This thesis is submitted to the School of Computing at Blekinge Institute of Technology in partial fulfillment of the requirements for the degree of Master of Science in Computer Science. The thesis is equivalent to 24 weeks of full time studies.

Contact Information:
Author:
Daniel Palm 850717-1976
E-mail: daniel.erik.palm@gmail.com

University advisors:
Dr. Johan Hagelbäck
School of Computing
Dr. Veronica Sundstedt
School of Computing

School of Computing
Blekinge Institute of Technology
SE-371 79 Karlskrona
Sweden
Internet : www.bth.se/com
Phone : +46 455 38 50 00
Fax : +46 455 38 50 57
In the wake of motion capture and visual animation one aspect seems to be lacking. A realistic representation of the dynamic and unpredictable human visual perception. A human examines her surroundings in an unpredictable, saliency based and top-down task oriented manner. In the field of computer science, interaction design and the industry of game development, great leaps have been taken when it comes to capture motion of bodies. Motion capture helps developers and film makers to portray realistic humans in virtual environments. Where motion has come far, eyes and perception has not. As of yet a virtual representation of human visual perception, has not been mimicked as close as body motion.

This thesis will examine perceived realism in virtual agents, with a focus on eye motion. In this study a virtual agent has been given eye movements of human beings and been compared to an agent based on current virtual agents in games. This is the first step towards synthesizing more than just human motion in virtual agents. It will provide future research with the data and tools needed to produce an algorithm based on the gathered data.

Prior context research includes a study of current games. Two participant experiments have been be conducted, both has recorded eye positional data for analysis. The first experiment helps build the second as it compares virtual agents using a Likert scale for a subjective rating of realism.

The results offers some very interesting data, indeed data that lie at the core of the study as well as data for further studies. While statistical analyses of Likert scales might be considered ambiguous this study has done so and reached a conclusion. A virtual agent enhanced with eye-motion based on human eye movements does portray a more realistic human like behaviour.

**Keywords:** Eye-tracking; virtual agents; gaze; realism; perceived realism; eye-motion capture; Likert scale; t-test; digital games, human computer interaction
Abstract

1 Introduction
   1.1 Background
   1.2 Reasoning
   1.3 Aim of Research
   1.4 Related Findings in Literature
   1.5 Practical Studies and Observations
      1.5.1 Assassins Creed III Analysis
      1.5.2 Deus Ex Human Revolution Analysis
      1.5.3 Fable III Analysis
   1.6 Early Findings

2 Experiment One
   2.1 A Gaze at Agents Experiment
      2.1.1 Source Material and Data
      2.1.2 Work Process
      2.1.3 Translating Observation Into Motion
   2.2 Experiment Setup
   2.3 Compilation of Data
      2.3.1 Fixations
      2.3.2 Timestamps and Duration
      2.3.3 Area of Interest
   2.4 Mean Gaze Plot

3 Experiment Two
   3.1 A Subjective Gaze Experiment
   3.2 From Data to Visuals
   3.3 Likert or Not
   3.4 Opinions with Gaze
   3.5 Experiment Setup
   3.6 Compilation of Data
   3.7 Likert Results
      3.7.1 Likert Analysis
      3.7.2 Correlation of Numbers and Opinion

4 Results
   4.1 Results from Experiment One
      4.1.1 t test
   4.2 Results from Experiment Two
      4.2.1 Mean and Standard Deviation Analysis
      4.2.2 Statistical Analysis
   4.3 Overall Results

Contents
5 Discussion

5.1 Conclusion .................................................. 35
5.2 Further Research ............................................. 37

A Appendix

A.1 Participants and Data Points ............................... 38
A.2 Fixations and Percentage ................................. 38
A.3 Data Calculations in R ................................. 40

References .................................................. 41
Chapter 1

Introduction

1.1 Background

Efforts to enhance realism of motion in computer generated characters has come a long way with the help of motion capture for both the film- and game industry. Large game worlds inhabited by moving and communicating agents has since long been a reality and has contributed to immersion and an increased sense of realism during gameplay. The visual aspects, motion and graphics are responsible to present a vivid world through the window of a screen. These are enough to provide a suspension of disbelief unless the narrative is lacking[10].

While engaged in play with an avatar, single- and multiplayer, one is highly likely to encounter non-player-characters (NPC’s). Their movement and pathfinding behaves realistic and dynamic enough to avoid obstacles in their path. The artificial intelligence (AI) helps the NPC navigate, react and adapt to many different situations, but it has limitations. The awareness of the NPC’s surroundings, the player character (PC) and additional NPC is very static, they see and react only what they are programmed to do. In some cases a visual cue is triggered when an NPC has stepped into an area, causing them to notice an event. In other cases the NPC receives a call from the code once the PC has entered the area or an area nearby the NPC. This effect in particular is more noticeable in online multiplayer games[13].

The result is that as a player controlled avatar you do not exist in the virtual world up until you initiate any kind of direct interaction with the world. This mentioned interaction can take one of many forms, be it entering an area, triggering a global event, unsheathing a sword or intentionally initiating a dialogue. It gives the impression that the player is part of an interactive narrative where you have little effect on the environment around you. Or rather that your environment is almost never aware of you as a player. This fundamental issue has always been present in games, and it can never be done away with. There is however methods to alleviate the solitude and linear mechanics of games.

The advent and development of motion capture technology has enhanced the cinematic and game industry to portray realistic characters and avatars to an audience. Movement of a virtual human has been enhanced thanks to motion capture, realistic eye movements in a virtual human have not. When interacting with a NPC in a virtual environment there is still an unnatural element, eye movements follow a predictable and predefined path across games or it seems as though motion is added just because eyes move. The sense of human curiosity, reaction and situational awareness cannot said to be present or noticeable. In the context of games the agent is ‘made for it’s environment’ which results in a double edged sword.
On one hand they have excellent obstacle avoidance and navigation (i.e. racing games). On the other they are highly predictable and outsmarted (i.e. shooter games).
In the context of cinematics the agent is made to mimic a real life human as close as possible in every aspect. It is of even higher importance to have the virtual human mimic human behaviour even closer to avoid the audience rejecting it.

Eye-tracking technology is still fairly young and has a similar role in the industry as motion control or motion capture. It serves as an input for games and applications or offline analysis[21]. Good examples of sibling technologies are the Kinect (Microsoft Xbox 360), Move (Sony Playstation 3) and Wiimote (Nintendo Wii). What is missing is the reverse application of the technology. The use of eye-tracking as a means to simulate realistic motion of eyes as motion capture has for bodies.

1.2 Reasoning

Does a virtual agent with human eye motion appear more realistic than current virtual agents?

Imagine walking along a street filled with other people, involuntarily you seek eyecontact with the majority of the crowd. This curious behaviour is inherently human, you feel watched and acknowledged as a person in a crowd. Quite the opposite is true in a virtual world in the same situation. Acknowledge of your existence takes place either when you initiate a dialogue or pass close to another virtual agent. The argument is that virtual agents exist at both ends of the perception spectra. They either lack or has a high perception of their environment depending on the situation. This inhibits a natural interaction between the audience and virtual agents.

Previous work has shown that when faced with virtual characters the gaze of the observer shifts towards the face[8], humans are predisposed to search for faces in their surroundings. If the observer is met with static eyes and vivid head movements that might be deemed abnormal or inhuman, she or he is less likely to form an emotional attachment or feel empathy towards the virtual agent. On the other hand, movements or behaviour that pass ‘too close’ to human-like might cause the observer to reject the agent entirely. This has the chance to occur if the virtual agent displays a ‘too good’ human-likeness in motion and appearance, but off by just enough to notice[5], this problem is associated with the concept the Uncanny Valley[16]. The Uncanny Valley is usually depicted as a graph that plots out familiarity along the Y axis and human likeness along the X axis. It is possible to overcome this obstacle with the appropriate balance of motion and human-likeness.

In relation to eye movement, other studies have shown that eye fixation and movements of virtual agents has shown to influence observers to a certain extent[17]. This approach could help provide a more dynamic interaction between an audience and virtual agents. If for instance a developer wishes to lead the gaze of the player towards an event using the shock of an agent. Or perhaps if a developer tries to have an agent fool or distract a player using only the agents gaze.

There have been models of human gaze behaviour developed and implemented into virtual agents[18][19]. This has as of yet only been done in a laboratory environment though and not yet tested for any entertainment audience or purpose. Said research did not intend to gauge realism of virtual agents based on human gaze. These are probability models based on saliency, what a person is more likely to look at. There is also research related to focus prediction[2]. An approach similar to this project, only with a comparison of saliency maps contra eye position data.
In relation to the mention of a strong narrative in virtual worlds, some attempts to provide a dynamic dialogue based on gaze has been successful. The approach is based on tagged objects in a scene instead of objects caught in a virtual agents gaze\[20\]. While it does provide a realistic curiosity behaviour of the agent, expressing opinions on objects lying around, it does not help or prompt the agent to explore by itself. It is merely the virtual agent reacting to tagged objects that cross its field of view. The technology has been in use since 2007 with the release of Team Fortress 2\[^1\], instead of gaze though it is used in conjunction with cross-hair focus of the player. Hovering over a tagged object and issuing a command causes the game to play certain audio clips depending on context.

The claim is that this human delay in gaze point to gaze point navigation or motion is missing from virtual agents and needs to be considered to create the next generation of artificial intelligent, interactive and curious virtual agents.

The interesting challenge is to implement human perception elements\[18\], sight or gaze, to a virtual agent and examine the effect of perceived realism from observers. The goal would be to have an agent experience the world as a human would. As humans we cannot detect an entire scene just by panning it with our gaze fixed, we move our eyes to examine objects blurred by our peripheral vision. Virtual agents do not, they experience their surroundings by panning the scene before them without the help, or delay of eye movements.

In the context of cognitive neuroscience salience is a measure of how (much) an object stands out in relation to its neighbours. The effect is subjective and the object in question affects different triggers in an observer depending on the observer. Most salient objects affect the visual system by colour and motion, a colourful object or an object in motion draws attention to itself. The process wherein the human brain chooses to shift the gaze and focus on an object is governed by this system. As such the system has built in flaws or rather delays in order to piece together a coherent overview\[7\].

For the purpose of discussion, the observable human gaze behaviour could be described to follow three stages Attention, Focus and Act in simple terms.

- **Attention** implies that something salient has entered the field of view (FOV), either something with motion or colour.
- **Focus** is defined as the object that triggered the attention moves from the peripheral vision to the foveal vision, in other words the eyes move and centres the object.
- **Act** suggests an action taken in relation to the object, which would depend on the object in question.

This process is a vastly simplified description of what happens each time your eyes move. The description assumes that the attention is externally triggered and not voluntarily initiated, like scanning the keyboard for the right letter for example. If you were to apply the process to a virtual agent you would notice some key differences. For one, an agent does not show any signs of going through the phase Focus. A current generalized virtual agent will go through the Attention and Act phase, but with some differences.

\[^1\]<http://www.teamfortress.com>
The entire FOV can trigger the Attention, as the agent does not have a separation of peripheral and foveal vision. From the virtual agent's point of view, the entire FOV is its foveal vision.

1.3 Aim of Research

Can a virtual agent enhanced with human gaze be perceived as more realistic than a virtual agent that has not? The head and face area determines the perceived realism of the virtual agent, by enhancing a virtual agent with human gaze behaviour it should be perceived as more realistic.

This research project intends to take a step towards a more balanced virtual agent. The first step is to examine how a virtual agent with human eye-movements is perceived by observers. Should the result prove to be positive, there is incentive to examine the prospect further.

The results presented from this research will be based on what an individual actually look at. Further research may lead to a model based on real gaze data; however, this project aims to show whether people will accept a more realistic gaze pattern of virtual agents at all, as a first step. This research project will examine and determine the key differences of perceived realism between a virtual agent with eye and head movements based on current games and a virtual agent with eye and head movements based on real eye-tracking data.

A baseline of gaze behaviour has to be established first. An in depth study of virtual agent behaviour, how developers have thus far dealt with head and eye motion of characters in general and interaction events specifically. In order to get a reliable median at least three sample games from different engines and with different characteristics were chosen.

From that an animated clip with a virtual character is created, based on the game research study. The first gaze experiment is performed where participants are asked only to view the produced clip. Their recorded gaze data is compiled and analysed, the data in turn is used to create another clip with a virtual agent, this time based on the gaze data. The second experiment has participants gauge the realism of 3 clips or rather their subjective opinion on the realism. One clip is identical to the first, the second clip is created from the gaze data and the third is a control clip. With the data collected and analysed a result can be shown and conclusion be stated.

1.4 Related Findings in Literature

Most of the related work and inspiration has come from McDonnell and her colleagues out of Trinity college Dublin. McDonnell et al. has done a lot of work related to virtual crowds and the perceived realism of them. However, determining the realism of an virtual agent with human gaze behaviour or the thought to do it has not been found among related publications. Within this specific specialization McDonnell's work could be considered closest.
Chapter 1. Introduction

The initial inspiration for this research however came from an entirely different field. Elan Ruskin (previously at Valve software) did a talk at GDC\textsuperscript{2} 2012 (Game Developers Conference) about a dynamic dialogue system that could help story writers create a fluid dialogue tree for game developers. The system also included capabilities to have virtual agents 'seem' to react to objects in their surroundings. This sparked the idea for having virtual agents actually made aware of their environment in a new way with the help of human interactions.

It seems as though the industry has become aware of this also as developers themselves has mentioned the very static and strapped in ride the player is taken on. Open world sandbox concepts are more frequent, at the same time linear structured games with little or no allowed interaction apart from what the player is presented with is still going strong.

1.5 Practical Studies and Observations

Three different games from different genres and systems were examined for this research project. Assassins Creed III, Deus Ex Human Revolution and Fable III. The mean interaction behaviour will lay the foundation for the animated clips used in experiment one. Each game had previously been experienced numerous times and as such, each game research session lasted a maximum of two hours.

In the following sections the term non-playable-character (NPC) will substitute the phrase virtual agent as it is necessary to separate player and computer controlled characters. The term player character (PC) will help compliment for the same reason.

It cannot be stressed enough that any distances, angles or metrics mentioned in the following sections are merely approximations due to different perspectives and scale of each game. This includes graphics and diagrams representing scale and or distance.

\textsuperscript{2}<http://www.gdconf.com>
1.5.1 Assassins Creed III Analysis

Assassins Creed III is a Third Person adventure/platform game developed by Ubisoft with balanced focus on story (interaction) and action. The game has vast crowds of dynamically spawned NPC, which means that minimal interaction is possible with each NPC. The crowd mechanic lets NPC’s flow past the PC, as the PC was one with the crowd.

In the neutral state NPC’s neither seek him out or avoid him completely. Non essential characters mostly ignore the player avatar. Important characters tend to walk past with occasional comments or greetings, in some cases these NPC’s fixates onto the player avatar, following the PC with a fixed gaze until the angle gets too sharp and the head swings back to face forward, as illustrated in Figure 1.1. It gives the impression that little interest is shown towards the player up until the point the player almost bumps into an NPC.

Figure 1.1: A NPC passing the player avatar in Assassins Creed 3

Figure 1.2 shows the perceived field of view for a NPC in the game. The diagram has been determined from observation of the passing agents, as such it is an approximation. Field A shows the FOV when the NPC is stationary, idling or walking forward. This was determined due to the fact that NPC’s did not seem to notice the PC until a certain distance, thus the assumption was drawn that the FOV is fairly narrow. Field B shows the NPC reacting to the player or player avatar. The observations here showed that when NPC’s were close enough to react to an action taken by the PC the FOV seemed to broaden, this correlates to the increased area of influence of the PC described in Figure 1.4.

Figure 1.2: The field of view for a NPC in AC3

The mechanic of AC3 forces NPC’s ignore the PC almost completely during a stroll through a city. NPC’s do notice and gaze towards the PC when performing ‘odd’ behaviour, for example running, climbing, galloping on a horse or fighting.

[^http://www.ubisoft.com]
NPC’s do not react to being stared down by the PC (standing still and observing a stationary character for an extended period of time). Throughout the cities there are groups of NPC’s that act as ‘hiding groups’ or blend groups (Figure 1.3). When interacting with one of these groups to hide or otherwise the groups ignore the PC except for an odd glances by chance. A Passing NPC gaze towards the PC first when entering a very close distance of the PC (approximately 0.2-0.3 m). Some odd behaviour occurs when an NPC passes the PC while running or if the NPC passes the PC too close. It has the same effect as in Figure 1.1, only at running speed it produces a comical effect with the NPC momentarily running backwards and at walking speed it causes very erratic and jerky head movements.

![Figure 1.3: Interacting with a blend group](image)

Figure 1.4 describes the difference in the area of influence that affects NPC’s in the vicinity of the player avatar. The left diagram shows the area of influence when the PC is neutral or incognito. At the distance $B$ from the player character they will ignore him, but once the distance is equal to or less than the distance $A$ they will fixate on the player avatar. The game code seems to tell the NPC to fixate its gaze on the PC, regardless of speed or minimal distance from the PC. As mentioned previously, this causes occasional odd behaviour where the NPC spins around mid stride and runs backwards for the duration the NPC is within distance $A$. The distance a NPC acknowledges the PC is very short, for the NPC the player practically materializes in thin air before him or her. Depending on the walking speed and again the distance from the PC the NPC passes, the head and eye movements appear jerky due to the FOV limitations, since the head and eyes of the NPC only fixates on a single point and tracks it.

While the PC is stationary and has a weapon unsheathed a similar situation occurs with only a few differences. The right diagram in Figure 1.4 shows that the area of influence is increased, but not gaze fixations. Primarily it has the effect of passing NPC’s loudly commenting that the PC has a weapon and that he should put it away. These NPC’s stop in their tracks at a further distance from the PC then act in two different ways. They either pass the PC in a wide circle with radius $C$ or they turn on the spot and walk away the same way they came. Those NPC’s unaffected by the drawn weapon tends to behave normally, only fixating on the PC from a further distance. The area of influence for reaction or acknowledgement of the PC do increase, but not when it comes to gaze fixation. NPC’s in this state only fixate their gaze when passing close by.

If the PC happens to carry around a body the same situation and area of influence can be observed, with a difference in comments from the passing or onlooking crowd. If the PC moves with a weapon unsheathed or carrying a body the area of influence can be seen to increase as passing NPC’s stop and fixate their gaze at the PC.
When the PC has mounted a horse NPC’s tend to notice the PC earlier regardless of speed. The minimum area of influence of the PC can be said to increase drastically. The reactions of NPC’s are more energetic, also dependent on speed they can be seen backing off or jumping out of the way.

During in-game cut scenes (scripted events) the gaze is fixed on the PC even though the head is not. This gives off an odd impression of NPC’s behaviour. Signs of stress, unease and nervousness to borderline disinterest on from the NPC. It is unclear whether the intent is to reflect emotions during these instances as it falls short.

In summary the area of influence affecting the gaze fixation of NPC’s is very narrow when the PC is stationary. It does increase slightly when a weapon is drawn and expands somewhat further when in motion. Still the overall experience is that the PC is invisible or non existent to the NPC’s until an action is taken on the players part.
1.5.2 Deus Ex Human Revolution Analysis

Deus Ex Human Revolution is a First Person Shooter game developed by Edios Montréal\(^4\) with a heavy story element. The crowd and NPC mechanic is similar to Assassins Creed 3. The game has a focus on 1 vs 1 interactions, as the amount of NPC’s are fewer and fixed, with the majority of NPC’s being named. There are no moving crowds, instead there are NPC’s in stationary groups, individuals in a looped path and stationary individuals. Most NPC’s are non essential offer limited interaction, however always physical and verbal.

Figure 1.5 shows the perceived field of view for NPC’s in the game. It is an approximation based on observations in the game. Distance and angle of the PC in relation to a stationary NPC helped determine the diagram. The angle \( A \) is the amount of degrees the head moves to focus on the player character, while \( B \) shows the maximum angle the eyes are able to view with a stationary head. Similar to the field of view of humans.

\[ \text{Figure 1.5: The FOV of an NPC in Deus Ex Human Revolution} \]

It is clear that NPC’s only focus on the PC when close enough to interact with the PC. Outside of this radius the NPC idles, gazing past the PC by chance without any noticeable interest. The distance is approximated at about one and a half the length of the PC (2.60-2.80 m). This give the impression that the NPC can sense the presence of the PC in its vicinity, but the boundary is fixed and quickly determined (Figure 1.6). As such the effect becomes noticeably static and binary as the gaze shift happens quickly (Figure 1.7). Engaging a NPC preoccupied with a task shows the NPC maintaining the task throughout the conversation. Neither turning its head nor eyes to fixate its gaze towards the PC. This behaviour does hold some merit to simulate a split focus of attention and a human-like action.

\(^4\)\url{http://www.ediosmontreal.com}
During conversation and dialogue the NPC has 'head' focus on the PC but lets the eyes wander (perhaps depending on the topic or emotional state of the NPC). The effect is more natural emotionally affected NPC which prompts empathy from the player.

Ongoing conversation between NPC’s cannot be interrupted simply by physically getting in between them. The PC is ignored completely until the scripted dialogue ends (Figure 1.8 and 1.9). After this the NPC’s does gaze towards the PC. This behaviour includes any conversation that involves the PC himself in third person. Showing that NPC’s does not notice the presence of the PC, regardless of the topic. Interrupting the conversation merely pauses it until the NPC’s has finished the dialogue with the PC. The scripted conversation picks up exactly where it left off.

When a conversation (without any dialogue options) is initiated the NPC’s head tries to turn towards the PC. If the PC stands at an odd angle in relation to the NPC, the head makes a snapping or jerky motion towards the direction of the PC, with the eyes trying to focus on the PC as well. The body of the NPC is stationary while the head can move freely to a certain point. This occurs when a NPC is stationary, facing a specific direction.
Chapter 1. Introduction

Throughout NPC encounters the eye as well as head movements does a good job to convey emotions. Fear, happiness, anger, sadness, satisfaction, frustration and stress. In some cases these changes of emotion are driven by global events, enhancing a dialogue or situation. However, no matter what global event the player has triggered, certain standard dialogue options nullifies this completely. Depending on the topic chosen for conversation the NPC might suddenly change emotional state as if the previous events had not happened at all.

In summary the area of influence of the PC is visibly clear when passing near NPC’s, but the distance at which this happens is closer to what is expected. Gaze and head movements in the game does mimic human behaviour good, except for a few details. The sense of solitude or non existence is less prominent, although player action is required to initiate any meaningful interaction.
1.5.3 Fable III Analysis

Fable 3 is a Third Person adventure/role playing game developed by Lionhead Studios\footnote{http://www.lionhead.com} with a heavy element of character development in conjunction with awareness of the PC’s renown and alignment. This game has a crowd mechanic that will swarm the PC, due to the fact that the PC is a central character (or will become one) in the game. This happens regardless of the chosen alignment.

The more renown the player earns the larger crowds will gather and as such an interesting aspect of PC - NPC interaction is available for examination. A multitude of NPC gathered around the PC provides a good chance to observe crowd behaviour in a more controlled environment than a moving crowd.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{f1_10.png}
\caption{The area of influence of the player character}
\end{figure}

If the PC is stationary NPC’s passing by will gravitate towards the PC at the distance \( A \) (Figure 1.10) and stop at 2-3 arm-lengths (1.80-2.50 m) and form a circle around the PC at approximately distance \( B \). They will either praise the PC (good alignment) or taunt the PC (evil alignment). Eventually a crowd will gather forming almost a complete circle around the PC (Figure 1.11).

The NPC’s will all consistently fixate their gaze towards the PC. If the PC is left stationary the gathered crowd will remain indefinitely and continue to grow to a full circle as more NPC’s enter the area. The game mechanic has the crowd expect displays from the PC regardless of alignment. The only time the crowd disperses is when the PC moves away, the NPC’s seems to be enthralled by the presence of the PC and cannot leave. The only exceptions are displays performed to any specific individual which causes them to leave the crowd satisfied or appalled. It gives the impression that NPC are more or less alerted of the PC entering an area and are dictated to gravitate towards the PC’s location in the area. When the PC is in motion NPC’s can be seen to fixate their gaze towards the PC with a snapping motion, either when stationary or mid strife as seen in Assassins Creed 3(1.5.1). Interacting with essential NPC’s such as shop keepers shows a noticeable delay of recognition.
Chapter 1. Introduction

In summary Fable 3 is a very crowd centric game with elements seen in both Assassins Creed 3 and Deus Ex Human Revolution. Each NPC can be interacted with equally, the crowds are large but not continuously spawned. Instead of passing by, the NPC stops to look at the PC until the PC moves on. The NPC can be said to display a trance or caught state which can only be broken by the player. Head and gaze movements are similar to AC3, but are more static in nature as looped animations.

1.6 Early Findings

A pattern that spanned across the three studied games was that no virtual agent would acknowledge the existence of the player unless interacted with first. The exception to this rule is if the virtual agent passes close enough to the player, enter the player avatar’s area of influence if you will. The odd thing is that the radius of the area of influence is very narrow, for all three games. Indeed there are differences in distances among the games, but none have a larger radius of 3 meters and beyond that radius the player might as well not exist in the virtual world of that game. On the other hand, acknowledgement of the player avatar was not consistent between the different games and even between different situations within each game.

The most telling observation was from Deus Ex Human revolution, where the NPC glances at the camera (player character) once at a distance of approximately 2 meters and then continues on (Figure 1.12). This was from an NPC that had a patrol around a floor and each time, without fail the same effect could be seen. This will become the archetype for how the animated clip for the first experiment should look like.

Other behaviours, while interesting to study, did not serve the core purpose of the project. The intent is to study virtual agent behaviour in relation to the player character and deduce a mean typical agent that could be reproduced and presented to participants in the experiment.

Based on these observations one can draw the conclusion that the average virtual agent (in recent games) is oblivious of the player avatar up to the point of direct or indirect interaction. Related conclusions are that the artificial intelligence of current non-player characters is to an extent unbalanced, where contextual awareness is high and situational awareness is low.
Figure 1.12: NPC passing the stationary player
Chapter 2

Experiment One

2.1 A Gaze at Agents Experiment

- Goal of 30 participants
- 1 clip of animated virtual agent (duration 10 seconds)
- Eye-tracking recording equipment

2.1.1 Source Material and Data

From the practical observations there was two main elements that could be translated into an experiment. The distance a virtual agent fixates onto the player avatar and where they fixate on the avatar. Consecutively from the three chosen games across three genres these two points were very similar. An average distance of 2 meters and fixation point was observed to be the head or the camera.

Through a student license it was possible to acquire Autodesk Maya as the primary source for animation and editing. A complete character model including joints was acquired through MocapClub\(^1\). As it stood it was a rather neutral body type of a male in the age bracket 35-40 (Figure 2.1. Initially a male and female model was considered, but due to the amount of work required to prepare a single mesh for animation only one was chosen.

\[\text{Figure 2.1: Source model from MocapClub}\]

\(^1\)<http://www.mocapclub.com/Pages/Characterizing01.htm>
2.1.2 Work Process

The animation was done as mentioned within the Autodesk Maya software, using a pre built model. A revamp of the model had to be done due to predefined skin weights of the model mesh and a skeletal structure not suitable for this particular animation project. Some re-texturing was done as an attempt to neutralize the model. Two additions were added to the model mesh, a left and right eyeball. An orientation constraint was added as well to be able to control the orientation. Some bevelling of the original mesh had to be done to simulate simple clothing, mainly shirt and trousers.

The re-texturing was the simplest step of the process, by only assigning a new material for the model mesh. The model had a high resolution texture, but the texture depicted the model wearing shorts and shoes. After a slight tweak of the UV-map of the mesh the new texture could be correctly applied.

The skeletal structure contained 54 joints in total, giving the model mesh good mobility and positioning when animating a walk cycle. Depending on technical constrictions or visibility of a model in-game the amount of joints can fluctuate between 46 to 60. The main differences are the hands, spine and potential accessories (such as hair, weapons or cloth).

Skin weight is the amount of influence a joint has on an area of mesh. It is one of the last steps taken before starting the animation process. It is a process wherein you assign the influence of individual joints to areas of the mesh. In other words how much the model should move when a joint moves. It can be painted or assigned individually either by joint or by vertices. A broad pass of the paint tool is usually enough per joint to have the mesh not deform when the model is flexing and moving.

The animation process is best to take in steps, animating one side of the model and then copy the values to the other side. Some animation curve smoothing was needed at the point the walk cycle looped, for a smooth transition. The walk cycle was animated only to portray a character in motion, more attention was given to head and eyes as they would be the primary focus of this experiment.

Figure 2.2: Halfway through the animation process
2.1.3 Translating Observation Into Motion

The constraints of the clip were duration and visuals. Earlier iterations had an animated virtual agent move in front of the camera in a quarter circle, as if the observer was a stationary player watching a character pass him or her in a large semi circle. It was instead substituted with a much simpler set-up. The camera (or ‘player’) was stationary facing forward, with the virtual agent passing by alongside.

When the scene starts a character in motion is seen at a distance of approximately 10 meters. For the duration of the clip the character is then seen walking towards the camera’s left edge, passing close to the observer. At a distance of 2 meters the character fixates on the camera, as if acknowledging the player’s existence.

Once boiled down to the essentials, this set-up is what was observed in the study. By removing potential visual distractions and having the virtual agent against a blank background the chance of reliable data increases. As a parallel, this is supposedly what the virtual agent sees, only with the acknowledgement happening earlier on. This rather simple set-up was still the most efficient and telling of what the earlier observations of games had resulted in. It was very similar in layout as this figure 2.4 as seen in the analysis of Assassins Creed 3, as well as the sequence seen in figure 1.12 in the Deus Ex Human Revolution analysis.

![Figure 2.3: Finished scene](image)

Figure 2.3: Finished scene

![Figure 2.4: Sequence of Experiment one, top down](image)

Figure 2.4: Sequence of Experiment one, top down
Chapter 2. Experiment One

2.2 Experiment Setup

The goal for number of participants for the experiment was set at 30 to get a reliable dataset. The experiment was designed to record how and where a person would look when facing another human like character. It consisted of a short instruction 'View the following clip', followed by a clip of an animated character walking towards the camera to the left side. The clip was ten seconds long and the experiment ended with an acclimation of gratitude for the participation.

Each test took on average two minutes including calibration, making participation in the experiment much easier. The experiment was performed with the Tobii T60 Eye tracker using the Tobii studio software present on the machine. Each participant was given an ID in the form of EPnn (experiment participant number) as well as two variables. Gender for analysis of gaze data and eyesight to eliminate any data loss due to reflections of glasses or lenses.

Prior to every recording a calibration of the software was done with each participant to maximize data collection. Since the purpose was to gather a large quantity of data no variations of the experiment were designed, even so each participant was brought into the lab one at a time. 35 participants in total were tested in the first experiment. 34 participants were recorded with high enough data validity for analysis. Out of the 34 participants 14 were female and 20 were male, a satisfactory balance of participants which provided a good foundation for further comparative studies apart from only gaze data. The set goal of participants for the experiment was 30, 15 females and 15 males.

Complete transcript for the participants in this experiment can be found in the appendix A.1.

2.3 Compilation of Data

For each recording the number of fixations, coordinates for fixations, timestamps, duration and area of interest was documented in a raw data file. Fixations are the individual data points where the eye stops and focuses after each gaze shift. With the aggregated data several different analyses could be performed. The most prominent one was to determine fixation per area of interests in preparation for the next experiment. The initial sift through the data showed no difference in data collection based on eyesight. There was an initial concern if glasses would affect the data collection since the hardware relies on being able to reflect infrared light of the participant’s iris.

The average percent of samples found was 85.7 for the female group and 88.95 for the male group. Total amount of fixations for the female group was 227, for the male group 241. The highest amount of fixations was 31 among the females and 20 among the males, while the lowest amount of fixations was 10 and 5 respectively.

The average amount of fixations was 16.21 and 12.05 for the female and male group respectively.

2.3.1 Fixations

Over a duration of time fixations form a gaze plot numbered in increasing order. These are generated by the Tobii studio software. With each fixation comes several data points.
The chronological number of the fixation, the X and Y coordinates of the fixation for both the left and right eye (the Tobii T60 Eye-tracker hardware is fixed at a resolution of 1024x768 px in a 4:3 ratio), a timestamp for the fixation, the duration start for the fixation (same data point as the timestamp) and the duration of the fixation. The average amount of fixations for all participants was 14 (13.74).

<table>
<thead>
<tr>
<th>Group</th>
<th>Head</th>
<th>Torso</th>
<th>L Arm</th>
<th>R Arm</th>
<th>Pelvis</th>
<th>Legs</th>
<th>Feet</th>
<th>scene</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>77</td>
<td>47</td>
<td>3</td>
<td>2</td>
<td>16</td>
<td>14</td>
<td>8</td>
<td>60</td>
</tr>
<tr>
<td>M</td>
<td>83</td>
<td>47</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>11</td>
<td>3</td>
<td>69</td>
</tr>
<tr>
<td>Both</td>
<td>160</td>
<td>94</td>
<td>13</td>
<td>10</td>
<td>26</td>
<td>25</td>
<td>11</td>
<td>129</td>
</tr>
</tbody>
</table>

Table 2.1: Number of fixations per area of interest

2.3.2 Timestamps and Duration

Each fixation in the replay was given several timestamp values. The duration start of the fixation (from the start of the recording in mm:ss:ms) and total duration of the fixation (in mm:ss:ms). With this data it was possible to determine how long the mean duration was for each area of interest. Both divided among the groups and as a total factor for the mean gaze plot.

<table>
<thead>
<tr>
<th>Group</th>
<th>Head</th>
<th>Torso</th>
<th>L Arm</th>
<th>R Arm</th>
<th>Pelvis</th>
<th>Legs</th>
<th>Feet</th>
<th>scene</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>83.3</td>
<td>17.32</td>
<td>1.51</td>
<td>0.65</td>
<td>8.12</td>
<td>5.24</td>
<td>3.01</td>
<td>15.36</td>
</tr>
<tr>
<td>M</td>
<td>144.1</td>
<td>18.43</td>
<td>3.07</td>
<td>2.97</td>
<td>3.99</td>
<td>3.16</td>
<td>1.53</td>
<td>17.59</td>
</tr>
<tr>
<td>Both</td>
<td>227.4</td>
<td>35.75</td>
<td>4.58</td>
<td>3.62</td>
<td>12.11</td>
<td>8.4</td>
<td>4.54</td>
<td>32.95</td>
</tr>
</tbody>
</table>

Table 2.2: Duration in seconds per area of interest

2.3.3 Area of Interest

Area of interest (AOI) is simply a term for defining parts of the scene where participants will look. The amount of AOI's was based on the work of McDonnell[15]. In their study they used a total amount of 14 AOI’s for a virtual character. This experiment has the goal to differentiate broader gaze patterns and as such the amount of AOI’s was reduced to \(7 + 1\) (Head, Torso, Left Arm, Right Arm, Pelvis, Left Leg, Right Leg, Feet and 'Scene' or 'not on character'). Each replay from a participant was examined and each fixation was tagged with an AOI. This was later used to produce a mean gaze plot or a probability table for fixations. It will be discussed further in section 2.4.

2.4 Mean Gaze Plot

For the second experiment a modified variant of the clip used in the first experiment had to be produced. The first objective was to determine the amount of fixations for the mean gaze plot. The mean fixation per participant is a factor of total fixations over total participants. It was worked out to be 13.76, rounded up results in a mean fixation of 14. For the purpose of the experiment the number of fixations was set to 14 over the clip duration of 10 seconds.
Next step was to determine what areas would be fixated upon. Two different methods was used to produce a mean gaze plot or probability table. The first matched each fixation in order, from 2 up until 15, 14 fixations total. This was done for all participants, tallying all areas of interest.

The reason for starting from fixation number 2 is that the first fixation for all participants is a lingering fixation from a previous segment of the recording. Thus it cannot reliably account for intentional gaze for the participant. The highest area of interest for each fixation was labelled as the highest probability and the second highest as a lower probability and so on in a descending order. This method proved inconsistent however. Since only the fixations 2 to 15 was included, data from participants with more than 15 fixations were left out of the mean gaze plot table.

The second method tallied each change of area of interest instead. With this more data from each participant could be utilized. Each new area of interest got tallied up in the total, skipping duplicate areas of interest. By applying this method more fixations were able to be included in the total, including more data points and giving more validity to the mean gaze plot table. There were six steps of probability, in table 2.3 below only steps one through four is shown. For each step the probability was lower or missing. As such, beyond column four the data becomes irrelevant.

<table>
<thead>
<tr>
<th>Fix #</th>
<th>AOI prob 1</th>
<th>AOI prob 2</th>
<th>AOI prob 3</th>
<th>AOI prob 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Head</td>
<td>Torso</td>
<td>L Arm</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Torso</td>
<td>Legs</td>
<td>Pelvis</td>
<td>Head</td>
</tr>
<tr>
<td>3</td>
<td>Head</td>
<td>Torso</td>
<td>Legs</td>
<td>Feet</td>
</tr>
<tr>
<td>4</td>
<td>Torso</td>
<td>Head</td>
<td>Pelvis</td>
<td>Legs</td>
</tr>
<tr>
<td>5</td>
<td>Head</td>
<td>Torso</td>
<td>Feet</td>
<td>R Arm</td>
</tr>
<tr>
<td>6</td>
<td>Head</td>
<td>Torso</td>
<td>scene</td>
<td>Pelvis</td>
</tr>
<tr>
<td>7</td>
<td>Head</td>
<td>scene</td>
<td>Torso</td>
<td>L Arm/Legs</td>
</tr>
<tr>
<td>8</td>
<td>Head/scene</td>
<td>Torso</td>
<td>Legs</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Head</td>
<td>scene</td>
<td>Torso/Pelvis</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>scene</td>
<td>Head</td>
<td>Torso/Pelvis/Legs</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Head</td>
<td>Torso</td>
<td>L Arm</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>scene</td>
<td>Head</td>
<td>Torso</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Head/scene</td>
<td>L Arm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>scene</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 2.3: Mean Gaze Plot chart*

With the help of this chart a gaze sequence for the virtual agent could be created. Each fixation lasts for a fraction of the duration of the clip divided by the amount of fixations. The mean gaze duration becomes approximately 0.71 seconds which correlates to the total mean duration of fixations for the experiment. The sequence differs from the chart due to the frequency of multiple fixations on one area. While humans focus on facial features, they tend to not linger too long or shift back and fourth to the same spot. A mix of column 1 and 2 helps give the sequence a more dynamic natural appearance.
The final sequence

1. Torso
2. Legs
3. Torso
4. Head
5. Torso
6. Torso
7. Scene
8. Torso
9. Scene
10. Head
11. Torso
12. Head
13. Left Arm
14. Scene

Some modifications, such as smooth head movements and duration changes had to be done. Fixing each shift in gaze to a maximum duration of 0.7 seconds would look very unrealistic and jerky. Some tweaking were added based on the chart guidelines. The choice to add gaze point off character was partly due to the chart, but also due to human nature. The inherent unpredictability of gaze shifts and the difference in salience triggers among people.
3.1 A Subjective Gaze Experiment

- Goal of 30 participants
- 3 variations of clips with animated virtual agents (duration 3 x 10 seconds)
- Likert scale questionnaire
- Eye-tracking recording equipment

3.2 From Data to Visuals

The standard current virtual agent was intended to be used again as a benchmark, so no further modifications were made to the existing clip. Two other clips to compare and rate had to be created though. Using the results of the mean gaze in 2.4 it was a fairly simple matter to plot out which body part in succession a person is most likely to look at a passing character. Having the length of the clip determined prior, the order of the gaze plot and the average duration for each point the work to translate that into the character began.

In the Autodesk Maya software a camera was mounted at the eye position of the virtual agent (facing the scene camera). A duplicate mesh of the character was put in place where the scene camera was to represent the participant. Gaze point by gaze point the orientation of the eyes of the character walking towards the scene camera was edited to match the data. Naturally the head movement had to be adjusted as well to mimic what a human would do as not to appear awkward or strained. This caused the virtual agent to appear to behave as if he was examining the observer as he passed by. With the data from experiment one visualized in the virtual agent there was, visually a clear difference in virtual agent gaze behaviour and recorded human gaze behaviour.

Another modification to the original clip had to be created as a control. The eye and head movements were animated in a jerky unnatural manner, with a short duration between gaze shifts. The entire character was made to move in a different manner as well. While the average duration for each fixation was worked out as well as the amount of fixation per clip in 2.4, some modification was done as well. Amount of fixations remained the same for the clips, however the duration was adjusted to be more spread out and contracted depending on which area was focused upon.
Having multiple clips to show the participants lets the participants know less about the intent of the experiment. By having three variations of the clips each new scene presented will provide a new perspective for the participants. If only two clips were being presented each participant might get too much insight into the experiment, a similar obstacle as with the questionnaire discussed in 3.5. The third clip would serve on one hand as an expected outcome, since the clip was designed to stand out, the results of the eye-tracking data and questionnaire results would support that. On the other hand it would serve as a clean slate, in between, before or after the other two clips.

Clip A was the representation of the standard virtual agent, Clip B was the control clip and Clip C the animation created based on the data from the first experiment.

### 3.3 Likert or Not

This was a project with focus on what one perceived as realistic or how something can be considered or felt realistic. To get a value of this it requires some sort of subjective test. Simply using a yes / no questionnaire would not do to get reliable results.

Widely used in many fields is the Likert test or scale. It lets participants rate the agreement or disagreement of a statement or question[1] simply put. The most common way to use a Likert scale in gauging statements is a five point scale where the middle point is neutral. There are variation as low as four points up to ten. The choice between an odd number of points and even is a hard one[1]. You do not wish to receive data that gives vague or no inclination to either of two opinions if you have an odd number of points. Neither is it positive to force an opinion from a participant by denying them a middle ground using an even number of points.

For this project it was decided to have six points, there are strong recommendations for using odd scales\cite{11}\cite{12}, however they do come with potential hazards. Choosing an even scale implies that the forced choice method is used. Using an odd number scale does even out the distribution of answers, but the ratio of pro and con responses remain the same regardless of odd or even scales\textsuperscript{2}.

Some argue that the prior knowledge of a subject influence the result, meaning that if the participant is aware of what is investigated and how it works, the participant should be able to categorize his or her answer in a positive or negative manner\cite{6}. The argument is that in the context of this experiment the survey covers both knowledge (how to decide whether something is realistic or not) as well as attitudes (if you thought it was realistic, how realistic was it?)\textsuperscript{3}.

In some extent it is up to the survey questionnaire designer whether to use an odd or even scale. With a higher than average number of points more variations would be visible and perhaps even more interesting conclusions could be drawn when correlating the Likert data with the visual data. Having an even number of points on the scale was to make sure to get an inclination towards one of the answers. Any neutral data in this test would neither gain or deprive the project, as such an even number of points seemed reasonable.

Apart from the number of points in the scale it is imperative to pose the correct statement otherwise the answers will be skewed. The actual reasoning and choice of statement is discussed in section 3.5.

### 3.4 Opinions with Gaze

Two variations of the experiment was available late in the project. Either have participants watch three clips on a computer screen and after each clip fill out a form. The other variant had participants once again enter the eye tracking lab and view clips while being recorded and fill out a form digitally. After some consideration the latter design was chosen for the experiment. This was the case for several reasons. The random but specific order of clips to be shown are easily documented and edited in the available Tobii studio software. A digital form is just as easily created and edited from within the software to perform single and multiple choice questions. It is also more convenient and appropriately placed time wise in relation to each clip, instead of a physical copy. This was how the Likert questionnaire was incorporated after each clip.

An added benefit is that eye tracking data can be recorded once again and provide arguments for the results. This would also give interesting data about gaze behaviour across both experiments. Finally, another strong argument for having the experiment in the lab is the exclusion from external distractions, in a hallway or public space there are a lot of things that could easily distract the observer which would affect the outcome.

\textsuperscript{2}\url{http://blog.evansanalytics.com/2011/09/placing-odd-vs-even-number-of-points-on.html}

\textsuperscript{3}\url{http://blog.evansanalytics.com/2011/09/including-dont-know-not-sure-not.html}
3.5 Experiment Setup

The goal for number of participants for the experiment was set at at least 30 participants. For this experiment three clips were used. Clips A, B and C were randomized into three sets of three. Clip A was the original 'game based' clip used in the first experiment. Clip B was the non-realistic control clip. Clip C was the 'realistic' clip, built from the gaze data gathered and analyzed from the first experiment. The experiment had three different variations in order to get as objective results as possible. All three sets got randomized by using an online random generator\(^4\). Set number one received the order B-C-A, set number two were A-B-C and set number three C-B-A. This was an unfortunate randomization as section 3.7.2 discusses further.

Each test was constructed identically with the same instructions and questions, except for the order of clips. To get an even spread of participants for each test, every new participant started with a different test. Participant number one was given test one, participant two test two and so on. By having a new test for each new participant all set of clips had a a good chance of getting viewed objectively and any margin of errors could be decreased if the question would be either to vague or leading.

The challenge was to formulate a relevant to-the-point question. The question cannot be to telling, neither too vague or general. Some inspiration for the type of question was found in the work of McDonnell[14]. In her work it was the formulation of the scale that helped get an answer. It did however very much help with the general idea of what the question should convey to the participant without telling them outright. The scale span was decided to be 6. The main reason behind this was to reduce the chance of participants choosing a neutral grade in an odd scale. A very common span for a Likert scale question is 5.

The initial phrase was 'The character who passed me displayed a realistic behaviour'. This statement while relevant, encompassed more than just the gaze of the character, which caused participants to scrutinize and rate the overall scene in front of them. Due to this the focus of the experiment got lost in the perceived realism of the scene. Instead of rating how they felt observed by the character, they rated the overall scene. In short the statement 'The character who passed me displayed a realistic behaviour', would not be accurate enough. The discrepancy was noticed early on of the experiment based on comments from participants. The experiment was put on hold until a statement that showed what the test ought to achieve could be formulated.

The statement 'I felt the character realistically acknowledge me as a person' was chosen. With an emotional statement like 'I felt[...]', the answer would more thoroughly give an indication to whether or not the participant felt like he or she was watching a clip of an animated character or if they felt like they were being observed by a person. The difference was noticed rather quick as participants were made aware the perception of the character instead of the motion or unrealistic rendering.

There is always a risk of the question being to leading, causing each participant to be more observant after each clip. Normally such an issue can be reduced by having multiple tests of different order. The three different sets ensures that the first clip of each test gets an objective result of a subjective experiment. See the table A.1 in the appendix for a participant index.

\(^{4}\text{http://www.randomiser.org}\)
3.6 Compilation of Data

A total of 30 participants were able to participate in the experiment. In the subjective gaze experiment there was not as high of a demand for an even distribution among gender groups as the first experiment. Among the participants 7 out of 30 were female, so a group comparison would be possible to perform if so inclined. In this experiment 100% of the data could be analysed.

Similar data point was recorded and documented as in the gaze at agents experiment apart from the Likert results. Number of fixations per participant, coordinates of fixations, timestamps, duration as well as area of interest. As the experiment was once again performed with the eye tracker this data could be collected. The total amount of fixations recorded totalled at 1052 of 30 participants at an average of approximately 35 fixations per participant. Total amount of duration was 878.34 seconds with the average duration per fixation at 0.83 seconds.

<table>
<thead>
<tr>
<th>Area of Interest</th>
<th>Fixations</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>371</td>
<td>626.03</td>
</tr>
<tr>
<td>Torso</td>
<td>169</td>
<td>57.27</td>
</tr>
<tr>
<td>L Arm</td>
<td>27</td>
<td>9.22</td>
</tr>
<tr>
<td>R Arm</td>
<td>20</td>
<td>8.56</td>
</tr>
<tr>
<td>Pelvis</td>
<td>38</td>
<td>13.82</td>
</tr>
<tr>
<td>Legs</td>
<td>27</td>
<td>13.24</td>
</tr>
<tr>
<td>Feet</td>
<td>12</td>
<td>5.68</td>
</tr>
<tr>
<td>Scene</td>
<td>388</td>
<td>144.52</td>
</tr>
</tbody>
</table>

*Table 3.1:* Fixations and duration per area of interest

3.7 Likert Results

Table 3.2 has the lineup of the amount of answers in total, the values are tallied from all participants, regardless of the order they have seen the clips in. The majority of the positive responses can be seen located with ‘clip C’ or the realistic animation.

At the same time the majority of the negative responses can be seen converge at ‘clip B’ which was the exaggerated animation, as well as ‘clip A’. Lastly ‘clip A’ can be seen having a rather high neutral response with results collecting among the middle grades. Coupled with this it should be noted that there are few strong pro or con values in the data.

<table>
<thead>
<tr>
<th>Viewed clip</th>
<th>(1) I strongly disagree</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6) I strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current (A)</td>
<td>1</td>
<td>4</td>
<td>12</td>
<td>6</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Control (B)</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Realistic (C)</td>
<td>0</td>
<td>2</td>
<td>8</td>
<td>9</td>
<td>7</td>
<td>4</td>
</tr>
</tbody>
</table>

*Table 3.2:* Likert results with the statement: "I felt the character realistically acknowledge me as a person"
3.7.1 Likert Analysis

First we compare the groups with each other. Group one and two, two and three and finally three and one. The first comparison, (group 1+2) had the median answer 3.5 for clip A, 3.4 for clip B and 3.9 for clip C. The second, (group 2+3) had a median value of 3.3 for clip A, 3.4 for clip B and 3.9 for clip C. Finally the third, (group 3+1) had a median value of 3.6 for clip A, 3.6 for clip B and 4.3 for clip C. The table 3.3 below shows the percentage of the answers.

<table>
<thead>
<tr>
<th>Scale rating</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clip A</td>
<td>3.3%</td>
<td>13.3%</td>
<td>40%</td>
<td>20%</td>
<td>20%</td>
<td>3.3%</td>
</tr>
<tr>
<td>Clip B</td>
<td>13.3%</td>
<td>16.6%</td>
<td>23.3%</td>
<td>16.6%</td>
<td>26.6%</td>
<td>6.6%</td>
</tr>
<tr>
<td>Clip C</td>
<td>0%</td>
<td>6.6%</td>
<td>26.6%</td>
<td>30%</td>
<td>23.3%</td>
<td>13.3%</td>
</tr>
</tbody>
</table>

*Table 3.3: Percentage of answers per clip*

When structuring the data by clip instead, the median and mode turns out different as seen in table 3.4.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>3</td>
<td>3.5</td>
<td>4</td>
</tr>
<tr>
<td>Mode</td>
<td>3</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

*Table 3.4: Median and Mode per clip*

There are several different ways to analyse Likert questionnaires[1]. Three analyses are needed as each group is a separate sub experiment of the Subjective gaze experiment. Due to the structure of the test and the scrambled order of clips several comparisons of Likert sets is needed. The analysis is structured by group and clip. The result is more tables and analysis data, however it is necessary to keep the data comprehensible and structured. The results chapter 4 has a more thorough breakdown of the data analysis and the meaning of it. It also goes in deeper and analyses the Mean and Standard Deviations of the Likert results for each clip.
3.7.2 Correlation of Numbers and Opinion

Before the results are discussed there are some interesting aspects to mention. Perhaps the most interesting branch in the data from this experiment is to look more closely at how the responses for clip C is depending on the order the clips are viewed. More specifically, if the amount of fixations and duration on the head area has any correlation to the Likert responses.

Due to unfortunate lack in scrutiny of the experiment design phase there is an inherent bias in the experiment as not all clips have three positions. This has resulted in a skewed distribution of the three groups or variation of clip order. As it stands the data in total can still be analysed, however some variables will remain unseen.

The order of the clips in each group needs to be considered when examining the table 3.5 below, it is important to know what clip precedes and succeeds to understand the difference in values.

Group order
1. B - C - A
2. A - B - C
3. C - B - A

The properties of each clip as mentioned in 3.2 are slightly different. Clip A represents the standard virtual agent, Clip B is the unnatural control and Clip C represents the enhanced virtual agent.

The theory is that the clip viewed first might on one hand be the most valuable sources of data as the observer is unaware of what to view or expected to look at. After the first clip the questionnaire gives some hints to what the experiment intends to find. The question is still vague enough to not elude any instructions for viewing. Similarly the second clip might be even more valuable as a data source. The observer has a hint as to what the experiment seeks, but not clear enough to skew the result.

A high amount of fixations might just as well mean that the observer has seen something unusual or abnormal. It might just be an interesting factor that indicates a point of interest.

Clip A has one primary and two tertiary positions.
Clip B has one Primary and two secondary positions.
Clip C has one primary, one secondary and one tertiary position.
Table 3.5 is a summary of the fixations and duration of the head area, as the focus of the survey has been just that. With these numbers correlations can be drawn with the Likert results.

In the primary position clips A and B shows a higher amount of fixations than clip C. The secondary position has clip B decline roughly half of the fixations while Clip C shows a sharp increase in fixation. In the tertiary position clip C reduces its fixations to the same level as its primary position. This happens to clip A as well in one instance, while in the other it elevates to the values of its primary position.
Chapter 3. Experiment Two

Table 3.5: Fixations and duration of the head area

<table>
<thead>
<tr>
<th>Clip</th>
<th>Fixations</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clip A</td>
<td>56</td>
<td>59.9</td>
</tr>
<tr>
<td></td>
<td>23 &amp; 57</td>
<td></td>
</tr>
<tr>
<td>Clip B</td>
<td>52</td>
<td>74.11</td>
</tr>
<tr>
<td></td>
<td>26 &amp; 39</td>
<td></td>
</tr>
<tr>
<td>Clip C</td>
<td>34</td>
<td>61.07</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>34</td>
<td></td>
</tr>
<tr>
<td></td>
<td>67.9</td>
<td></td>
</tr>
</tbody>
</table>

The fixation values of clip A in group one follow that of a linear decline as its values are halved, while in group three they remain almost constant. The duration values both increase in group two and three. The fixation values of clip B mimics the behaviour of clip A in group one and two, but in group three it can be seen not to drop as sharp. The duration values for clip B decline in group two as the fixations, but remain almost constant in group three. The fixation and duration values of clip C can be likened to a symmetrical bell curve.

In its secondary position clip C seems to draw the most attention as seen in the increase of fixations compared to the others. This is only true though when clip C is preceded by clip B in the first group. A possible explanation for the sudden rise in fixations for clip A in its tertiary position might be that it preceded clip C. Clip B seems to only have highest attention when it is viewed in its primary position.
Chapter 4

Results

4.1 Results from Experiment One

In this section the results not directly involved in the creation of experiment two are brought up.

4.1.1 \( t \) test

Additional analysis of the data could be done thanks to the amount of data. This test was performed to examine the difference of male and female fixations. The amount of fixations for each group was compared. The data was calculated using an online \( t \) test resource called GraphPad\(^1\), it is an unpaired \( t \) test.

Null hypothesis: There is a difference of perception between females and males.

P value and statistical significance.

The two-tailed P value equals 0.0282, by conventional criteria this difference is considered to be statistically significant.

Confidence interval

The mean of F fixations minus M fixations equals 3.9

95% confidence interval of this difference: From 0.45 to 7.35

Intermediate values used in calculation

\( t = 2.3023 \)

\( df = 31 \)

Standard error of difference = 1.693

<table>
<thead>
<tr>
<th>Group</th>
<th>F fixations</th>
<th>M fixations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>16.21</td>
<td>12.32</td>
</tr>
<tr>
<td>SD</td>
<td>5.73</td>
<td>4.01</td>
</tr>
<tr>
<td>SEM</td>
<td>1.53</td>
<td>0.92</td>
</tr>
<tr>
<td>N</td>
<td>14</td>
<td>19</td>
</tr>
</tbody>
</table>

*Table 4.1: \( t \) test table*

With the value received from the \( t \) test the null hypothesis holds true. In other words, based on the data collected it is reasonable conclude that females has more fixations than males in this context.

\(^1\langle http://graphpad.com/quickcalcs/ttest1.cfm\rangle\)
4.2 Results from Experiment Two

The order of the clips seems to have had some effect on the outcome, letting participants perhaps know too much about the experiment and skewing the results. The unfortunate experiment design might have influenced the data as well, however the goal is to compare the total amount of opinion replies. Only looking at the answers might provide a misleading hint and would not be enough to conclusively state neither a positive nor negative support of the hypothesis.

4.2.1 Mean and Standard Deviation Analysis

Figure 4.1: The Mean and Standard Deviation of the Likert results for each clip

Figure 4.1 shows the mean for each clip as well as the standard deviation from each of the means, while Table 4.2 lists the values. In this context a higher mean shows a higher likelihood of a positive response and a lower mean would therefore signify a negative response. The standard deviation tells us how much the values are spread out from the mean, a lower value would show that the mean is more reliable or rather that there is less spread among the values. A higher standard deviation would show signs that the values fluctuate a lot more and that perhaps the mean is less reliable. Clip A has a mean of 3.5 and the standard deviation from the mean is 1.6386144975. Clip B has a mean of 3.4 and the standard deviation is 2.1932498325. Clip C has a mean of 4.1 and a standard deviation from the mean of 1.4069776853.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>3.5</td>
<td>3.4</td>
<td>4.1</td>
</tr>
<tr>
<td>Stdv</td>
<td>1.638</td>
<td>2.193</td>
<td>1.406</td>
</tr>
</tbody>
</table>

Table 4.2: Mean and Standard Deviation table

C has the highest mean which would support the hypothesis, it also has the lowest standard deviation which lets us know that the mean is more reliable. B has the lowest mean and the highest standard deviation, this tells us that B has high and low values at both ends of the spectrum. This could be said to be expected from a control object. Neither a strong positive or negative result. Finally A has a marginally higher mean than B as well as a marginally higher standard deviation than C. While the mean is low, it is still fairly reliable as with C.
The difference in standard deviation between the clips is not relevant to compare. It only shows how spread out the values for each clip is. The means are the important values to keep an eye on. The questionnaire used had an even scale so there is no middle point value, 1 through 3 are considered con answers and 4 through 6 are considered pro answers with a virtual midpoint of 3.5.

It can thus be said with confidence that in rising order of perceived realism comes first clip B followed closely by clip A and finally clip C. These results are not statistically significant enough to support the hypothesis, however they do indicate a tendency. With a different experiment design and a new set of trials may provide more statistically significant results.

4.2.2 Statistical Analysis

Further statistical analysis is done to get a better understanding of the data. For Likert data there are a few options on what to use, as well as what not to use[3][9]. There is an ongoing debate in the social sciences about how to go about analysing Likert data and how to treat it. Either as ordinal or interval data. Likert scales results cannot be assumed to be normally distributed. Each sample is related and paired and as such a non-parametric test needs to be used.

The open source software R was used to calculate this². The software R lets one use several kinds of statistical analyses, it has been very helpful in determining statistical significant differences when analysing the data gathered. Other resources online has been helpful in finding the right test and how to go about analysing it (socialscience-statistics.com³ as well as Johnny Deng⁴).

The Kruskal-Wallis test tells us first whether there is any significant differences between (two or more) groups, but not where. So to start off the data will be run through the Kruskal-Wallis test.

Within the context of this test a p-value greater than 0.05 signifies that we accept the null hypothesis, which is that there is no significant difference between two sets of groups. In contrast a p-value less than 0.05 shows that there is a significant difference in the values of two groups.

For the Kruskal-Wallis test the data was listed in a table with 3 columns where each row shows the response of a participant for clip A, B and C (3 columns, 30 rows). In R the data for each column was tied to a variable, the 3 variables (A, B and C) were then listed into another variable (columnlist). With this setup the Kruskal-Wallis test could be performed.

²<http://www.r-project.org>
Kruskal-Wallis rank sum test
Kruskal-Wallis chi-squared = 3.7766, df = 2, p-value = 0.1513

This result tells us that between the 3 groups there is no statistical significant difference. Further down each group is compared in pairs.

The statistical analysis cannot help support the hypothesis.

The code used to calculate in R can be found in section A.3 in the Appendix.

4.3 Overall Results

From both experiments it is clear that humans tend to seek eye contact with another human like character. The amount of fixations around the head area are significantly higher than any other are throughout both experiments (see Table 4.3 and Table A.2), when including the areas of the character. Discounting the fixations not focused on the character (scene fixations) the rest of the areas follow this order Torso, Pelvis, Legs, Left Arm, Right Arm and Feet.

The total amount of fixations throughout both experiments is 1520 across 64 participants. This gives an average of approximately 24 fixations per participant. The total duration is 1207.69 seconds which gives an average duration per fixation of approximately 0.79 seconds.

The heatmap 4.2 below is a visual representation of the data. The fainter the colour the lower the percent. Table 4.3 below is an approximation of the percentage for each area. Calculation of the data can be found in the appendix A.2.
Figure 4.2: Heatmap of fixations

Table 4.3: Approximate percentage of fixations per area across both experiments
Chapter 5

Discussion

To some the aforementioned experiments might only reinforce previous assumptions about realism in games, movies or robotics. While this is not false neither was it the only goal of the project. It is a prominent presence in games that the player is mostly led through an interactive story with limited impact and effect. This is of course one of the many limitations of storytelling in a confined linear medium. There are however tricks to dispel the linearity and induce a more realistic and immersive feeling in such a setting.

Most of the related research have dealt with identifying realism in characters through participant studies, determining the current position of the precipice of the uncanny valley or to some extent having virtual agents react and adapt to the gaze of a participant or even trick a participant (this field includes but is not limited to games as applications). Gaze prediction based on saliency maps heralded from cognitive science are indeed fairly accurate and potentially useful, if not less resource intensive. They are however still predicted model and to an extent predefined and predictable patterns unlike the human nature of gaze patterns.

The claim lies not that these models are by any means inferior to actual gaze data models as a comparison has not yet been performed. Little of the current research deals with adopting human gaze to virtual agents or determining the realism of virtual agents enhanced with human gaze.

Eye-tracking as an input to games, applications or devices has been around for some time. It has not been common to use eye-tracking technology to record motion in the way motion capture is used. The process is as of now rather time consuming, and bulky if a project is in need for individualized eye motion. The positive aspect is the quick set-up and calibration, the downside is the amount of material ‘to be looked at’ for the actors in each particular case.

One interesting aspect is the current industry opinion regarding this particular subject. While as noted limited research has been done there are voices airing the current lack of immersion in games. Even going so far as comparing self proclaimed story heavy games of being but dioramas without any ability to affect few or any events [4].

5.1 Conclusion

In general people examine a character in a jagged pattern (see Figure 5.1). For example, Head, torso, head, legs, head, head, head. From a distance the observer tries to take in the entire character as quick as possible. At closer distances the observer tends to focus on primarily human features like the face and head area. The curve resembles the symbol for square root very distinctly.
Regarding the subjective opinion on realism it seems that the effect of enhancing a virtual character with human eye and head motion indeed noticeable. The effect has not been comparable to moving closer to the cliff precipice of the uncanny valley, but neither has the response been highly positive. It is a subtle response of positive feedback when encountering a virtual agent that has eye motion and behaviour similar to the observer. These experiments have not been able to distinctively show that an enhanced virtual agent is more realistic than what current virtual agents display. This is not to say that the experiment has failed or been performed badly. The subjective opinion on what is considered realistic is just that, subjective. It is difficult to quantify a subjective reaction to stimuli using only questions or ratings. Such an experiment could have benefited more with either an EEG test or a galvanic skin response that might more truthfully reveal an observers reaction.

In the long run, a combination of technology including the dynamic dialogue system mentioned by Ruskin, with an increased area of influence for the player avatar, visually affected and altered appearance of the player as damaged clothes or wounds and virtual agents capable of examining each other and a player by human like gaze would open up for more realistic scenarios.

Imagine the player walking down a crowded street with NPC vendors. Just as you are able to distinguish individuals from a crowd so would they depending on the immediate visual need that your avatar would convey. You are carrying a bow and quill with no visible arrows, suddenly a vendor cries out from across the street through a gap in the crow while fixating on your character 'I've got arrows for you!'. In another scenario your character is visibly wounded or the armour he or she is wearing is cracked or damaged, a dug vendor or a smith might have seen you from a distance and approached you to inform you of their wares and services. The interaction between the player avatar and virtual agent will no longer be initiated by the immediate presence or voluntary interaction on the players part, but by the fact that the virtual agents are aware of the presence of the player avatar by sight.
5.2 Further Research

An experiment to test participants EEG or galvanic skin responses to similar stimuli would be an interesting compliment to this experiment. A higher number of participants could also give more insight as to whether subjective realism can be measured in a scientific way.

There are several ways to move on in this field with this project as a basis. One branch that would be the most interesting is to create a reliable model from eye motion in virtual agents that can compete with predicted eye models based on saliency maps. Such a model could then be implemented into characters to enhance the human behaviour of them. It would not necessarily change the artificial intelligence of a game, but add another depth to its characters and repertoire of tricks for the player to experience. All characters that you encounter do not behave in a similar manner, neither in game nor real life. However those characters that do matter should be able to help to establish a certain level of immersion.

There is a potential to develop safety gear with strategically placed information based on where a person is more likely to look on a human body. Today there is already gear with the most important information on or at head level, a good example is a firemans helmet. The information could also be used to draw the focus from a certain area using colour or shapes to where one wishes the focus to be. If you know the probability of focus, then it is possible to prepare or account for it.

A different approach would be to apply the model to robotics research developing androids or similar human like constructs to let them more easily recognize and examine humans beings. There are interesting developments in the field of virtual agents that examines patients through a web camera and registers data. SimSensei\(^1\) developed by the University of Southern California (USC) at the Institute for Creative Technologies (ICT). In this case the software is able to pick up human behaviour and fairly accurately determine them. However, reversed it still gives of an uncanny valley presence that may push potential patients away.

\(^1\)\url{http://ict.usc.edu/prototypes/simsensei/}
Appendix A

A.1 Participants and Data Points

Table on the Left tallies the Gaze at Agents experiment.

- Eyesight: 1: No glasses 2: Glasses 3: Lenses
- Sex: 1: Female 2: Male

Table on the right tallies the Subjective gaze experiment.
The variation column tells which order the participant viewed the clips.

- Eyesight: 1: No glasses 2: Glasses 3: Lenses
- Sex: 1: Female 2: Male
- Variation: 1: B-C-A 2: A-B-C 3: C-B-A

See table A.1.

A.2 Fixations and Percentage

The raw data table for total amount of fixations. Each area is divided by the total to get an approximate percent of the total. This data is comprised from both experiments. See table A.2.
<table>
<thead>
<tr>
<th>P ID</th>
<th>Eyesight</th>
<th>Sex</th>
<th>Group (1-3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP01</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EP02</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EP03</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP04</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EP05</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP06</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EP07</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP08</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP09</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP10</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP11</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP12</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP13</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP14</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP15</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EP16</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP17</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP18</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP19</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP20</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP21</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP22</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP23</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EP24</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EP25</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EP26</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP27</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EP28</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP29</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EP30</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EP31</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EP32</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>EP33</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EP34</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>EP35</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Table A.1: Left table: Gaze at Agents — Right table: Subjective Gaze
<table>
<thead>
<tr>
<th>Area</th>
<th>Fixations</th>
<th>Fix/tot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>531</td>
<td>0.34</td>
</tr>
<tr>
<td>Torso</td>
<td>263</td>
<td>0.17</td>
</tr>
<tr>
<td>L Arm</td>
<td>40</td>
<td>0.02</td>
</tr>
<tr>
<td>R Arm</td>
<td>30</td>
<td>0.01</td>
</tr>
<tr>
<td>Pelvis</td>
<td>64</td>
<td>0.04</td>
</tr>
<tr>
<td>Legs</td>
<td>52</td>
<td>0.03</td>
</tr>
<tr>
<td>Feet</td>
<td>23</td>
<td>0.01</td>
</tr>
<tr>
<td>Scene</td>
<td>517</td>
<td>0.34</td>
</tr>
<tr>
<td>Total</td>
<td>1520</td>
<td>0.96</td>
</tr>
</tbody>
</table>

*Table A.2:* Percent of each area, approximate to two decimal places

### A.3 Data Calculations in R

Kruskal-Wallis calculation:
Each set of group data is added, and then sorted into a list. The variables contain the Likert values of participants 1-30 for each column (or clip).

```r
A <- c(3,3,3,4,3,5,3,5,5,3,3,2,2,2,4,5,4,3,5,3,3,5,4,2,6,4,1,4)
B <- c(3,1,1,3,1,4,3,5,4,5,2,1,4,5,2,3,2,4,6,5,6,5,3,2,5,5,3,5,4)
C <- c(2,3,4,4,6,3,4,4,5,3,4,3,3,4,4,5,5,5,3,5,6,5,5,2,3,6,3,6,4)

compareABC <- list(g1=A, g2=B, g3=C)

kruskal.test(compareABC)

Kruskal–Wallis rank sum test

data: compareABC
Kruskal–Wallis chi-squared = 3.7766, df = 2, p-value = 0.1513
```
References


