levels, and was implemented in the final build of the game. The increase of complexity is only true for the first half of the game though. Since we had to have a way to test for the effect of the scaffolding, the second half of the game would repeat the tasks from the first half, but in a different order to avoid the players would recognize the order as a pattern.

As a supplement to the scoreboard, an ancillary reward system was implemented as an extra carrot on the stick for the players to chase. If a player completes a level in perfect manner i.e. taking no damage from the enemy, when they reach the scoreboard phase they will be rewarded with a crown that the player character will wear during the following levels. The crown is lost if the player character is harmed, and the player can then try and regain it in a later level. Besides being a cosmetic change to the player character, the crown has no further effect in the game.

4.3 Decision Making

As suggested by Martin J. Osborne and Ariel Rubinstein (1994) a game needs constraints and decision making. Even the simplest form of game you can think of implements some sort of decision making for the player, whether it is a roll-and-move game like Ludo, in which the player has to decide which piece to move, or a place-and-go game like Four in a row where the player has to decide where to place the next piece. If there were no decisions to be made, the “game” would merely just instruct the player in what to do, and the “player” would change role to become a user of a system instead. Worth noting is that within a game, no matter how strict the rules are and how persistent the game is trying to hint the player in how to go about the problem solving, the players always have the choice to disobey and neglect the system in order to try and solve the problem in their own way. However, whether or not a game supports multiple solutions is a question of how the game is designed.

To give the players the opportunity for decision making in the game we implemented a set of different tools in form of three different weapons to choose from. The weapons differ in range and capabilities and are designed so that some are better for certain enemies than others. This decision was made to make it possible for the players to experiment with the tools at their disposal in order to solve the task. Some enemies were designed to be impossible to defeat with a certain tool, but the game does nothing to prevent the players from using that tool regardless. To understand which tools to use for certain enemies the players have to experiment and experience for themselves and it can in some cases come down to a matter of preference.

4.4 Core Gameplay

The game is a third person tactical action game. The players take control of a player character and have to make their way through twelve stages. Each stage is a circular room which contains nothing but an enemy that the player has to defeat in order to progress. There are six enemies that differ in both behaviour and visual appearance. As mentioned previously, from stage one through six the players encounter the enemies in an order that is based on the complexity of the enemy; the enemy in stage two is more complex than the enemy in stage one etc. The enemies in stage seven through twelve repeat the same six enemy types, but in a different order.

The players have three health points that deplete as they take damage from an enemy. If the players take three damage during any single stage they are defeated and have to replay that stage from the beginning (Figure 9). To avoid players getting stuck on a certain stage making them give up and leave the test before they have reached the end, an option to skip a level was implemented. The skip level option was available for players after having been defeated three times in any single stage, but would only be available for that stage. At the start of each stage the players' health would reset to three health points so that the players would initiate each stage with the same options and starting point. After each stage the players will be taken to a scoreboard in which the score from the just cleared stage will be added to the total score. The score is a simple composite of damage taken and time spent in a stage. The only function the score has is as a motivator for the players.

Figure 9: The player receives the last point of damage during a stage causing the player character to die. The player now has to replay the stage.

4.5 Tutorial

At the start of the game the players have to complete a four step tutorial that teaches the players the basics of how to control the game (Figure 10). In the first step they are taught how to move and are tasked with moving to a goal. At this point in the tutorial the players are unable to do anything but move the player character, since every other function is locked. In second through fourth step the players are introduced to the weapons one by one. In each of these steps the players are only able to use the weapon they are tasked with using in that step. Thus the players are forced to familiarize themselves with the different weapons they have access to throughout the game.

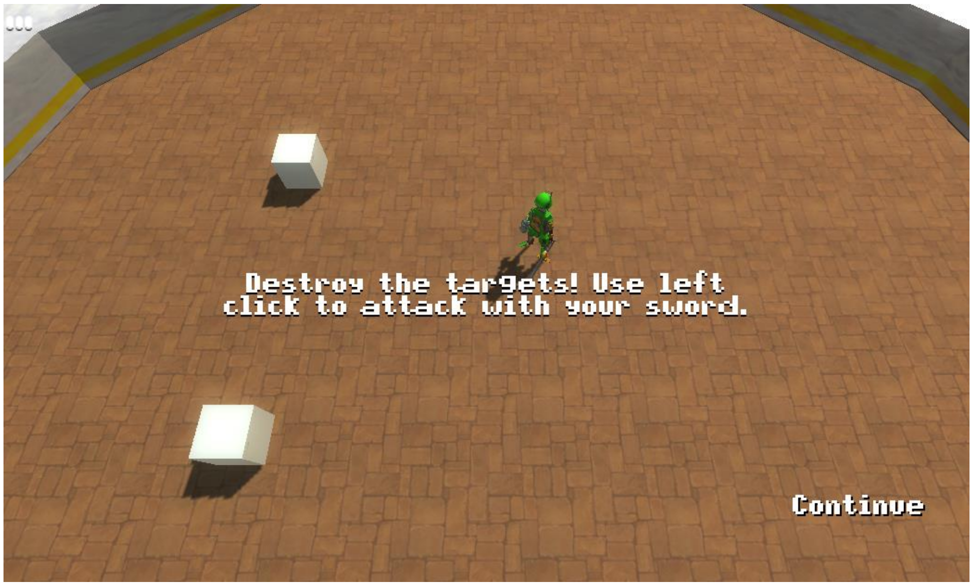


Figure 10: The players are guided by text prompts during the tutorial to teach them the basic controls and weapons.

4.6 Controls

To make the game accessible for as many users as possible it was decided to focus the controls for keyboard and mouse only. Since the test was supposed to be conducted on the internet with no form of supervision from facilitators, there would be no control over who had access to game controller, and who had to make do with mouse and keyboard. To include game controller support could prove bringing unwanted bias to the results and such risk was deemed unnecessary.

The players control the direction of the player character's movement in with the W, A, S, D keys on the keyboard (Figure 11). To use the sword, the player has to press the left mouse button, and right mouse button is used to shoot. To place a bomb, the player has to press SPACE. Moving the mouse will control in which direction the player character will attack in.

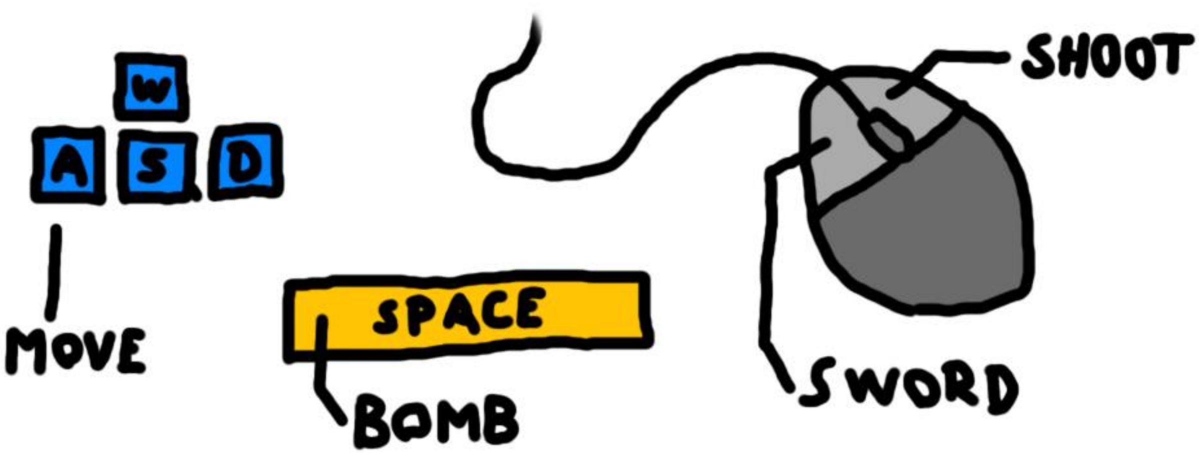


Figure 11: How to control the player character in the game.

4.7 Weapons

The players have three weapons at their disposal; a sword, a hand-cannon, and a bomb (Figure 12). The weapons were designed to be equally useful over the course of the game. This decision was made to encourage the players to experiment with the weapons in order to discover what weapon would be best in a certain situation, instead of what weapon was best overall.

The sword is a melee weapon that works at short range. It can be used at all times and has no further restrictions beside its range. The hand-cannon is a long range weapon which shoots projectiles in the direction the player points. It is restricted to only being able to carry three rounds at a time. Each time a round has been used there is a cool down timer of one second before that round will be reloaded into the hand-cannon again. This was implemented to prevent the players from rapid firing with the hand-cannon which would make it too powerful compared to the other weapons. The bomb is a timed weapon that explodes within three seconds after deployment. It deals a great amount damage to everything in an area including the player character. It has a cool down of four seconds before it can be used again.

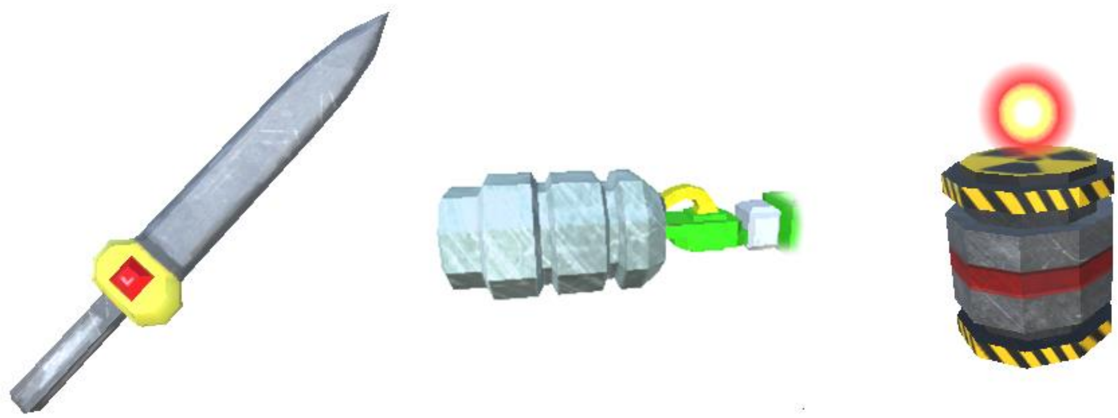


Figure 12: The weapons at disposal for the player character. From the left; the sword is a short ranged weapon, the hand-cannon is a long ranged weapon, and the bomb is a heavy damage dealing drop weapon.

5 Experiment

5.1 Experiment Design

After the game was created, variations had to be made so that the different versions of scaffolding could be held up against each other. The two methods we used were direct and indirect scaffolding. We wished to test the scaffolding methods individually to isolate the effect of each, instead of just testing them against each other. We also wanted to test the effect of the two methods in combination, and furthermore had to make a control version of the game in which neither direct or indirect scaffolding method would be present. The information the players were provided about the enemies varied as such across variations (Table 2).

	Picture	Name and Description	Appearance
Randomized	Red	Red	Yellow
Indirect	Green	Red	Green
Direct	Red	Green	Red
Direct + Indirect	Green	Green	Green

Table 2: A table showing the information the players were provided in the different versions of the game. Green = present, Red = Not available, Yellow = present, but incorrect.

In the control version, which will be referred to as the randomized version from now on, the appearance of the enemies was shuffled, so the players could meet enemies with the behaviour of a Chaser, but the appearance of a Flyer, or an enemy with the behaviour of a Shooter, but the appearance of a Teleporter etc. There was also no other information about the enemies provided for them to rely on, so they would have to get re-accustomed with whichever enemy they were facing in every level.

In the version with indirect scaffolding only, the players were not given any information about the enemies before they start the encounters, except for a picture of the enemy. The name and description of the enemy was unavailable. That means that when they started the encounter they would have to rely on the visual appearance of the enemy as well as experience fighting it to figure out how to defeat them, i.e. the enemy with the propeller probably flies, and the enemy with the big shield probably blocks your attacks.

In the direct version, the enemies all looked the exact same - like grey cubes, so the players could not rely on any visual indicators to figure out what they had to do (Figure 13). Instead, the game informed them via a text prompt before each level how they had to approach the encounter that level.

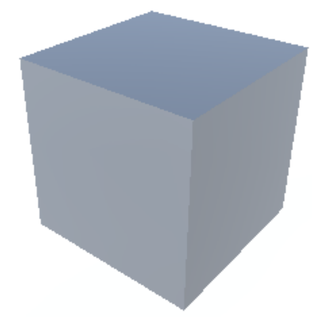


Figure 13: Every enemy appeared as grey cubes in the Direct version

Finally, there was the version containing both types of scaffolding. In this version the players were both provided the text prompt from the direct version as well as the correct visual appearance of the enemy. Thus the players had received all relevant information about how to combat that enemy (Figure 14).

Figure 14: The start screen for each stage. Top left: the randomized version. Top right: the indirect version. Bottom left: the direct version. Bottom right: the direct + indirect version.

In every test the players encountered an enemy of each type exactly twice. Levels 1-6 contained each enemy once, and then each of those enemies appeared again in levels 7-12. In levels 1-6 the order of enemies was picked so that the least complex enemies were encountered first, allowing the players to familiarize themselves with the mechanics of the game while they were playing. In the second half of the test, the order in which they met the repeat enemies was randomly selected. The order was always the same across variations.

The reason for each enemy appearing twice was to allow players to learn the appearance, movement and attack patterns, and the names and tips for defeating the different opponents. This allowed us to investigate which type of information given to the player proves to be the best approach for preparing the players for the subsequent fight with an enemy of the same type.

The full design of the game was based on the wish to be able to carry out the test online, in order to get a higher number of participants, as well as a more varied group of participants than ‘other people from the same study’. This meant that the experiment had to be completable without any interception by a test facilitator, and also could not be too long, since if people got bored and stopped playing they would not provide any useful data. These limits were the reason for the design of the tutorial and informative text prompts at the start and end of the game, as well as the overall time the experiment ended up taking.

5.1.1 Questionnaire

When test participants finished the final level, they were presented with a couple of things they were asked to rate on a scale of 1-7. The answers were used to allow us to investigate some more qualitative aspects of the experiment, without having test facilitators present, such as how much the players enjoyed themselves, and how difficult they thought the game was, though we were able to back up the difficulty ratings with gameplay data also. The questions were as follows:

How would you rate your enjoyment of the experience on a scale from 1 - 7 in which ‘3’ would equal the experience of playing the tutorial?

How would you rate your desire to continue playing the game on a scale from 1 - 7 ?
1 = I definitely don’t want to play again
7 = I definitely would like to play again

How would you rate the challenge of the game on a scale from 1 - 7 ?
1 = Really difficult
7 = Really easy

On a scale from 1 - 7, how interesting would you rate the overall experience to be?
1 = Really uninteresting
7 = Really interesting

5.1.2 Gameplay Data

In addition to the questionnaire, a lot of separate statistics from the player's gameplay were collected. These statistics include: Accuracy with each of the three weapons in the game - measured by counting times the weapons were used and times enemies were hit with the weapon, damage taken by the players in each level, time spent by the players completing each level, as well as the score the players got in each level (even though score is a composite of time and damage already). Of course, it is also tracked which game variation the players played, so that we can compare the other statistics across variations.

6 Results

89 persons participated in the online experiment, but to ensure the same amount of samples in each variation, excess samples were removed till exactly 21 samples remained in each variation. The samples kept were always the first 21 in each variation. No demographic data was collected from the test participants during the experiment. Since the experiment was conducted online we had no means of checking the credibility of the participants' statements. Thus this kind of data was considered too unreliable even though it could have been useful.

All of the data was analysed using Shapiro-Wilk tests to figure out if it was normally distributed, and then, since it was not, Kruskal-Wallis tests were used for comparisons of the entire data sets, and Mann-Whitney U tests were used for the single variation comparisons.

	Both	Direct	Indirect	Randomized
Both				
Direct				
Indirect				
Randomized				

Table 3: Shows the possible between variation comparisons that can be made.

Since the Kruskal-Wallis tests run a total of six comparisons between variations, as seen in Table 3, we will make use of the Bonferroni correction for single between group comparisons to make sure our results still prove significant. This means that the Mann-Whitney U for the between group tests will need a p-value of 0.05 / 6 = 0.00833, or at least 0.01, for us to be able to tell if they are significant or not.

6.1 Questionnaire Results

Before looking at the data from the questionnaire a Shapiro-Wilk test was carried out to figure out if the data is normally distributed or not. As it turns out, it is not, with p-values for all of the questions being <0.001, so instead of the ANOVA test you would use for a normally distributed dataset, a Kruskal-Wallis test will be used to compare the ratings.

The first variable we collected from our questionnaire was how enjoyable the players found their experience of playing the game. Although plotting the data and visually inspecting the graph shows randomized as being quite a bit lower than the other three, this difference is not large enough to be significant, with the Kruskal-Wallis test finding no significant correlations across any variants (p-value 0.47) (Figure 15).

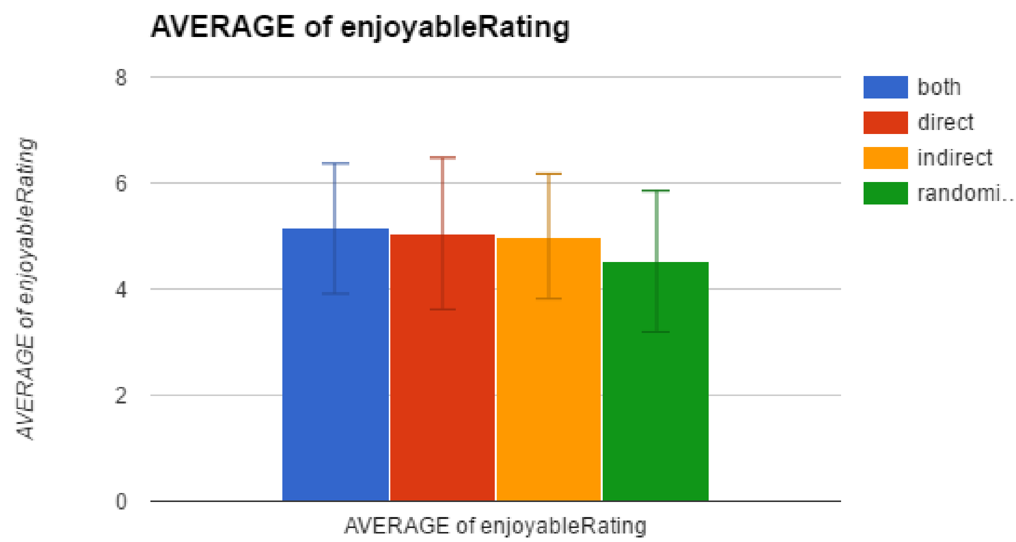


Figure 15: Graph showing the average ratings of how enjoyable players found their game experience, with a bar representing each variation.

Next, participants desire to continue playing the game was measured. Interestingly, a visual inspection of this graphed data shows randomized as the highest by a small margin, but with no significant differences between the variations, and very high p-values across the board in a Kruskal-Wallis test (0.95), the differences can most likely be chalked up to some outliers (Figure 16).

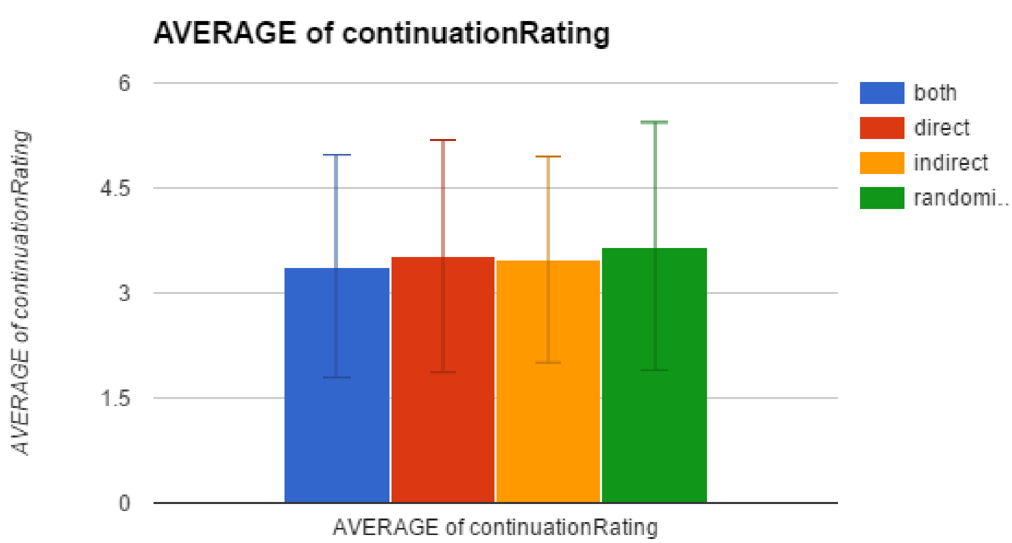


Figure 16: This graph shows how players rated their willingness to continue playing the game.

The third question asks the players how challenging they found the experience. The Kruskal-Wallis test carried out on this found that there was no statistically significant difference between the groups (p-value of 0.29). Randomized looking a lot higher than the other three prompted us to do an extra test, but a Whitney-Mann U test found that this difference was not large enough to be significant. ‘Randomized’ versus ‘both’ was the closest, but only had a p-value of 0.09 (Figure 17).

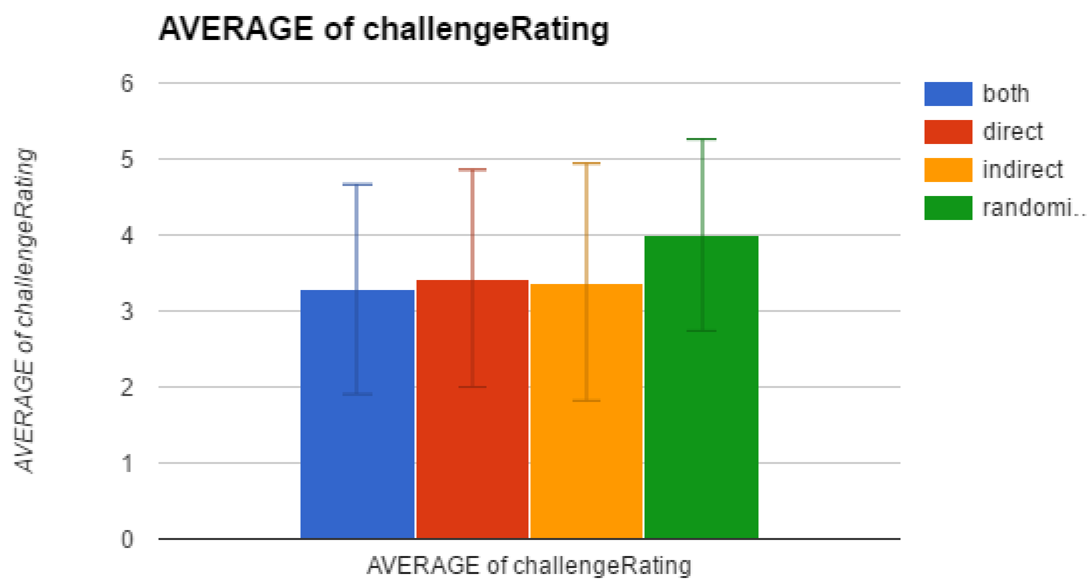


Figure 17: Graph showing the average ratings of how challenging players found the game.

Finally, players were asked how they thought the overall experience of the experiment was. Overall, the Kruskal-Wallis test of this also proved statistically inconclusive (p-value of 0.29), but a Mann-Whitney U test carried out afterwards did find that ‘indirect’ was higher rated than ‘both’ with a p-value of 0.05, but taking the Bonferroni correction in mind, this is still not small enough to consider significant (Figure 18).

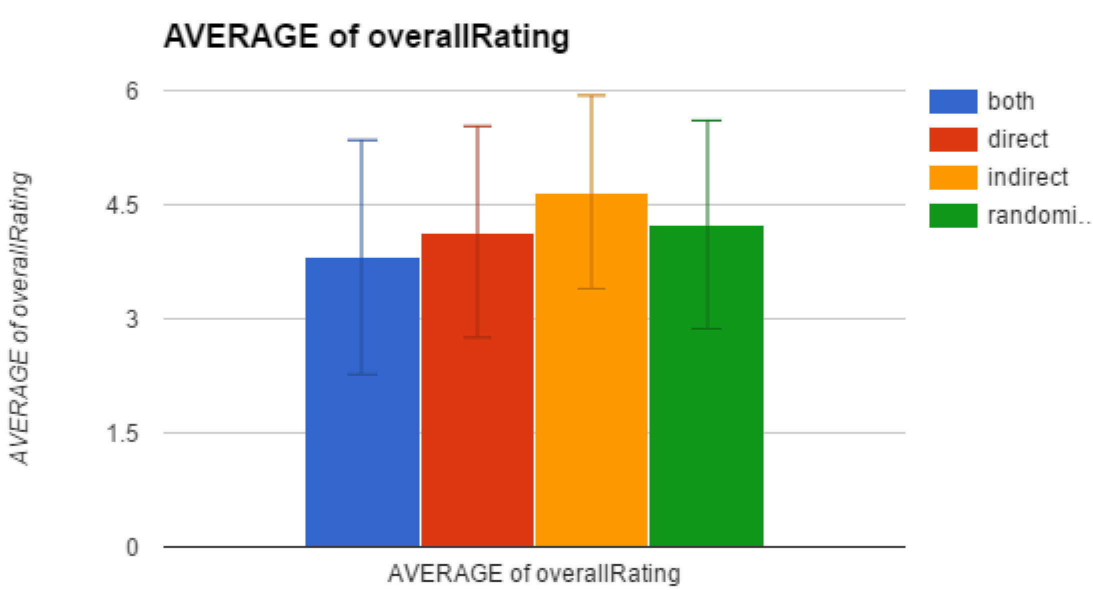


Figure 18: This graph shows the test participants overall rating of the game experience.

6.2 Gameplay Data Results

The first factor we look at from the gameplay data we collected is the player's accuracy with the different weapons across the variations. A Shapiro-Wilk test was used to figure out if the data was normally distributed or not, and with p-values of 0.002 or less, none of the accuracies were normally distributed. Then, comparing the accuracies for each weapon across variations we found no significant differences for Bomb and Blaster, with a Kruskal-Wallis test giving us p-values of 0.27, 0.15 respectively. The differences in Sword accuracy however, proved big enough to be significant, with a p-value of 0.01 (Figure 19).

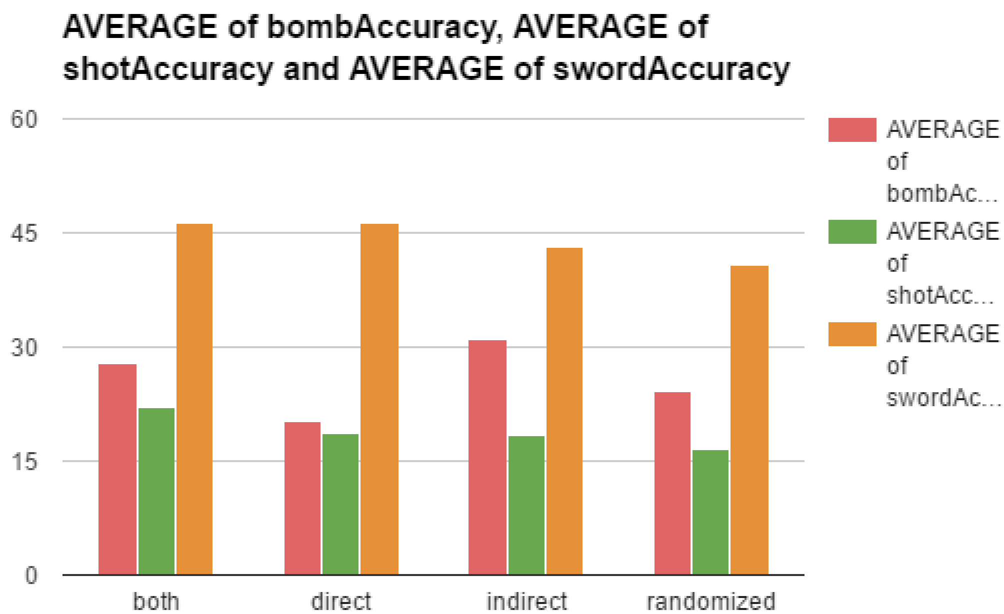


Figure. 19: This graph shows the test participants accuracy with each different weapon in the game. The bars are grouped by game variation, and each colour represents a different weapon.

Secondly, we look at the damage taken by the players in the different variations. A Shapiro-Wilk test confirms that this data is also not normally distributed, with a p-value of <0.001 for player damage. A Kruskal-Wallis test of this data set shows that the difference between variations is highly significant, with a p-value of <0.001 between groups. A Mann-Whitney U test was then used to compare some of the groups directly. Players playing 'indirect' did not take significantly less damage than players playing 'direct' (p-value of 0.2). They also did not take significantly more damage than players playing 'both' (p-value of 0.1). However, players playing 'direct' did take significantly more damage than players playing 'both' (p-value of 0.003) (Figure 20).

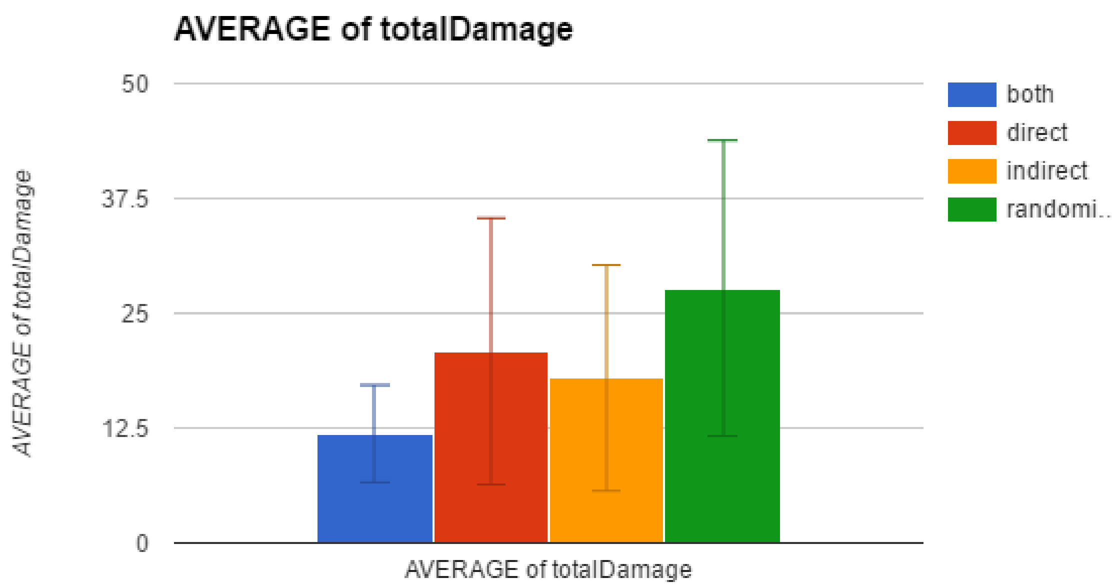


Figure 20: This graph shows how much damage the average test participant took over the course of the entire test.

Going even further into the numbers from the damage taken, here is a look at how much damage players took in the first and second half of the test. All variations showed that players had taken significantly less damage in the second half of the test compared to the first, with p-values from a Mann-Whitney U test of 0.05 for 'both', <0.001 for 'direct, <0.001 for 'indirect', and finally 'randomized' with a p-value of 0.05 (Figure 21).

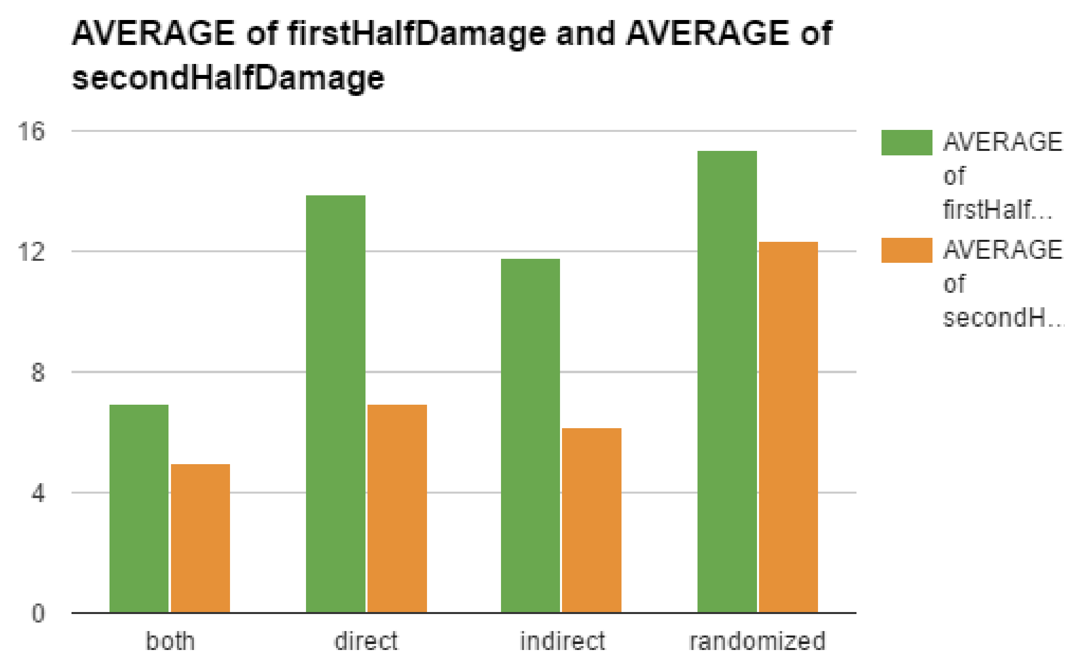


Figure 21: Again a graph showing the damage taken by participants in each variation, but this time it is further separated into how much damage they took in the first and second halves of the test.

Next up is the time spent by players on each level. This data is also not normally distributed, with a Shapiro-Wilk test giving a p-value of <0.001. Visually this graph looks fairly similar to the previous one, which shows us that in the levels that were quickly completed by the players, they also tended to take less damage overall. The difference in how much time the players spent across variations is statistically significant, with a p-value of 0.01 from a Kruskal-Wallis test carried out on the samples (Figure 22).

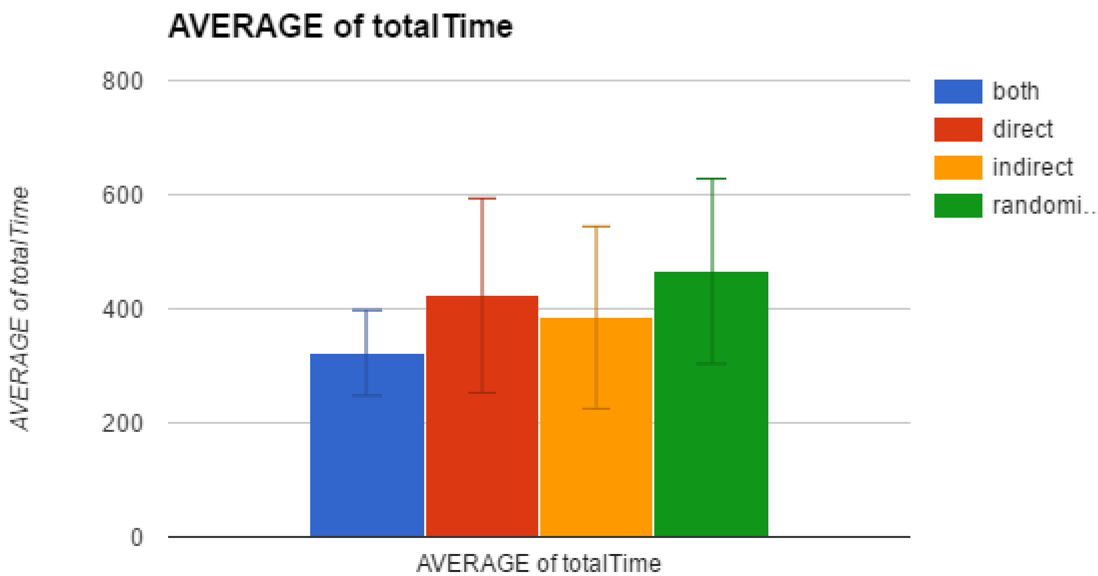


Figure 22: This graph shows the average time participants spent on the entire game period.

Similar to what we looked at for damage taken, this chart shows the average time taken in the first and second half of each player's play through. 'Randomized' is the only variation in which no significant improvement can be found, with a p-value of 0.09. Players playing 'both' do not show as big of an improvement as the last two versions, but their p-value still comes in at 0.022. For 'direct', this number is 0.005, and finally, players of 'indirect' show the biggest improvement by far, with a p-value of <0.001.

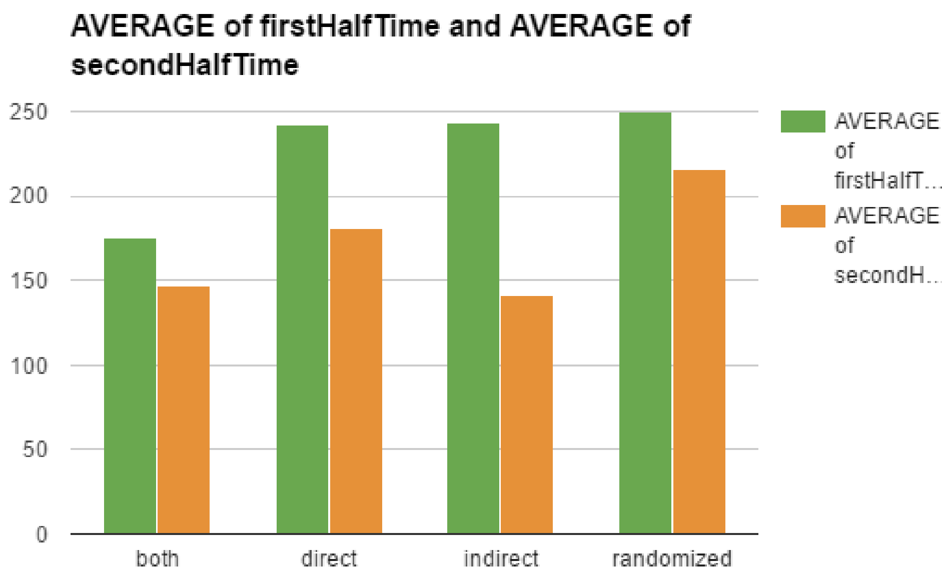


Figure 23. This graph shows the time spent by participants, but separated into first and second halves of the test.

The last factor in the data we have collected is the player's score. It is not as vital as the previous two, since score consists of a composite of damage taken and time spent, but we wanted to include it regardless so all the data we are using for conclusion and discussion is present. The difference in player's score is also statistically significant, with a p-value of 0.01 from a Kruskal-Wallis test.

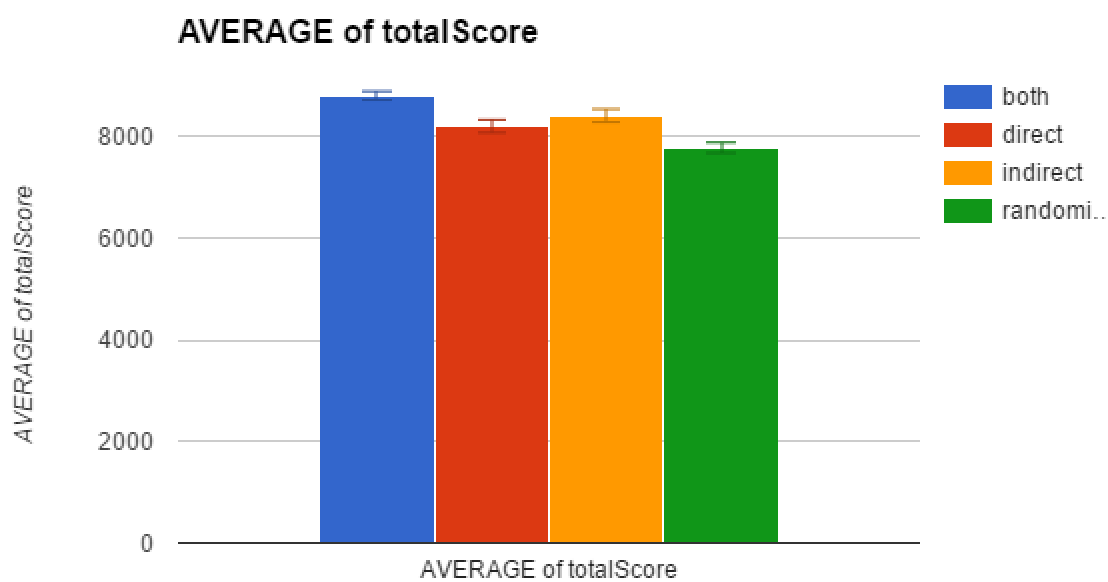


Figure 24: Graph showing the average score achieved by players in each variation.

7 Discussion and Conclusion

The answers we got from player's questionnaires proved mostly inconclusive, with no real significant answers to speak of. A cursory look at the graphs mostly tells the story we expected though, with 'randomized' receiving the highest rating for challenge, and the worst rating for enjoyment, 'indirect' receiving the highest overall rating (though from our expectations this could have gone to 'both' also), and finally, nothing really winning out over anything else in the continuation desire ratings.

Moving on to the data collected throughout the experiment, we found that the difference in gameplay variations had no significant impact on the player's accuracy with two of the three different weapons they had access to. The one that proved significantly different across variation was the sword, and it can reasonably be assumed that the largest reason the accuracy with the sword varied so much was because the hitboxes in the randomized test often did not line up perfectly with the visuals of the enemy.

The next significant result shows up when we look at the damage taken by players in the different gameplay variations. Player's took the most damage in 'randomized', second most in 'direct', then 'indirect', and finally, the least damage in 'both'. This shows us with a large degree of certainty that the scaffolding methods we have worked with provide a good way of teaching the players how to act against enemies of different types, and allow them to predict their behaviours on subsequent encounters.

The notion that the players learn how the enemies work better in the versions with better scaffolding is supported by the next data points that we looked at, which is how much damage players took in the first half and second half of the experiment respectively. In 'both' players had the best initial performance, but they did so well that they did not have a large amount of room for improvement, leading to the improvement in this scenario not being as significant as the ones shown by 'direct' and 'indirect'. In 'direct' and 'indirect' however, players start out with pretty poor initial performances, just like 'randomized', but unlike 'randomized', the player's performance in the second half is markedly improved, and actually comes close to rivalling how well players did in 'both'.

Following this, we get to the other major statistic that was tracked for players across the whole experiment, and that is how much time they spent playing. This data mostly mirrors the damage taken, with the scenarios in which the players took the most damage also taking them the longest time to complete. In order of slowest to fastest they are; 'randomized', 'direct', 'indirect' and 'both'. It makes a lot of sense for players to take longer if they also take more damage, since if they take three damage in one level, they die and will have to start over, while their time keeps counting up.

Just like with damage taken, we again separate the time spent into the first and second half of the test to see if the players are showing an improvement over the course of the test. As with damage taken, players playing 'randomized' show no significant improvement, players playing 'both' do well initially and only improve a little, and players in 'direct' and 'indirect' start out doing quite poorly, just like the players in 'randomized', but improve a lot over the course of the test. In fact, the improvement of the players in 'indirect' is so big that in their second half their performance matched, if not bested, the performance of the players in 'both.

For good measure we also made sure that the significant results from damage taken and time spent match up with the player's total scores, and fortunately it did, seeing as the scores are based on damage and time.

Overall, the data shows us that while players provided with a lot of information right from the very start, such as in 'both', tend to do very well from the beginning of the test, the players provided with slightly less information, such as in 'indirect', are quite quickly able to catch up and perform to the same level as the players who also had description. In our ~20-minute test, 'indirect' players were able catch up with the performance of 'both' when they got to the second half of the test, meaning that in most scenarios it will be perfectly fine for game developers to not provide direct information to their players, and instead rely on player's ability to figure out what to do for themselves - since it will most likely not take them very long.

One factor that we did not control for during the experiment was how players controlled the game. The game was designed to work well with keyboard and mouse, but while the controls were explained in the tutorial, players were not prohibited from playing the game with a keyboard and touchpad for example. We assumed that playing the game this way would be too frustrating for players to actually play through, but since we do not have a way to actually check for it, that might be the cause for some outliers in our statistics.

In the future, it could be interesting looking into a good way to apply the knowledge gained over the course of this project to a game with more varied levels, and multiple enemies at a time. There is still the chance that the test experience did not really feel like a game to some test participants, as the 'one enemy at a time' approach can end up feeling a little like a test scenario. Making something that feels more like a real game, but is still experiment friendly in terms of the data you can gather from it is a challenge, but might prove to give insight into some things that this project did not.

Furthermore, the concepts of direct and indirect scaffolding provide a very robust framework for future work in the same area, and is applicable to almost all aspects of a game, such as level design, enemy design and more.

8 References

Halkjær, Kasper. *Investigating the effect of scaffolding in modern game design.* 2015

Osborne, Martin J., and Ariel Rubinstein. *A course in game theory*. MIT press, 1994.

Salen, Katie, and Eric Zimmerman. *Rules of play: Game design fundamentals*. MIT press, 2004.

Lopez, Mike. *Gameplay design fundamentals: Gameplay progression*.Gamasutra, 2006. http://www.gamasutra.com/view/feature/130188/gameplay_design_fundamentals_.php. Accessed May, 23, 2016.

Rogers, Scott. *Level Up! The guide to great video game design*. John Wiley & Sons, 2014.

Gee, James Paul. *What video games have to teach us about learning and literacy*. Computers in Entertainment (CIE) 1.1 (2003): 20-20.

Wood, David, Jerome S. Bruner, and Gail Ross. *The role of tutoring in problem solving**. Journal of child psychology and psychiatry 17.2 (1976): 89-100.

Weppel, Sheri, Mj Bishop, and Hector Munoz-Avila. *The design of scaffolding in game-based learning: A formative evaluation*. Journal of Interactive Learning Research 23.4 (2012): 361-392.

Barzilai, Sarit, and Ina Blau. *Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences*. *Computers & Education* 70 (2014): 65-79.

Savant, Nathan. *Enemy Design in Link to the Past.* Gamasutra, 2016, http://www.gamasutra.com/blogs/NathanSavant/20160104/262944/Enemy_Design_in_Link_to_the_Past.php. Accessed May, 23, 2016.

Nintendo, *The Legend of Zelda: A Link to the Past.* https://en.wikipedia.org/wiki/The_Legend_of_Zelda:_A_Link_to_the_Past, 1991