Aalborg University Copenhagen

Semester: 10th

Title: The Elements of a Visual Effects Set Extension –

Evaluating a spectators level of Believability for VFX shots through a synthesized Model and Methodology framework.

Project Period: 01/02/2012-24/05/2012

Aalborg University Copenhagen
Lautrupvang 15, 2750 Ballerup, Denmark

Semester Coordinator: Stefania Serafin

Secretary: Lisbeth Nykjær

Semester Theme: Master Thesis

Supervisor(s): Rolf Nordahl, Niels Christian Nilsson

Project group no.: N/A

Members:

Jonas Rasmussen

Asbjoern Svenning Andersen

Abstract:

The following report outlines a 10th Semester Medialogy Thesis at Aalborg University Ballerup.

The thesis undertook the process of synthesizing a theoretical model and test methodology to evaluate upon a spectators level of *Believability* in regards to Fore- and Background elements in a Visual Effects Set Extension. The relationship was hypothesized to be relevant as valuable work and time could be saved if a high amount of background elements were found to be unnecessary when the foreground presented an element of interest.

Relevant perception topics were addressed extensively as a means to create the model that would allow us to evaluate the otherwise very subjective topic of *Believability*. Through the analysis we created a qualitative method for evaluation involving a questionnaire with respective guidelines for interpreting a subject's answer. The end result was a per-case evaluation with one of two answers; believable or unbelievable.

In addition to the model of Believability we created a test methodology that strived to examine whether the Fore- and Background elements in the shot were related and to what extent. Various significance tests including the *Chi-square, Factorial ANOVA for Mixed Designs* and *McNemar* were used to test for these relations. Even if numerous considerations in terms of the results must be reflected upon the thesis showed that the Fore- and Background Elements were indeed related. When the Foreground has an element of interest, the level of detail of detail in the Background is less important.

Thus the thesis addressed that valuable work/time may be saved in terms of VFX Background details when the foreground presents an element.

Copies: 4
Pages: 113
Finished: 24/05/2012

Preface

This thesis was developed during the 10th Semester for Medialogy at Aalborg University Copenhagen during the period 01/02/2012-24/05/2012.

Reader's Guide

The report is divided into various sections. Initially we motivate the theme for the thesis which concludes in an initial problem statement. This is investigated through a preliminary analysis which results in a final problem statement. What follows is a thorough analysis that examines the theoretical foundation of the topics relevant to the thesis ultimately ending with a set of supporting hypotheses. Based upon these we cover relevant Test Methodology which in turn motivates the design and implementation of 6 iterations of a science fiction clip to be used for testing purposes. Conclusively we analyse, discuss and conclude upon the results.

All references throughout the report use the Harvard Angila* standard. A reference will therefore look like (Brinkmann, 2008). Whenever we reference other sections or the appendix the text and section number will bold as follows, Section 2.1.4 or Appendix 8.1.

Full bibliography and list of figures are included at the end of the report.

A DVD is included in addition to the report. The DVD includes supplementary elements to the thesis (questionnaires, full test results and references) and the 6 iterations of the video clip. Elements on the DVD are referenced as DVD/Test Results/results.xlsx. Full digital pdf and text version are included in the root of the DVD.

The printed version of the report has a quality based on the printer used – for some images throughout the report we encourage the reader to open the digital version as the level of details is much greater.

Special thanks

We would like to thank in particular Niels Christian Nilsson and Rolf Nordahl for great supervision in terms of the project work. Furthermore we would like to thank Ian Bach and Tine Nikali from Slowmotion fx for their technical know-how and great insight into the realm of VFX!

1 Table of Content

2 List of Figures

3 Introduction

3.1 Methodology

3.1.1 Overview

This section will address the methodology behind the thesis with emphasis on the various steps that were taken to conduct an appropriate experiment in terms of a problem statement and various hypotheses.
The thesis involved six distinct phases

- Pre-analysis
- Problem Statement and Analysis
- Hypotheses
- Test Methodology
- Design and Implementation
- Results and Analysis

3.1.2 Pre-Analysis

An initial problem statement addressing Visual Effects set extensions in terms of *Believability* and the amount of production work involved was proposed and such topics were thus explored. The outcome of the preliminary analysis was a series of factors that all contributed to the amount of work for set extensions and as such could provide some guidelines as to what one should pay attention to. Additionally, the pre-analysis helped us start outline a definition for the main theoretical subject of *Believability* in terms of *Perception*. Set extensions are used in many different films and shots and it was thus an interesting field to delve into.

3.1.3 Problem Statement and Analysis

Based on the preliminary analysis a final Problem Statement was proposed:

"To what extent will believability be affected when altering the elements in the foreground and background in a set extension?"

The *Analysis* thus explored relevant topics to great depth. Based upon a multitude of various theorists we synthesized a definition of *Believability* and created a model to use for evaluating a subject's impression of believability with respect to a visual effect set extension. Moreover we addressed film theory in terms of deciding upon the particular type of shot and genre of the set extension.

3.1.4 Hypotheses

With a foundation in the *Analysis* and the *Problem Statement* various hypotheses were proposed. Such hypotheses make for a structured way to address a problem statement as they typically break down the problem into smaller more tangible sub-topics. Three such were proposed:

- If a foreground element is present a test subjects focus lies on that, especially if the foreground presents an action. Therefore the need for a complex background is lessened to achieve believability.
- Subsequently, if the foreground is not present the focus will lie on the background therefore a more complex background is needed.
- If the subjects fail to realise that the setting is sci-fi the execution of the set extension may be regarded as unsuccessful.

3.1.5 Test Methodology

The following describes the methodology behind evaluating upon the hypotheses and problem statement.

A *Mixed Factorial Design* was chosen as a means of evaluating upon six iterations of a similar clip as motivated by addressing the problem statement. A mixed design is a combination of *Independent Groups Design* and *Repeated Measures* and it is primarily ideal for situations where the number of samples might fall short. It furthermore inherits the advantage of repeated measure where the results will be more powerful. Each subject in a particular test session would thus watch two clips (two levels of background details) for one single level of foreground details. The order in which they were presented with the background detail was randomized to reduce bias. The result of this approach was thus repeated measures within each level of foreground and independent groups between the levels of foreground.

A questionnaire comprising five questions was used to evaluate whether a subject found a clip believable or not. The questions were made open in nature as not to lead subjects onto particular answers. What followed was a careful per-case examination of each subject's answers in regards to a particular clip.

3.1.6 Design and Implementation

The resulting data was eventually subject to inferential statistics as a means to analyse whether any apparent relationship would be significant or not.

Entirely motivated by requirements proposed by the hypotheses and problem statement we designed and implemented six different iterations of a short science fiction scene. The various iterations involved two levels of background elements and three levels of foreground elements which were then cross-combined to produce six unique clips. The purpose of these was to afford an evaluation of each clip in terms of the Believability model. Therefore it was of importance that the differences between each clip were distinguishable.

3.1.7 Results and Analysis

The experiment was conducted using two samples; A *Peer Group* comprising friends and their friends and a *Campus Group* using people found scattered around the University campus. Regardless of which group a participant belonged to the procedure was similar as they would each watch two clips hosted online and answer questionnaires for each clip at similarly an online host. The main difference between the two approaches is that Campus Group might be regarded as more regulated as compared to people conducting the tests at their homes at their own accords. The difference in regulation was however regarded as less important than obtaining a large sample number.

In the end 144 participants participated in the experiment with a total of 288 clips watched. What had not been anticipated however was that in order to conduct significance tests across multiple independent variables for binary frequencies (number of 'believable' clips) it proves hard to do so when the experiment involves repeated measures. In order to counter this we halved the original sample thus obtaining 6six independent samples for each of the clips. The *Chi-square* test was following used to test for a significant relationship. To test whether the differences within each group of foreground elements was significant the *McNemar* test was used on all of the original data as it assumes paired-samples. Finally one of the questions involved how much a participant liked the clip (1-5) and for this a *Factorial ANOVA for Mixed Design* was used as the data was not binary frequencies.

4 Motivation

Since the very beginning of film creation there has been a desire to create the imaginary with whatever means at ones disposal. One of the pioneers in creating such visual trickery was the French magician *George Mèliés* who during the early 1900's was responsible for hundreds of films (Ezra, 2000). Common for all these was the fact that they typically included creative use of perspective shifts, frame inserts and other in-camera effects to produce some spectacular effect. One of *Mèliés* most famous films was also one of the earliest narrative films and typically referred to as one of the first *science-fiction* film *'A Trip to the Moon*' (Le Voyage dans la lune, 1902). This particular film included a famous moon landing scene created by visual trickery and the scene is so iconic in itself that an illustrated adaption serves as the logo for the Visual Effects Society (VES, 2012).

At the time of *Mèliés'* effects there were obvious technological limits compared to the presence namely in terms of everything being carried out analogous. Similarly to other industries the world has witnessed a wave of digitization throughout most if not all of the visual effects (VFX) domain the past 20-30 years. Procedures that used to be time-expensive or even impossible are much more readily accessible and as a result the actual no. of VFX have increased (Okun & Zwerman, 2010; Finance & Zwerman, 2010) for both film and TV productions as well as the desired quality.

The average moviegoer typically associates VFX with visible effects as depicted in blockbusters such as *Avatar* (2009) wherein the audience is aware that it is in fact an effect but accepts and believes in it nonetheless. However the notion of VFX is a broad one including elaborate computer generated imagery (CGI) but also includes an entire sub-genre of invisible effects where the audience is not meant to notice anything out of the norm. Whether this actually comes true depends entirely on the execution of the particular effect.

In relation to visible effects, VFX were typically used cinematically to *amaze* an audience through few yet impressive looking effects (Finance & Zwerman, 2010, p.3) but recently there has been a surge in films that are driven entirely by VFX. Films like *Avatar* and the more recent *Green Lantern* (2011) uses VFX as a means to create entire virtual *mise-én-scene* sets where VFX could be regarded as the most dominating visual element. While the VFX budgets of such Hollywood film are exceedingly large (Khouri, 2011; Thomas, 2004) not all film productions have the same budget to work within. However, audiences tend to become picky and comfortable with high production value (Mitchell, 2004)and it is therefore difficult for the smaller studios with lesser available budgets to keep up. Even audiences for TV productions have come to expect VFX of high quality but the budgets do not follow suit (Bickerton, 2010). TV productions like *Game of Thrones* (2011)

and *Boardwalk Empire* (2010) feature effects that come close to film production value and they have thus upped the level of quality for TV.

Common for both the aforementioned film- and TV productions is that they deploy numerous VFX procedures. Amongst such techniques are those of *set extensions* or *virtual sets*. These are increasingly used throughout multiple types of productions and furthermore carry great personal interest. A set extension typically includes a foreground shot in-front of an easy separable background. However, the topic is more inclusive than just that. Depending on the situation, entire virtual sets may be created for maximum flexibility as exemplified in *Transformers 3* (2011) and *Avatar* (2009) whereas in other circumstances artists may choose to extend only part of distant scenery as used in *Game of Thrones*(2011). In addition to the uniqueness of each situation, audiences have continually come to expect more (Mitchell, 2004) and the VFX artists must therefore think creatively to get around the limited personnel, time and money. As initiated by this problematic situation this thesis strives to examine whether the believability of a particular set extension can be upheld in a non-Hollywood scale production. More specifically, we are interested in investigating whether it's possible to identify various factors that determine an audience's level of appreciation and believability. If it is indeed possible to identify such factors we hope to present various proposals that relate the work involved in a set extension and the level of believability.

With the previous discussion in mind an initial problem statement can be formulated.

4.1 Initial problem statement

"Set extensions are highly sought but the artists have trouble keeping up. Is it possible to identify factors that lower the work required for a particular shot whilst maintaining high believability?"

5 Preliminary Analysis

The goal of preliminary analysis is to get a deeper understanding of the keywords presented in the initial problem statement as means of narrowing down the statement further. The preliminary analysis is divided into the following sections:

- Definitions of VFX, Work and Believability
- VFX overview
- Set Extensions including relevant State of the Art cases

5.1 Definitions

5.1.1 Definition of VFX

Effects have always had a large role in film and television (Finance & Zwerman, 2010, pp.3-18). A common misconception however is that visual effects (VFX) and special effects (SFX) are mere synonyms of one another, which would be wrong to assume. SFX may be regarded as what takes place during shooting on set as exemplified by the use of pyro-techniques, weather effects, elaborate camera or wire rigs amongst others (Okun & Zwerman, 2010). In comparison, Charles Finance & Susan Zwerman define VFX in 'The Visual Effects Producer: Understanding the Art and Business of VFX' as:

> ***"A visual effect is the manipulation of moving images by photographic or digital means that creates a photorealistic cinematic illusion that does not exist in the real world." (Finance & Zwerman, 2010, p.4)***

Through the use of digital means it is easy to understand why SFX and VFX could be understood as the same thing. Effects like explosions which were traditionally achieved through SFX are often done digitally today. The important distinction is that VFX is something that is typically not recorded on set, but rather done in *Post-Production.* However, it should also be remembered that VFX is not only digital effects but also includes effects made by photographic means, such as the aforementioned moon landing done by Georges Méliès in 'A Trip to the Moon' (1902).

The definition for VFX in this project will not follow the one put forward by Finance & Zwerman but rather one by Fink & Morie.

> ***"Visual effects is the term used to describe any imagery created, altered, or enhanced for a film or other moving media that cannot be accomplished during live-action shooting." (Okun & Zwerman, 2010, p.2)***

The reason for the distinction between the two definitions lies in the use of 'photorealistic' by Finance & Zwerman. It may be argued that VFX do not necessarily need to be photorealistic in order to function. This shall be covered when defining believability.

5.1.2 Definition of Work

The amount of work put into a VFX shot is tough to generalise. The 'safe' assumption would be to measure it in the amount of man-hours put into a particular shot, but a problem occurs when discussing how to reduce the work required for a particular shot. One thing is that VFX shots vary greatly in type, but also that there is no set techniques for how to create specific VFX shot. As VFX creation fall under a creative work form, everything is very much up to the artist and his/her preferences, experience and talent. In order to define work required as a measureable unit, we will apply the assumption that it is related to the amount of content that need to created and composited into a single shot. In the context of a set extension for this project, work will relate to the creation and compositing of the elements that the set extension consists of. It will also be considered that some techniques involved in set extensions and VFX in general require more work in terms on work hours and manpower. An example of a work-heavy nature is that of CGI which can be regarded as work intensive as it typically requires modelling, texturing/shading, lighting, rendering and animation.

5.1.3 Definition of Believability

The goal of a VFX shot is usually to create a believable illusion, integrated into film. But what is believability and how is it defined? The foremost answer is to equate believability with realism, as VFX shots often try to replicate or simulate reality and thus aim to be invisible VFX (Bickerton, 2010). Realism is a well explored area in relation to CGI which carries over into VFX to some extent.

5.1.3.1 Realism in CGI

Photorealism in CGI is a substantial area with a lot of effort being put into improving algorithms, but also into research of understanding the perception of photorealism. A study from the University of North Carolina showed that photographs of real objects with changing properties (changes in texture and softness of shadows amongst others) also meant that the perceived realism changed, despite every photo being real (Rademacher et al., 2001). This is an interesting observation when considering Charles S. Peirce's notion that:

> ***"Photographs, especially instantaneous photographs, are very instructive, because we know that in certain respects they are exactly like the objects they represent . . . they . . . correspond point by point to nature. In that respect then, they belong to the second class of signs, those by physical connection." (Prince, 1996, p.28)***

Stating that photographs are a good example of realism, it's interesting to consider that photographs can convey an unrealistic affordance even without the use of photo manipulation. An important aspect to consider is that the experiment from the University of North Carolina required participants to rate each picture as being real or not real, which could be inclined to give the participants the preconceived notion that

they would encounter both real and non-real pictures. Peirce's notion also implies that one is aware that the picture at hand is indeed a photograph, in that they represent objects.

The preconceived notion of the medium in question raises an interesting point for realism but also the definition of believability as a whole. What about the cases where VFX and/or CGI do not try to portray something we can classify as realistic? Such are unreal in nature or unreal due to the viewer having no pre-existing knowledge of the element e.g. a monster in a film. How would this fit into believability? The preconceived notion of the medium (in this case being film) is something that correlates well with the concept of willing suspension of disbelief (Hooks, 2011; Schaper, 1978) where the viewer willingly accepts the film as 'reality' despite being aware of its fictitious nature which in turn is why a film can provoke emotional responses (Schaper, 1978).

5.1.3.2 Perception of the real and unreal

An interesting question to address is how we perceive elements we know are real but also those we know are unreal. Willing suspension of disbelief is one part of the explanation, but it cannot be regarded as the sole element that explains why and how we accept what we perceive as being unreal. To examine this further a look at how we perceive objects in general on a perceptual level, will be made.

On a higher perception level, theorists argue how we recognise objects. Some have claimed that we build up an extensive mental library of three-dimensional objects (Coren et al., 2004, pp.320-21) while others propose that the library consists of much smaller three-dimensional parts or components that we use to quickly assemble into something we recognise. This also leads onto how we as spectators perceive what happens in a film; Stephen Prince argues that:

> ***"An extensive body of evidence indicates the many ways in which film spectatorship builds on correspondences between selected features of the cinematic display and a viewer's real-world visual and social experience." (Prince, 1996, p.31)***

Prince further argues that this is the reason why we can perceive something obviously unreal as being real in a film context. As an example Prince highlights the dinosaurs from *Jurassic Park* (1993). While we know dinosaurs from books, museums etc. no one has ever seen a dinosaur and no one can truly know how they moved or accurately looked. But we still accept them as being 'real' as we form our own expectations for their movement and look, from a composite of the aforementioned visual and social experience.

> ***"Such images display a nested hierarchy of cues which organize the display of light, colour, texture, movement, and sound in ways that correspond with the viewer's own understanding of these phenomena in daily life." (Prince, 1996, p.32)***

Given this theory it would explain why we would accept the transformer robots in *Transformers: Dark of the Moo*n, *the Na'vis* in *Avatar* and so forth.

This also correlates with the perception theory of *Conceptually Driven Processing* (Coren et al., 2004, p.320) which states that our patterns in our stimulus input is guided by experience from past events, memories and general organization strategies; e.g. the perception of a dark flash through the air could be initially perceived as a bird, if we were at a quiet park, or as a ball if we were at a busy playground. This goes to show that the human perception system is dependent on several factors, making it a very complex system to dissect and understand.

Going a step back to realism as a term, it should be remembered that while films like Transformers and Avatar deal with unreal entities in shape of giant robots and aliens they are still created with a goal of being realistic and believable. In terms of Prince's notion there must thus be some correspondences to real-world visual and social experience. With respect to Avatar this is visible in the fact of the great amount of work that went in to motion capture performance as well as creating very complex skin-shaders (Sandal, 2010) in order to archive a very high degree of realism. In regards to VFX and films, *Final Fantasy: The Spirits Within* (2001) tried to achieve photorealistic CGI humans. However the film was a box office flop which some critics attribute to the fact that it dropped into the uncanny valley(Mori, 1970; Monnet, 2004). In relation to believability this adds an important point, as the threshold for when a given object is perceived as being believable, is very intangible. As described with the uncanny valley the difference between achieving a believable and realistic result compared with failing and thus obtaining an uncanny result, when creating CGI humans, is very slim. This comes down to the fact that we ourselves are humans, and constantly perceive others around us, which gives us a very detailed expectation to how humans move and look, making it very easy to spot any mistakes and shortcomings. *Møller* and *Pellengahr* (2011) investigated how people had a more consistent and clear expectation on how fire looked compared to smoke. This meant that people were more critical towards various iterations (in terms of simulation resolution) of simulated fire. Whereas with smoke there was a broader acceptance as the expectations, and thereby comparison basis, were not as defined. This adds to the idea that believability is a very context dependent element in terms of the preconceived notion of the medium as implied by Peirce (Prince, 1996, p.28), and now as well in terms of existing expectations of a given object.

With all of this a definition of believability can be constructed:

"Believability is to accept an element as being real within its own context and medium. Thus covering both unreal elements and truly realistic elements."

With the keywords defined, a general look into VFX will be made, in order to gain knowledge of the pipeline as well as precautions that needs to be taken. This will help determine factors that can influence the amount of work needed in a VFX shot, and eventually help isolate such factors.

5.2 VFX Overview

5.2.1 Introduction

As previously mentioned, VFX is put to use in practically all TV/film productions and a general discussion with respect to *pipelines* proves relevant. Even if the thesis emphasizes on the notion of *set extension*s, this section will provide a quick overview over VFX as a whole providing insight into the general workflow and pipeline of typical shots. The aim is to provide the reader with an idea of the magnitude and presence of VFX as a cinematic tool.

5.2.2 Overall

Even if the individual workflows for VFX shots differ there are norms and trends that typically get deployed. A typical VFX pipeline of such norms is outlined in Figure 1. Such a pipeline most generally consists of three overall stages, namely *pre-production, production* and *post-production*. As shown on the illustration, the various procedures may overlap the different stages as said stages are only indications of whereabouts in the flow a shot is currently at. In relation to *set extensions* in particular a similar pipeline may effectively be used to describe the flow as extensions can include any number of elements. There are extensions that purely require 2D work whereas others may require individually rendered 3D CGI elements to be composited. Even if a set extension may involve a combination of all the procedures most work typically resides in the final node, *compositing*.

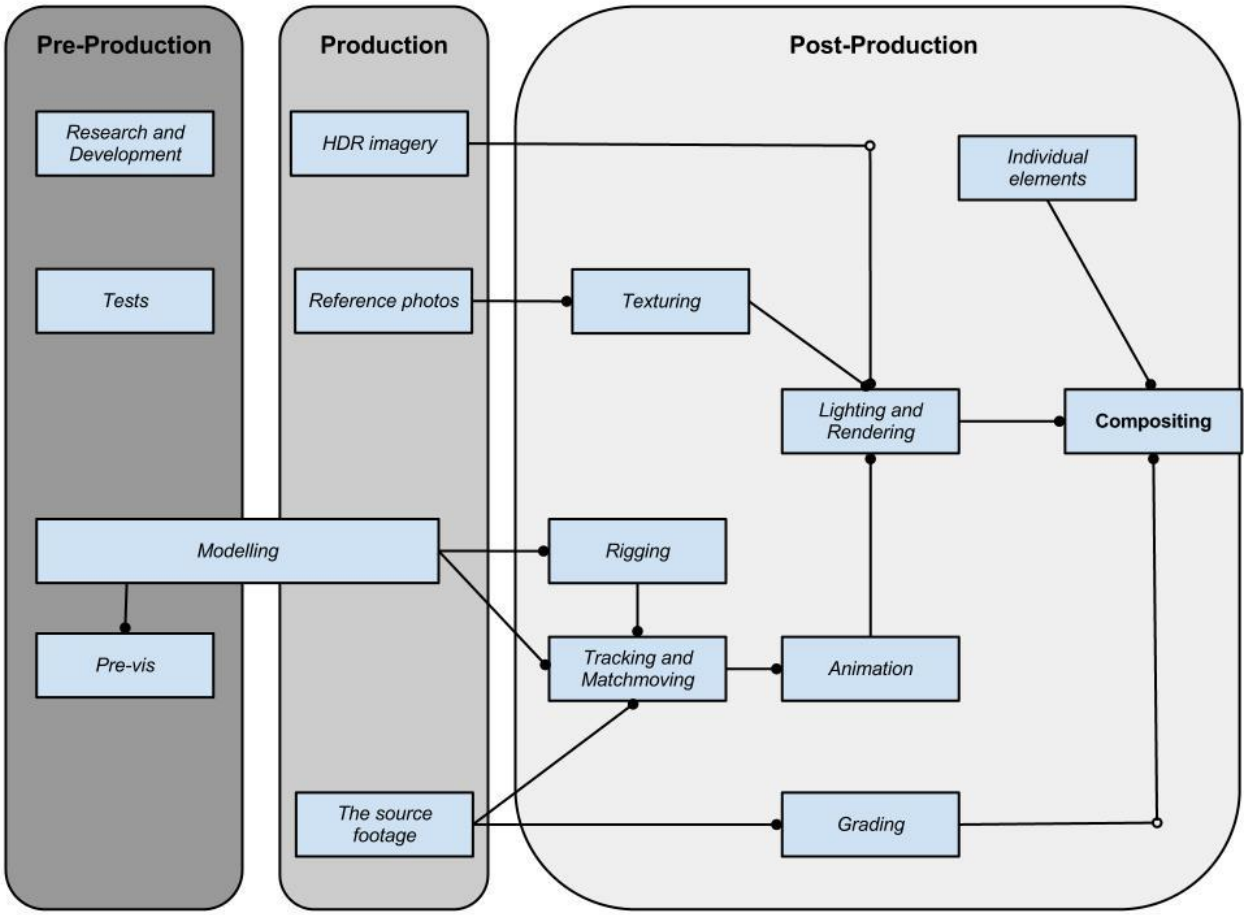


Figure 1 VFX Pipeline

Evidently, VFX is an expansive topic that includes numerous sub-topics. This particular pipeline provides a general overview over VFX addressing the various areas of interest. For the sake of simplicity we will further divide VFX into two main categories; namely that of visible and invisible effects.

5.2.3 Visible Effects

An audience typically associates CGI and VFX with *visible effects* such as the lush planet Pandora in *Avatar* (2009), an invaded Chicago in *Transformers:* Dark *Side of the Moon* (2011) and the scores of roaring apes in *Rise of the Planet of the Apes* (2011) (see Figure 2).

Figure 2 Rise of the Planet of the Apes © 20th Century Fox, 2011

Common for all three examples is the need for *CGI* to varying extents. The process of creating such elaborate CGI involves tremendous amounts of *work* (see section 5.1.2 for definition) and should only be considered when necessary. As part of preliminary research we orchestrated an interview (Appendix A: Interview with Slowmotion FX) with the small scale Copenhagen-based VFX studio, *Slowmotion fx,* as a means to include observations and considerations from the industry. What was apparent from the interview was that CGI effects in general were avoided wherever possible. In small productions alternative ways of solving shots are often sought instead. As an example of such a workaround the lead-compositor at *Slowmotion fx* informed us how he'd solved a shot for a short that included steam immersing from a container. One way to solve this would have been to create CGI particle simulations followed by combining the end result of that with the original plate. The way it was solved however was to film the immersing steam from a coffee machine in front of a black velvet screen, and then separate the two elements via a simple luminance key. This type of ingenuity is what makes VFX an interesting topic to operate within but also why it is exceedingly difficult to normalize techniques - there is no *golden solution*.

5.2.4 Invisible Effects

VFX also spans over a whole division of invisible effects. Such effects include set extensions as exemplified by *Game of Thrones, Avatar* and *Twilight - Breaking Dawn - Part 1* (2010)(see Figure 1). The latter involves greenscreen footage combined with an exterior-set matte painting in what results in a seamless transition. The quality of such an effect depends entirely on the execution; if the match-moving is unsteady, keying is imperfect or the colour integration falls behind chances are the shot fails completely.

Figure 3 Green Screen in The Twilight Saga: Breaking Dawn Part 1 © Summit Entertainment, LLC

Apart from such seamless set extensions, invisible effects also include repairs and fixes. Following primary photography there is often a need to fix problematic footage or the director may have had a change of heart. The support cable for a stuntman may have to be removed, the main actor is caught looking into the camera or a particular piece of clothing is thought to be the wrong colour. Either way, VFX is used extensively to take care of such things. Some fixes involve isolating and cloning areas, frame-freezing, time-blending and stitching footage back together whereas other fixes may require entirely different techniques.

5.2.5 Other considerations

An efficient workflow for VFX creation also requires regular communication between the parties involved. Based on the interview with *Slowmotion fx*, it was apparent that communication is of utmost importance to any production. Valuable time may be wasted if the director doesn't make his requirements clear or if the VFX supervisor makes decisions without confirmation. However, it was also emphasized that excessive communication may have an advert effect - the director may change his mind repeatedly and it should thus be made clear from the start what the deal actually entails.

As the broader approach to VFX has now been examined, the next section will focus on set extension in particular, how it came to be, and which techniques are commonly used in modern VFX production. This will be done in order to present what a set extension typically consists of and to get an understanding of how to approach it.

5.3 Set Extension

At its core set extensions build upon the traditional matte painting technique. Pioneered by Norman Dawn in the early 20th century (Finance & Zwerman, 2010) the process involved filming on location through a glass

onto which the desired elements would be painted creating the desired extension of the set. The process evolved with the negative matte (Mattingly, 2011) where a black matte (mask) would be placed in front of the camera during the shooting, blocking out the part which the matte painter would later paint in his studio. Walter Percy Day was one of the pioneers of this technique which became the standard of matte paintings for the years to come (Figure 4). Matte painters were gifted artists that were able to create highly detailed paintings which blended seamlessly with the life footage. Such techniques resulted in saved costs as the amount of set construction could be reduced.

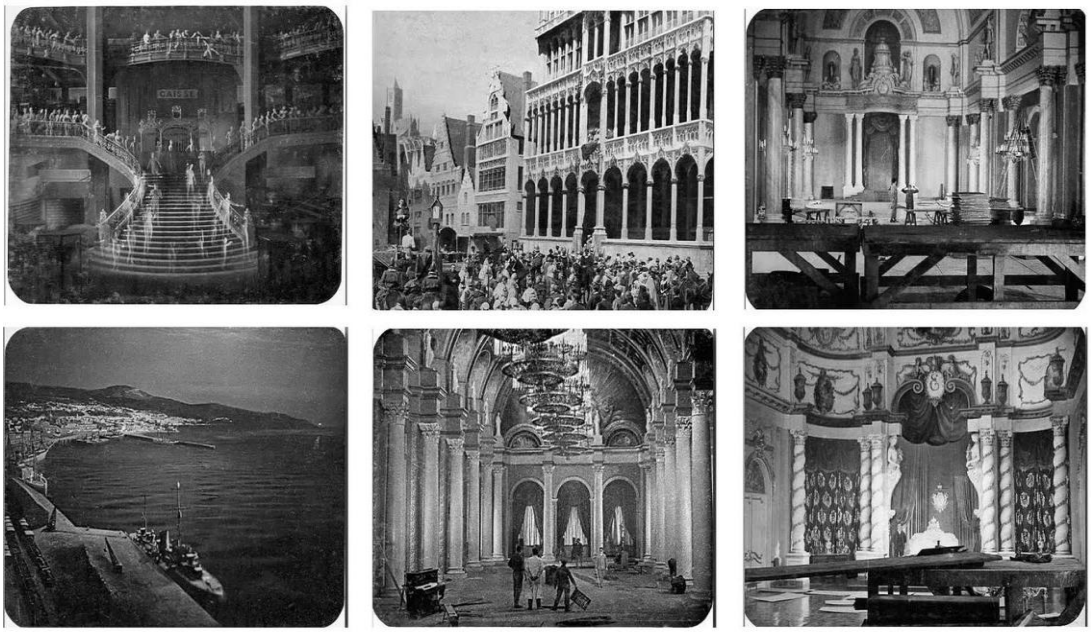


Figure 4 Collection of matte paintings by Walter Percy Day

Matte painting has evolved with the introduction of digital possibilities as matte paintings for the most part no longer is done with brushes and paint, but rather with a mouse and tablet . The technique of using a 2D image as a set extension is still a core element in the set extension toolbox, but it has been vastly expanded with the advent of CGI and techniques such as camera projection, keying and other techniques. The important thing to remember is that traditional matte painting technique is still used, albeit in digital form and that the possibilities of a set extension have increased vastly. Such possibilities will be explored to gain a better overview of set extensions.

As discussed in previous sections, VFX and set extensions are used in a broad variety of film and television productions. As with VFX in general, the degree to which set extension is used can be very different from shot to shot, from more or less full set creation to small additions in the background. The following will therefore be a rough categorisation of set extensions based on the amount of content created digitally.

5.3.1 Basic Set Extension

Matte painting similar to the traditional way is still widely used, although today the 2D image is created digitally followed by being roto-scoped into the appropriate location or replacing a green screen. This is used as it always has been e.g. as an extension to a background or a building. As a slight advance to the classic matte painting usage, instead of doing a 2D composite of the matte paintings, they can be placed and layered in 3D space. This gives the possibility to create parallax in the set extension, creating more depth and possibly adding to the believability of the set extension. Creating parallax in film by layering images in 3D cannot be regarded as something entirely new with the introduction of digital VFX. The technique is very reminiscent of Disney's use of the *Multiplane* camera (Johnston & Thomas, 1981) creating parallax in their animation films by painting and filming elements on several glass plates, which gave the parallax effect, a rather revolutionary technique at the time.

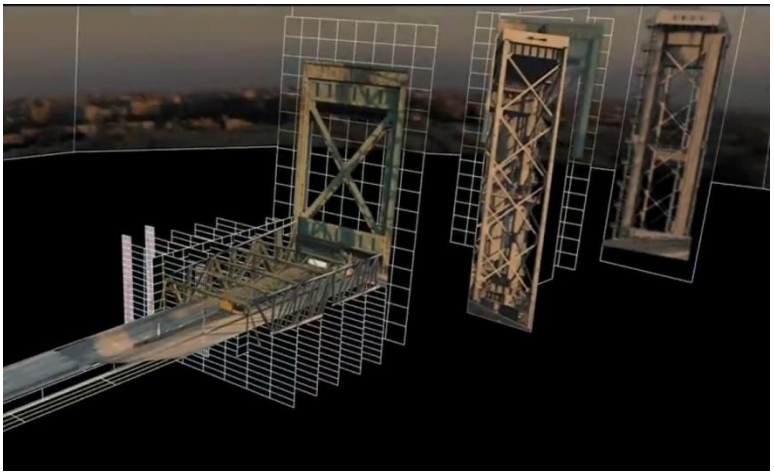


Figure 5 3D composite setup from Transformers: Dark of the Moon © ILM

5.3.2 CGI Set Extension

The use of 3D space for set extensions has become a very essential tool, and the use of CGI in combination or as replacement for the 2D matte painting. CGI gives even more depth than the 2D image and gives room for more camera movement in all directions, where the 2D matte paintings have their obvious limitations. The downside to the use of CGI in set extensions is that it can be a rather time consuming affair that includes modelling, texturing, lighting and rendering. All of these depend entirely on the scope of the set extension. CGI set extensions are also often the foundation for the more elaborate set creation where the majority of, or the whole set, is created by VFX.

5.3.3 Camera Projection

In between *The Basic Set Extension* and *CGI Extension* is the hybrid type extension known as camera projection. The basic concept with camera projection is to have a still image of the object needed for the scene (in this case the camera projection) e.g. a building. For this very simple example, CGI geometry of the building is created and aligned with the image such that the virtual camera matches the position of the real camera that took the picture. From here the virtual camera is used to project the still image onto the geometry, creating a CGI model which is textured and lit. Furthermore camera projection corrects the perspective such that the model can be viewed from different angles with only little distortion. While camera projection has some of the same flexibility as CGI, in terms of moving a camera around the objects, it also has some limitations as geometry not covered by the image will have no proper texture from the projection (Figure 6).

Figure 6 On the Left: Original picture with aligned geometry. Centre: Projected texture. Right: Limits of camera projection.

With the basis for set extension explored, as well as some of the techniques covered, a look at how set extensions are used in state of the art productions will be covered next.

5.4 State of the Art

The following section will provide insight as to where the industry is currently at with respect to set extensions in particular. In addition to the productions briefly introduced in the *Motivation (section 4)* the following will go through a few other high quality productions.

5.4.1 TV productions

There has lately been a surge in high quality productions made for the TV as previously exemplified by *Game of Thrones* (2011) amongst others. Audiences have come to expect much more as of late and the studios simply have to keep up (Mitchell, 2004). Fortunately the technologies evolve pretty swiftly also. Examples of such techniques include the previously mentioned camera projection and full CGI set extensions.

Another great example of such techniques put to use is for the critically acclaimed HBO series *Boardwalk Empire* (Boardwalk Empire, 2010) wherein *Brainstorm Digital* was responsible for recreating an Atlantic City some hundred years ago. In an interview with Fxguide (Failes, 2011), VFX producer Richard Friedlander discusses how the set extensions came about. Most notably, but also to be expected, is the number of various techniques that were used for the production, as Friedlander says:

"Most of the shots were a combination of everything from a visual effects point of view". (Failes, 2011)

Depending on the particular shot they would use traditional 2D matte paintings for the far background and full CGI for the mid to near-ground. Camera projections would be used interchangeably to project reference footage onto geometry. An example of this is shown in Figure 7. A large practical set was built outdoors with a blue spray-painted wall defining the limits of the set. According to Friedlander, the actual keying of the blue wall was harder than had they shot indoors but they chose against it to obtain natural light inflections and the imperfections that are crucial to sell a shot. An interesting point made in the interview was that of matching the colour palette from the live action plate with the extension. They were aiming for a look some hundred years old but all reference footage was black-white and they thus had to improvise as it was important to portray the right look to the audience.

Figure 7 Broadwalk Empire © HBO

Another use of set extensions for a TV production was for the 2011 *Pan Am* series set in a 1960s environment. Differently to *Boardwalk Empire* all the set extensions were shot inside a large studio. Based on historical photographs *Stargate Studios* recreated the iconic 1960s Pan Am terminal as seen in Figure 8. To actually carry out the extensions a combination of the aforementioned techniques were put to use. Furthermore, the studio used a real-time compositing system that would allow the director to see what the final composite would look like on-set (Failes, 2011). This required the studio to create the virtual sets and assets prior to shooting but this also meant that final compositing would be the more relatively quicker.

Figure 8 Pan Am © ABC

5.4.2 Film productions

The set extensions created for Avatar are still amongst the most successful even if they date back to 2009. A more recent example is the Martin Scorsese *Hugo* (2011) that put set extensions to use throughout most of the film. The film takes place in a 1930's Parisian train station and similarly to the previous examples it therefore required the team to aim for a particular 'old touch/feel'.

It is interesting to note that traditional effects were combined with new cutting edge techniques. Miniature effects were put to use as exemplified by an iconic train crashing scene as seen in Figure 9. This particular scene is in fact homage to the Granville-Paris Express derailment from 1895. On a side-note the film includes many other homages and the film is in all actuality a homage to *George Méliès,* an effects pioneer. Even if director Martin Scorsese does not belong with the newest wave of directors he pushes the limits of the possible. This must in part be credited to the fact that he embraces modern techniques and combines them with more traditional techniques.

Figure 9 Comparison between the real crash and the recreation on the set of Hugo

Technically with respect to choosing miniature effects, there are considerations in terms of frame rates, how gases react at a lower scale, and the notion of reproducibility/repeatability that should be investigated prior to shooting (Seymour, 2011). However the VFX department is also given a lot of 'free' elements that help sell a shot. If the shot involves a collision with an actual practical prop there is that element of randomness and imperfection that computer simulations struggle to obtain. The use of miniatures for this film worked very well but it must be carefully considered if such are to be used for the thesis.

In addition to miniature effects, the film also included a multitude of set extensions similar to *Boardwalk Empire* and *Pan-am*. The production companies would build rather extensive practical foreground sets and these were thus augmented with backgrounds created through full-CGI and camera projection techniques, as seen in Figure 10. As with all other effects in this film, the quality was exceedingly high with seamless transitions albeit the look was characteristic in nature.

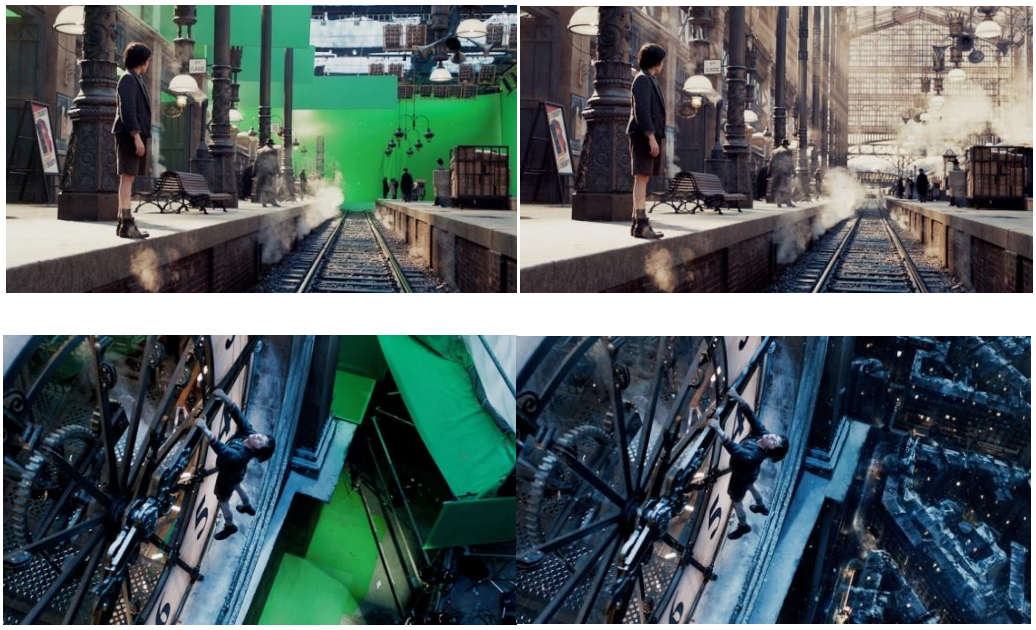


Figure 10 Hugo © Paramount Pictures

Another movie that features high quality set extensions and acts as one of the primary motivations for this thesis is the 2011 *Transformers: Dark of the Moon*. Even if the story presents the audience with hordes of futuristic machines and city-scapes in ruins, the film includes some of the most believable and photorealistic imagery currently available. In contrast to the three prior examples this particular film presents itself as realistic whereas the others appear characteristic in nature either in terms of colour nuances or slightly stylized CGI. The latest *Transformers* excels in this regard as what is presented appears real even if it does involve demolished buildings and machines.

Industrial Light & Magic recently released a VFX breakdown of the work they did for the film (Industrial Light & Magic, 2012). Figure 11 presents a series of steps that all contributed to the creation of a Chicago city-scape in ruins combined with machines.

The raw plate was initially shot and this was then *matchmoved* with respect to CGI geometry that would match the buildings exactly. Based on the matched geometry, particle elements, debris and others such would be appropriately inserted. Alternative textures would be projected and added to the buildings mimicking ruined windows and pieces falling. The machines would then be inserted and made to interact with the buildings wherever required. Finally the end plate is graded and lens flares/volumetric fog effects are added.

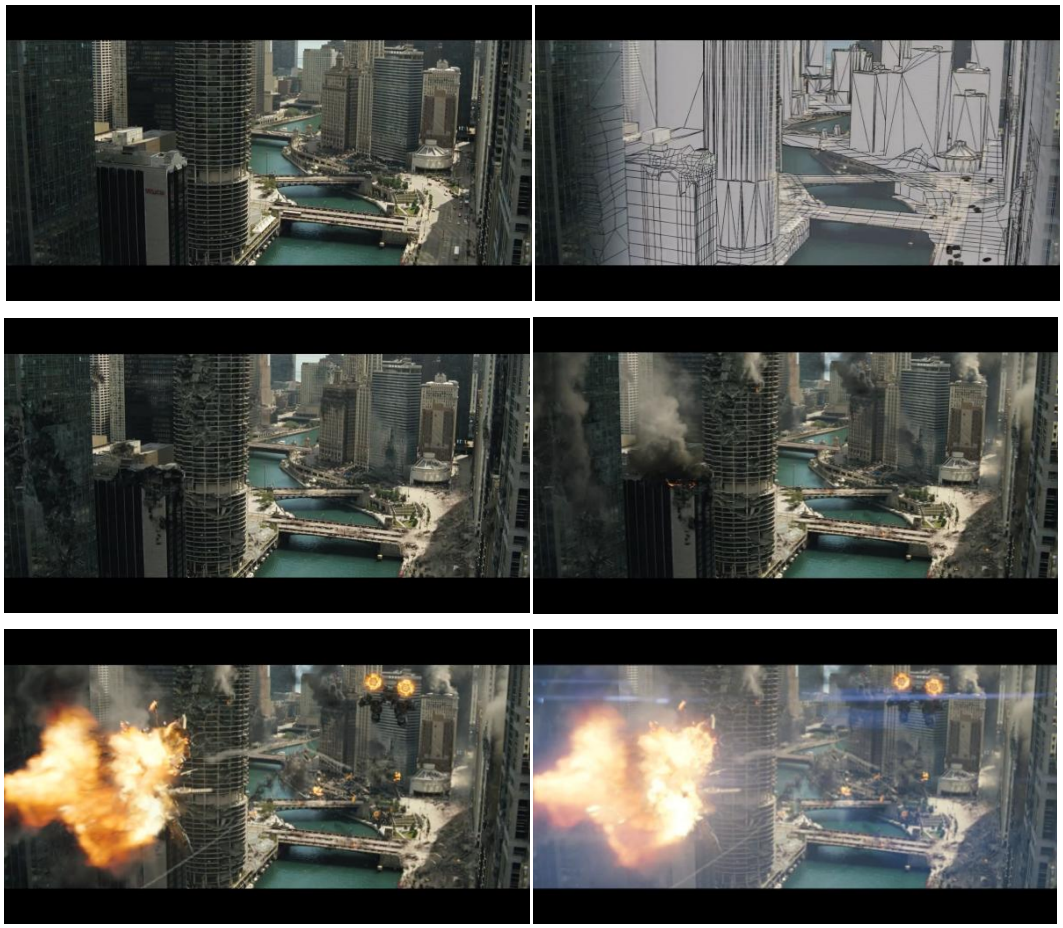


Figure 11 Transformers: Dark of the Moon © Paramount Pictures

The previous example is a great example of the power of using set extensions and provides inspiration for the thesis at hand. It is an example of all of the different techniques at work which in turn provides for the ultimate flexibility.

Common for the aforementioned examples is the fact that audiences continually expect more (Mitchell, 2004). As new productions come out the boundaries of the possible are pushed ahead and the smaller studios have trouble keeping up. *Boardwalk Empire* (2010) and *Pan-am* (2011) are both examples of TV productions that have managed to keep up but the issue at hand is that the budgets do not follow suit (Bickerton, 2010). This is obviously a path wherein the smaller studios in particular will not be able to stand long term. As we looked through the various productions that put set extensions to use we sought in part to identify what was responsible for the most *work* involved for a given shot. What quickly struck us was that the number of individual elements (fog, smoke, debris etc.) varied greatly depending on the scope of the production. For each of the TV productions and also *Hugo* (2011) in part, the number of elements was relatively small as compared to *Transformers* (2011). For the latter all scenes involve a very large number of fore-, middle-, and background elements - if one closely examines the background there is movement everywhere. The three prior examples are all unanimously praised however and it could thus be that there is no actual benefit from having excessively too many elements in a scene.

The prior discussion provides valuable observations with respect to back- and foreground elements. It is particular interesting to consider whether it is possible to identity a relationship between a given set of back- and foreground elements in relation to how believable a set extension presents itself - is it actually necessary to create a multitude of background elements if the foreground comprises dominant elements? *Work*, as defined in section 5.1.2, is related to the number of elements in a shot and if it is possible to reduce the number of such elements the saved work can be put to use elsewhere. In relation to this we will work with fore- and background in relation to each other. The background in this regard is composed of what traditionally can be defined as back- and middleground and the elements within these. An element in this context is that of any object that a back- or foreground is comprised of and furthermore relates to the context of the shot itself. It is a somewhat abstract notion and it will be properly addressed at a later point but for simplicity an element could be anything from smoke/fire/debris to props or even people (see section 10). It came about as a result of a particular scene in *Transformers: Dark of the Moon* wherein the singular heroine takes up the foreground whilst the background features lots of elements. Based on this example we were led to consider that the number of elements in the background possibly could be lowered without necessarily lowering the believability of the shot as the dominating foreground appears to be where one focuses.

5.5 Pre-Analysis Summary

The following section will address the results of the pre-analysis. In particular we will focus on actual considerations that may influence a VFX shot in regards to a production. Only considerations of direct

interest to a set-extension will be discussed here. It should not be viewed as a definitive or all-encompassing collection, as many of these can vary from shot to shot, nor is it an attempt to create an definitive collection, but we rather strive to identify the more common ones. Throughout the pre-analysis we were interested in identifying considerations that contribute to the amount of *work* in order to be aware of what one should consider in the creation of set extensions. The following will summarise such considerations including what has been addressed in previous sections and furthermore additional ones as identified for this particular section. For simplicity we have broken the section into various phases - it should be noted that these typically overlap. Furthermore, wherever possible we break each consideration into actual factors that influence the work involved.

The considerations and factors have been identified through an open dialog with *Slowmotion fx, The VES Visual Effects Handbook* and own observations.

5.5.1 Pre-planning

This phase is crucial to any production as it has a direct effect on the rest of the production pipeline. It may be decided that the entire production is to be stereoscopic 3D which will have a large influence on the VFX in particular. Furthermore, deciding whether said stereoscopic is to be filmed in stereoscopic or done in post-processing will influence the pipeline even further. In terms of set extensions in particular the pre-planning phase is crucial as most technologies and techniques are determined. The producer may overrule the directors wish of having a lush CGI environment due to limited funds or perhaps it is decided that interaction between the real set and CGI is unnecessary making the relatively inexpensive 2D matte paintings a feasible solution. Typical things that should be considered and that potentially may influence the work involved in set extensions are shown in Table 1.

Topic	Considerations	Comments
Choosing Technologies	Practical effects (rain, snow, ice, fire), Animatronics, Stopmotion, Miniatures, Motion capture, Stereoscopy (Shot in 3D? Post-conversion?), Motion Control Camera etc	Shot execution varies greatly depending on the technology used. The motion control could remove the need for matchmoving all together but it is expensive.
Choosing VFX techniques	2D Compositing(Repairs/fixes, stabilization), 3D CG, Matchmoving(markers?), Matte Paintings etc.	Similarly, the shot execution will vary depending on the VFX techniques used. For some shots it's possible to predict which VFX technique will work well. This will have a large say in actual shooting.

Table 1 Pre-planning factors

Even if the mentioned considerations only contribute a small selection of what one should consider it is evident that they may influence a set extension greatly. Depending on the scope of the production in question and whether the director/producer have particular preferences in terms of technologies the VFX pipeline might change entirely.

5.5.2 Designing

The next phase is the actual *Design* of the various effects. Depending on what is decided in the previous phase a particular shot may involve full CGI extensions or perhaps it is decided that the duration of the shot is sufficiently short to make do with a 2D extension. Either way there is a multitude of considerations and actual factors that directly affect the work involved in a particular shot, as shown in Table 2.

Topic	Considerations	Factors	Comments
Overall	Shot dependencies	Scale, detail, speed, physics, length of shot, no. of shots	Quicker shots may be more forgiving. Also, if there are lots of shots it is important that the given effect is consistent throughout all(continuity)
2D	2D effects	lens flares, colour grading, repairs/fixes, stabilization etc.	
	Rotoscoping	Motion blur, duration, complexity, no. of wires?	
3D	3D CGI	Animation, texturing, rigging, modelling, rendering, particle fx	The amount of work involved in 3D effects is very dependent on the level of detail/quality sought after for each. Each of the factors may be subdivided into smaller segments.
	Set Extensions	Set Extensions - 2D or 3D, matte paintings, Greenscreen elements, no. of background elements, no. of foreground elements	
	Matchmoving	markers, parallax, handheld camera, enough features, resolution of camera, how much noise is present	

Table 2 Design factors

Even if Table 2. addresses considerations and factors for general VFX most are equally relevant in terms of set extension creation. As briefly mentioned in the *State of the Art* section (see section 5.4) the number of background elements in relation to the foreground could be of interest to the thesis. If it is found that the difference in terms of believability does not match the amount of the relative more *work* that goes into creating additional back- and foreground elements then it would be possible to suggest that the invested *work* could be spent elsewhere or saved altogether.

5.5.3 Acquisition & Projection

Acquiring the raw footage may also have a large say in the amount of work needed for a particular effect. If there are no VFX-oriented people on-set and therefore is a lack of proper reference footage or the camera-operator is unaware of important notions the VFX pipeline might be seriously hindered. Furthermore if the production goes ahead with shooting regardless of a sudden weather change from one shot to the next there may be put unnecessary extra work on the VFX department. In addition to acquiring footage it is also important to include the projection of the end result. Whether the end medium is a TV, mobile device or cinema screen, it may change the level of required detail altogether. A selection of considerations and factors

are put forward in Table 3 - similarly to previously, they hold true for set extensions even if they apply to VFX in general.

Topic	Considerations	Factors	Comments
The camera	Technicalities	Aperture(DoF), Exposure settings, Codecs', resolution, framerate, film vs. digital video, real-time/slowmotion, colour space, telephoto/wide-angle lens,	h.264 codec issues, Rolling shutter issue
	Movement	Zooming, camera steadiness, horizontal/vertical movement? Locked off?	Movement of the camera exhibits perspective shifts. Depending on this a set extension will vary greatly in difficulty. Little perspective: 2d tracking is enough, big perspective: matchmoving, 3d compositing etc.
	Physical Filters	warm/cool filters, skylight filters,	Sometimes the DP may choose to add filters to the lens to produce an effect like lens flares say. This could pose issues for the VFX shot.
On-set	References	No. of environment maps/ clean plates/lens distortion cards/location/texture references/elements/character scans	The number of references from set can drastically reduce the time spent of VFX afterwards
	Natural Conditions	Time of day, lighting, weather (rain/foggy/clear/snow etc.)	Depending on say the weather an effect may be difficult/easy to carry out
	Artificial Lighting	No. of light sources, quality of light	
Projection	The projection medium	TV/Cinema/etc.	The medium at which the shot is to be projected onto has a big say the level of detail necessary for the shot

Table 3 Acquisition and Projection factors

In terms of set extensions it proves particularly important to mention *camera aperture, camera movement* and *on-set references*. The aperture of the camera in question controls depth of field (DoF) which effectively may blur the background elements if so desired. This could have a large say as to the level of detail required for elements. As mentioned we found it interesting that the latest *Transformers* used particularly many elements and a shallow DoF could perhaps be put to use to counter this. Furthermore, the motion of the camera directly determines what type of set extension is required. If the camera is locked-off, 2D extensions are typically sufficient whereas if the camera is to be moved drastically then a full matchmoved solution must be sought. Finally, references on set are crucial to any set extension as they may be used to actually create the virtual set. If a greenscreen element is poorly executed on-set or if a particular HDRI environment map doesn't get recorded, this may pose unwanted strain on the VFX department. The remaining considerations and factors in the table are also important and should be kept in mind during production.

5.5.4 Perception/Believability

The definition of believability also presented some factors that need some to be considered a later point in the project.

Topic	Considerations	Comments
Perceived realism	Preconceived notion of the medium	Is the viewer aware that he/she is watching a film?
	Real things not always perceived as being real	
	Constructing own reality to fit context/Accepting unreal elements	E.g. accepting a dinosaur despite it not existing anymore, and we can only speculate how it exactly looked and moved.
Perception theory	How we recognize objects and perceive and categorize 'unknown' objects.	According to the theory of Conceptually Driven Processing our patterns in our stimulus input is guided by experience from past events, memories and general organization strategies.
Willing suspension of disbelief		

Table 4 Perception and believability factors

The main factors are listed in Table 4 and now need to be seen in a more specific role in relation to set extensions. The topic of perceived realism is going to play an important role in relation to both the creation and eventual testing of a set extension. The preconceived notion of the medium is likely going to be influential on the result, one thing is if the viewer is aware that he/she is watching a fictitious shot, but also to what extent the viewer is aware of the use of VFX. The perception of the real and unreal related to Prince's notion and the final definition also provides some valuable considerations that unreal objects can be accepted in films, but has to have some correspondence to the real-world. This is important to consider if it is decided to implement elements in such a direction. Furthermore perception theory on how we perceive and recognise objects will serve as a basis for further research into the perception theory regarding the relation between foreground and background in a scene, to get a better understanding on how believability might be changed when altering these.

As motivated by the initial problem statement we carried out a pre-analysis:

"Set extensions are highly sought but the artists have trouble keeping up. Is it possible to identify factors that lower the work required for a particular shot whilst maintaining high believability?"

The pre-analysis was used to identify parameters and do preliminary research in relation to VFX and set extensions in particular. Furthermore, we researched believability in order to get an initial idea as to how VFX shots are perceived and which theoretical considerations should be addressed. Evidently as shown in Table 1 through Table 4 the topics include a multitude of considerations and actual factors that contribute to the work involved in the execution of a shot. As partially motivated by *Transformers: Dark of the Moon*(section

5.4.2), the notion of a relation between the fore- and background elements was found intriguing as valuable work could be saved if fewer elements could be identified as producing the same relative level of believability for a given shot. For simplicity we acknowledge and consider all that is presented in the tables. However as a means to eliminate the many unknowns the focus shall be on the relationship between fore- and background elements while using the remaining considerations and parameters as guidelines for the design and implementation.

This leads to the formulation of the final problem statement:.

6 Final Problem Statement

To what extent will believability be affected when altering the elements in the foreground and background in a set extension?

6.1 Delimitation

Before proceeding with the analysis some limitations will be made in order to refine the scope of the thesis.

- Believability will be dealt with respect to visuals and visual perception. The creation of the set extension will similarly focus on the visuals. Therefore audio will not be focused on during the research, beyond fundamental necessities and implications that may arise.
- Furthermore the set extension will be produced to facilitate an experiment to examine the problem statement and it should therefore not be seen as a fully-fledged standalone film.
- While several factors can influence a VFX shot, as listed in table 1 through 4, the experiment will be aimed at altering fore- and background elements in relation to one another whilst keeping the rest constant wherever possible.

7 Analysis

The aim of the *Analysis* section is to fully explore and research the topics involved as set forward by the problem statement. It is evident that in order to actually answer the statement various tests must be conducted that in turn look to study the relationship between believability and foreground/background elements in a given set extension. The thesis may thus be regarded as an experiment of sorts wherein a product acts as a testing bed for the experiment rather than the documentation of creating a product and then testing that in terms of usability/viability etc.
In order to facilitate and carry out the necessary tests this section will cover and concretize the involved topics with respect to the final problem statement. In addition to such topics we will furthermore address *film theory* which will prove valuable come design and implementation.

At the end we will propose various hypotheses as based on the problem statement and as facilitated by the analysis which will dictate what product must subsequently be created.

7.1 Scene Perception

The following section will examine believability and perception in relation to set extensions. The aim is to establish a more specific hypothesis on how believability might be affected when altering the elements in the foreground and background. This should help in outlining the test. In order to establish this hypothesis, perception theory in regards to how high level scene perception ties in with believability, will serve as the starting point.

7.1.1 Semantic consistency and congruence

Semantics play a large role in how a scene is perceived and how easily objects are identified. Two terms, in regards to semantics, that will be discussed in this section are semantic consistency and congruence. It should be noted that the definitions for semantic related topic vary widely from paradigm to paradigm; semantic consistency and/or congruence might mean one thing within programming and something entirely different within perception studies (Spilotros & Parzy, 2010). The main paradigm for this thesis will be that of perception. As semantic consistency and congruence can be rather convoluted terms the definition of both, for this thesis, will be determined. While semantic consistency and congruence can be seen as two overlapping entities it is necessary to specify a clearer distinction within the context.

Semantic consistency goes by many different terms; Loftus and Mackworth (1978) denotes it as informative areas, Henderson et al. (1999) as semantic consistency while Davenport and Potter (2004) denotes it as semantic consistency but also a slightly broader term in form of scene consistency.

All of these terms deal with how objects in a scene relate to one another and the context of the scene itself, much like the definition of believability (section 5.1.3). An example of a semantic consistent scene could be that of a kitchen with a glass of juice on a table. Such an object is what one would expect/accept to encounter in such a context. On the other hand an inconsistent scene could be the same kitchen scene, but with an anvil (being the inconsistent object) on the table instead of the glass of juice.

Whereas semantic consistency in this context deals with the relation between an object and a scene, semantic congruence will in this context be used to describe the relation between objects. The formal definition is that of 'agreement or harmony' (Oxford University Press, 2012) and in relation to perception studies, *Spilotros* and *Parzy* define semantic congruence as follows:

> ***"Semantic congruence could consequently be defined as a notion of agreement, harmony, equivalence or correspondence between the meanings of several components."(Spilotros & Parzy, 2010, p.5).***

This is the definition that will be followed in this thesis, as it provides a useful complementation to semantic consistency. It should be noted that both semantic consistency and congruence will primarily be dealt with in regards to visual perception rather than other stimuli.

Davenport and *Potter* conducted a series of experiments focused on the identification of objects (Davenport & Potter, 2004). The participants were presented with a series of images that were either examples of a consistent or inconsistent scene wherein the consistency was based on the semantic relation between the foreground object and the background. A consistent scene could be the background of a football field and a foreground element of a football player, whereas an inconsistent scene could be a background of a church with a football player in the foreground. After a brief viewing of the images (80 ms) the participants were asked to name the foreground object. The results showed that it was easier to identify the object in the consistent scenes than that in the inconsistent scenes. Similar results showed in subsequent experiments where participants were asked to identify the background instead of the foreground object, as well in experiments where they were to identify both fore - and background.

In another research experiment regarding semantic consistency *Henderson et. al* investigated the effect on eye movement (Henderson et al., 1999). Two experiments were conducted: The first had the participants instructed to observe a line drawing in preparation for a memory test, the second experiment with the instruction of locating a specific object within a line drawing. The intent with these experiments was not memory testing nor object locating, but rather to see how semantic consistent and inconsistent scenes affects

eye movement. The hypothesis for both experiments was that an inconsistent object in the scene would increase the fixation period upon that object and act as an initial fixation point in the scene. The observations from the experiments were as follows:

> ***"(a) Initial fixation placement in a complex, natural scene is not controlled by a peripheral semantic analysis of individual objects in the scene, (b) Once an object has been fixated, the eyes tend to remain fixated longer on that object if it is semantically informative (inconsistent) than uninformative (consistent) in the context of the scene, (c) The eyes tend to return to semantically inconsistent objects in a scene more often than to consistent objects, (d) Search paths to a specified object tend to be shorter to objects that are consistent with the scene than to objects that are inconsistent with the scene. Taken together, these results support a model of eye movement control during scene viewing in which the eyes are initially driven by visual factors and global scene semantics, with cognitive and semantic aspects of local scene regions playing an increasingly important role as scene exploration unfolds." (Henderson et al., 1999, p.226)***

Henderson's et. al. research correlates with similar studies (Yantis & Jonides, 1984; Friedman, 1979). The thing to take note of is how semantic inconsistent objects can control the eye fixation and as Davenport and Potter showed, affect the initial object identification. These findings provide some interesting parallels to the established definition of believability. As *Prince* (1996) points out real world correspondence to the perceived object is an important aspect in order for believability and acceptances to work. In that regard semantic inconsistency might violate such a correspondence. A semantic inconsistent object can be recognisable and have a real world correspondence, but this correspondence might be violated if the object does not fit within the scene thus being semantic inconsistent. There is also a scenario where the object is semantic consistent and recognisable but still not believable. This is where semantic congruence plays a role, the case where the object does not fit in relation with the other objects e.g. different lighting, shadows, proportions etc. This is something closely connected to set extensions and VFX in general as it is often such semantic incongruence that can hurt the believability in the shot.

In relation to our definition of semantic congruence *Biederman et. al.* presents a list of *five Relational Violations* (Table 5) which as Biederman states *:*

> ***"...may be sufficient to characterize the difference between a display of unrelated objects and a well-informed scene." (Biederman et al., 1979, p.145).***

Relational Violations	Example of violation
Support	A floating fire hydrant. The object does not appear to be resting on a surface.
Interposition	Building in the background passing through the hydrant. The background appears to pass through the object.
Probability	The hydrant in a kitchen. The object is unlikely to appear in the scene.
Position	The fire hydrant on top of a mailbox in a street scene. The object is likely to occur in that scene but it is unlikely to be in that particular position.
Size	The fire hydrant appearing larger than a building. The object appears too large or too small relative to the other objects in the scene.

Table 5 Biedermans Relational Violations

These violations serve as an ideal basis for expanding on how semantic incongruence can influence a set extension and VFX shots. Within the VFX realm, technical factors such as shading, lighting and rendering may also be regarded as Relational Violations; e.g. wrong lighting of an object or shading that does not mimic the surface property it is intended to. This also correlates with *Rademacher's* study (section 5.1.3.1) (Rademacher et al., 2001) which showed that changes in texture and lighting of real objects in real photographs could result in a perception of the object being non-realistic.

Henderson mentions an interesting study made by *Guy Buswell* (1935) on how people look at pictures. Like Henderson's study it was an experiment conducted by eye tracking (although in Buswell's case on a much earlier state of eye tracking, with Buswell being one of the first to utilise a non-intrusive eye tracker (Babcock et al., 2002)) analysing people's eye movement when viewing pictures. Part of the conclusion, as Henderson also points out, is that people tend to fixate their view more on foreground elements rather than the background (in the pictures in Buswell's case on people in the foreground rather than the background).

7.1.1.1 Attention

Another interesting topic to examine is parts of the theoretical aspects of attention. Much like the eye movements used for the fixation described by Henderson, a similar phenomenon exists for attention. Having several different names for this, *Coren et. al.* uses the term *Attentional gaze* (Coren et al., 2004, p.393) which will be used here as well. Attentional gaze takes the assumption that one's attention can 'gaze' independently of where the eyes are looking. Naturally physical eye fixation and attentional gaze goes hand in hand as attentional gaze can be seen as an extension to fixation. The important part to remember is that while change in fixation often leads to change in the attentional gaze, the separation between the two still exists. Thus the attentional gaze can differ from the fixation which can have an influence on scene perception. This visual attentional gaze goes for auditory attention as well. It is also worth considering the fact that usually neither the eye movement nor the attentional gaze can be drawn to more than one point at a time

J. E. Hoffman conducted a study on the attention for different levels of details (Hoffman, 1980). The level of detail is ordered into two distinctions; the global level of detail and the local level. E.g. a face would have a global level (the head) comprising local levels of a mouth, nose etc. Subsequently a local level can become a global level and have its own local levels e.g. an eye would have an iris, pupil and so forth as local levels. While we are capable of selectively focusing on a global or local level Hoffman's study showed that when focusing on one level it is harder to process information from other levels at the same time. This draws some parallels to Buswell's point of people focusing on the foreground object rather than the background; one thing is the foreground object being the initial fixation point but it is also worth considering that it is harder to process information from the background at the same time (if we regard the fore and background as two separate levels) which in turn also correlates with the idea of attentional gaze.

Considering that the aim is to lower the required amount of work in a set extension while maintaining believability - if a foreground element attracts the eye fixation like Buswell points out, can the required background work be lowered? Attentional gaze and detail level perception could further support this hypothesis, as they also indicate that it is likely that the viewer will pay less attention to the background, if they are focused on the foreground. This is also supported by research in the field of *inattentional blindness*. One of the more prominent studies was conducted by *Daniel J. Simons* and *Christopher F. Chabris* (1999). Simons and Chabris conducted an experiment where the participants were to watch a video of two groups of people (black and white team) passing along two basketballs and count the number of times with which the white team passed the ball. At a certain point in the video a woman in a full gorilla costume walked through the playing field. Out of test participants (n=192) 46% did not notice the woman in the gorilla costume at all, indicating the inattentional blindness. A similar experiment conducted by Ulrich Neisser (1979) which served as the inspiration for Simons and Chabris yielded a similar result, and the experiment and results have been replicated since as well. This only adds to the possibility that less attention might be paid to the background in certain circumstances but also poses an interesting paradox regarding the previous discussed semantic inconsistency. Taking Simons and Chabris experiment and looking at it with semantic inconsistency in mind one would suspect that the gorilla would draw a lot of attention as it is a rather inconsistent object in the scene. To some extent it did as 54 % of the participants noticed it, but still the remaining 46 % did not and there is no telling if those who did, saw it immediately. One explanation is likely found in that the attention is on one very specific level of detail and on a very specific task. If the participants were not instructed to count the number of passes with the ball, but merely observe the scene, the gorilla would most likely have been noticed by everyone instantaneously.

All of this also ties into our definition of believability; the following section will thus examine the subject further in order to establish a possible test method for believability.

7.2 Believability Model

The following will once again examine believability, however this time with the aim of establishing a suitable test method. This will naturally be done in regards to the definition from section 5.1.3, but also by further examining studies within perception of realism and suspension of disbelief as they carry similarities to believability.

The definition of believability was as follows:

"Believability is to accept an element as being real within its own context and medium. Thus covering both unreal elements and truly realistic elements."

In order to gain a better overview, the definition will be combined suitably with perception theory (section 7.1) to form an appropriate model.

7.2.1 The Model

The following model is an attempt to connect the different elements from the believability definition and the perception theory. The model is rather abstract but it is meant to provide an overview of some key elements that are useful for testing believability. It is important to remember that this it is not an all covering model of the entire perception/psychological process, as that is a very complex matter. It should furthermore be mentioned that many of the topics may contain some overlapping as the process in itself is somewhat intertwined.

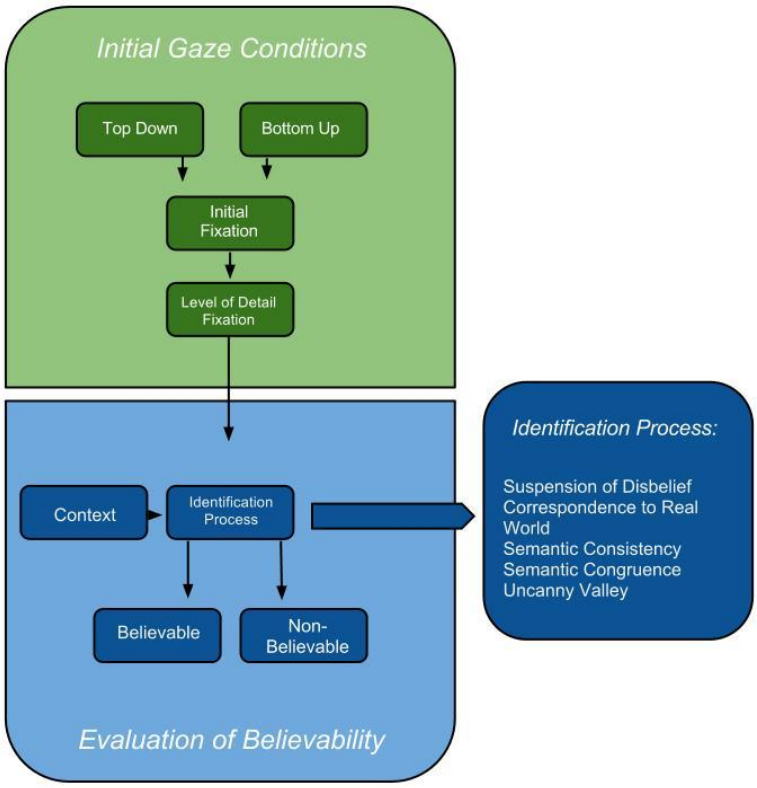


Figure 12 The Believability Model

7.2.1.1 Initial gaze conditions

The initial gaze conditions are the factors that influence the believability before and during the first look at the scene.

Top Down or Bottom Up perception

Dependent on whether the scene is perceived top down or bottom up can have a significant influence on the overall perception and thus the believability. *Top Down* perception is a task driven process (Hasic & Chalmers, 2007) where our eyes search for any given element related to the task. Such a top down perception is what Simons and Chabris described in their gorilla experiment (Simons & Chabris, 1999). As shown a top down perception can have a large influence on how the rest of the scene is perceived. Likewise Cater et al. presents a study where when given a search task people did not notice changes in two versions of the same image (Cater et al., 2003). This also plays into believability as while focused on one particular task the viewer might be less likely to notice other elements that could violate the believability. The task at hand is not necessarily only related to experiments such as locating an object or counting as with Simons and Chabris. In relation to film and VFX the task might as well be a result of the narrative, such as following a conversation, or other action performed by a subject in the foreground.

On the other hand there is the *Bottom Up* perception which is the stimulus driven process (Hasic & Chalmers, 2007), where our eyes and attention are drawn to size, shape, brightness and other factors that help make an object stand out. This would also include semantic inconsistent objects as pointed out by Henderson (Henderson & Hollingworth, 1999), and the semantic incongruence that often stems from the size, shape, brightness etc. An example of the *Bottoms Up* is what Buswell mentioned regarding eye movement and fixation when viewing pictures, where the viewer is more likely to fixate on people in the foreground at the initial gaze (Buswell, 1935). This is in correlation with the *Bottoms Up* perception, as a person in a scene (and the foreground) often carries different characteristics than that of the background in terms of shape, size, colour etc.

Fixation

The perception approach in the previous step is of great importance for the initial fixation point in the scene, which naturally determines which is processed first in regards believability. It should be noted that the fixation does shift during the viewing of the scene. However if the initial fixation is caused by a top down perception or a very dominant object in the bottom up perception, the fixation tends to remain longer on that given point, and tends to shift the fixation back to that point more often.

Level of Detail Fixation

The next thing to consider is the level of detail fixation as described by Hoffman (Hoffman, 1980). As Hoffman states; fixation on a particular level (being global or local level) makes it harder to process information from other levels at the same time.

7.2.1.2 Evaluation of Believability

The evaluation of believability is the phase where, as the name implies, the current element being perceived or the scene as a whole are deemed believable or non-believable. First of however is some external factors that is not a direct part of the perception, but rather a consideration of the aim and properties of the scene.

External Factors

As already discussed, the perception approach (top down or bottom up) depends on the intent of the scene. However the aim in terms of realism also plays a role. James A. Ferwerda presents three varieties of realism in computer graphics (which is relevant considering VFX is a part of computer graphics) being; *physical realism, photorealism* and *functional realism* (Ferwerda, 2003). Physical realism deals with an accurate point-by-point representation of reality. As Ferwerda points out this is a very computationally heavy process and often overkill as it tends to go beyond the eye's perception capabilities. Photorealism is the aim to create something indistinguishable from a photograph of a scene. A very common approach in VFX as discussed in section 5.2 and often the goal for invisible effects and set extensions. Finally Ferwerda mentions functional realism e.g.

the graphics in a flight simulator for the training of pilots, where the graphics is not realistic by the standard described by the two other definitions, but provides the necessary visual information required for the scene.

The interesting aspect to consider here is that believability can be obtained in any case, while the visuals in a flight simulator (functional realism) is not realistic as such, the result for the pilot can still be very believable. This goes back to the definition of believability for this project, mentioning that:

"Believability is to accept an element as being real within its own context and medium...."

Again, the keywords are that of context and medium. Before entering the flight simulator the pilot has a preconceived notion of the medium and context he/she is about to experience (knowing that it is a simulation) which helps to set the bar for what believability is in that case. Obviously in a flight simulation a variety of other stimuli such as haptic, tactile and auditory feedback is a vital component of creating a believable experience, which might also lessen the need for visual realism. But for film and VFX this preconceived notion of the medium and context is also an important thing. If one expects to see a science fiction film, spaceships and explosions in space helps on the believability, however if some expects an authentic drama film or documentary, spaceships and explosions in space likely have the opposite effect.

Even though VFX shot lies more in the realm of photorealism, Ferwerda's notion of different kinds of realism is interesting to remember as there is different kinds of photorealism within VFX depending on the aim, context and preconceptions of the scene.

Context

As mentioned in section 5.1.3. Møller & Pellengahr (2011, p.43) presents the Expectation and Belief Model following the structure of Torben Grodal's PECMA model (Grodal, 2006), which describes how we have expectations and notions on how things look, which serves as a basis for how believable we find the given thing. This is something that is very similar to the Uncanny Valley , as it also has a basis in the expectations about the human look. This ties into Prince's notion of relation to the real world, where it has to be considered that these expectations about objects are a part of the formation for a real world correspondence. This all ties in to the context of a given scene e.g. the believability of a fantasy scene would partially depend on what expectations the viewer has about fantasy elements. As presented in the Expectation and Belief Model, these expectations will be influenced by repeated exposure to a particular representation of a given object. This could mean that people that have a favour towards a particular film genre, and watch a lot films in this genre, could have a very different opinion of believability in that genre, than people who do not watch or favour the genre.

Identification

The final and crucial step in determining believability for a scene is identification which is an umbrella term that shall be used. All of the previous steps are essential for the outcome of this identification process. Identification consists of three main components: semantic inconsistency, semantic incongruence, suspension of disbelief and correspondence to the real world. Semantic consistency and congruence are recurring themes in this thesis so far, however in this part semantic consistency and congruence should be regarded as much more harsh entities. Here semantic consistency (or inconsistency) is the notion of whether or not a given object fits in the scene where semantic congruence is whether the object fits visually in terms of texture, lighting, size etc. This is a very crucial part for VFX and set extensions in particular, especially dealing with invisible VFX as they should be indistinguishable from the scene. Therefore if an object is perceived as being semantic inconsistent and/or incongruent to such an extent that it does not succeed under one's preconceived notion of the scene (see External Factors), the result is a non-believable result. Suspension of disbelief is a big part of film in general allowing us to accept small inconsistencies both, narrative and semantic, as well as unreal elements. For VFX and set extensions, suspension of disbelief is integral for the more visible effects, exemplified by CGI robots and monsters, and for set extensions elaborate science fiction or fantasy landscapes that obviously do not exist in real life.

As noted by Prince (1996) in order to accept unreal (non-existing) elements there has to be correspondence to the real world. This means that the object in question has to carry real life properties as well as follow physical law (unless explained otherwise by the narrative). This ties into the suspension of disbelief as these correspondences are a requirement for the suspension. A violation in real world correspondences also correlates with semantic inconsistency in the visuals (Biederman et al., 1979).

The entire process should be seen as a recursive process when viewing a scene, as the fixation can shift from object to object. But it can, as already argued, also remain on one object throughout the scene.

With this model established the question that still remains is how should believability be tested for a set extension? Semantic (in)consistency should not be defined in the sense that; semantic consistency = believability and semantic inconsistency = non-believability. But rather that in combination with the other elements it goes a long way in indicating believability. In a hypothetical test scenario the first step would be to establish such semantic inconsistency as perceived by the viewer. One way to do this is by examining the initial visual fixation point and why it was so. Was it a matter of a dominating object (thus intended initial fixation point) or because the object felt out of place or otherwise distracting. Taking a direct approach and asking people to describe and how they felt about a given scene might also provide some useful information. By keeping such a question rather vague might tap into the suspension of disbelief - if people are content to describe the narrative (if present) in a scene this could mean that the scene was believable enough for them to focus on what the scene was intended for. The downside to such an approach lies in the subjectivity of the

matter, and the implications of an experiment scenario. When watching a film at home or at the cinema, the suspension of disbelief, as well as a preconceived notion of this being a film comes naturally, and it can be questioned if this changes in an experiment scenario. A full walkthrough of how believability will be evaluated in terms of the thesis is described in section 8.2.

7.3 Film Theory

As both set extensions and VFX are an integral part of film production, film theory will now be examined in order to address considerations that need to be made before constructing the set extension.

7.3.1 Type of shot

The type of shot, for which the set extension in this project will be made, is an important aspect to consider. It is not possible to define which type of shots set extensions mainly appear as set extensions are applicable and used in a wide variety of shots. In order to narrow the project down, the following will focus on shots that allow for a set extension to play a more prominent role. Such a shot could be an extreme long shot or an establishing shot which would allow for a great display of the set extension itself. However wanting to explore the relation between fore - and background, a foreground element has to play a visible role as well. In that regards this does cause some limitations in shot selection as a close-up shot would reduce the visible background (thus the set extension) too much. Therefore the set extension in this project should be integrated in a long or medium long shot, to ensure a good visibility balance between fore- and background.

The foreground element also has to be considered. While the theory is that a foreground element draws the fixation, different elements will likely have different degrees of influence. It has to be assumed that a visually pleasing element (e.g. a beautiful person) will draw and ensure a longer maintained focus compared to a neutral or dull object (e.g. a grey box). As briefly discussed previously, action should have the same effect. It is well known that movement can attract the attention and fixation (Coren et al., 2004, p.391), but as discussed with the top down perception, if a task is present, attention and fixation will likely stay on that task. Therefore it seems plausible that if the foreground element performs an action, attention and fixation will stay on that.

This proposes some different scenarios; the effect of an appealing and unappealing foreground as well as a static foreground as well as one with an action.

7.3.2 Sound

Sound is without a doubt a large part of film and VFX, where sound may be used as an integral part to emphasise the effect at hand. However producing proper sound can require as much work as the work on the visual side. Since this is a project focusing on the visuals, sound will not be considered quite as much. The problem by disregarding sound is that lack of any sound, could cause quite some complications. If no sound at all is present in the shot, then all focus will be on the visuals thus making it more likely that visual mistakes will be noticed. The decision is therefore to avoid diegetic sound and use background music instead, with that said the choice of music have to be thoroughly considered, as it can have quite an impact on the mood of the shot.

7.3.3 Choice of Genre

The type of genre for the scene also has to be determined. As with the shot type a set extension does not afford one particular genre over another. Set extension have been used in everything from the fantasy genre (Game of Thrones (2011) over the historical genre (Boardwalk Empire (2010)) to modern action genre (Transformers (2011) and Science Fiction (Green Lantern (2011). The criterion for which the genre is chosen is therefore based on personal interest and in consideration to believability.

The believability definition states believability is:

"...to accept an element as being real within its own context..."

The context will often be defined by the genre and aided by a narrative. Therefore it must also be assumed that different genres would have different requirements for a set extension to appear believable. A set extension for an action film set in a modern urban environment could require a high degree of photorealism as such an environment is easily recognisable and relatable. For a fantasy setting there would be more creative freedom in terms of content, but that does not necessarily make it easier to achieve the believability as the notion of correspondence to the real world would have to be fulfilled. While it could be argued that most people would correctly identify a fantasy or science fiction, there is still room for a misinterpretation of the context (compared to contemporary modern setting, which mimics the 'real' world) which ultimately would hurt the overall believability.

With these considerations in mind the chosen genre will be science fiction, due to personal interest and content flexibility.

7.3.4 Science Fiction and its Elements

This following section will address the chosen genre and what it comprises. As the thesis examines fore- and background elements in a scene it is critical to know specifically which elements it is possible to alter.

Science fiction includes many different sub-genres including all-in robots/spaceships set in space type scenarios, post-apocalyptic dystopian low-key and all around mind boggling alternative universes. For these reasons there is no one true index of what a science fiction film may or may not include but common for most is the notion that the elements typically are unlike what a spectator observes in his everyday living. Two definitions, by Susan Sontag and Joanna Russ respectively, come close to the style that will be aimed for:

> ***"Science fiction films are not about science. They are about disaster, which is one of the oldest subjects of art."(Sontag, 1976 , p.116)***

> ***"Science fiction, as I mentioned before, writes about what is neither impossible nor possible; the fact is that, when the question of possibility comes up in science fiction, the author can only reply that nobody knows. We haven't been there yet. We haven't discovered that yet. Science fiction hasn't happened." (Russ, 1995, p.22)***

Sontag's definition is something very close to what can be described as dystopian and apocalyptic science fiction, and it is also work within this field (e.g. Blade Runner (1982), Children of Men(2006) and District 9 (2009) that will serve as the main design inspiration. The following pages will show a small selection of the visual style that inspires the project. At the end common elements and elements that can be altered in regards to the fore- and background, will be discussed.

Figure 13 Pictures by Laurent Menabe & Marek Denko

Figure 14 Pictures from Blade Runner © Warner Bros. Pictures

Figure 15 Pictures from Children of Men © Universal Pictures

Figure 16 Pictures from District 9 © TriStar Pictures

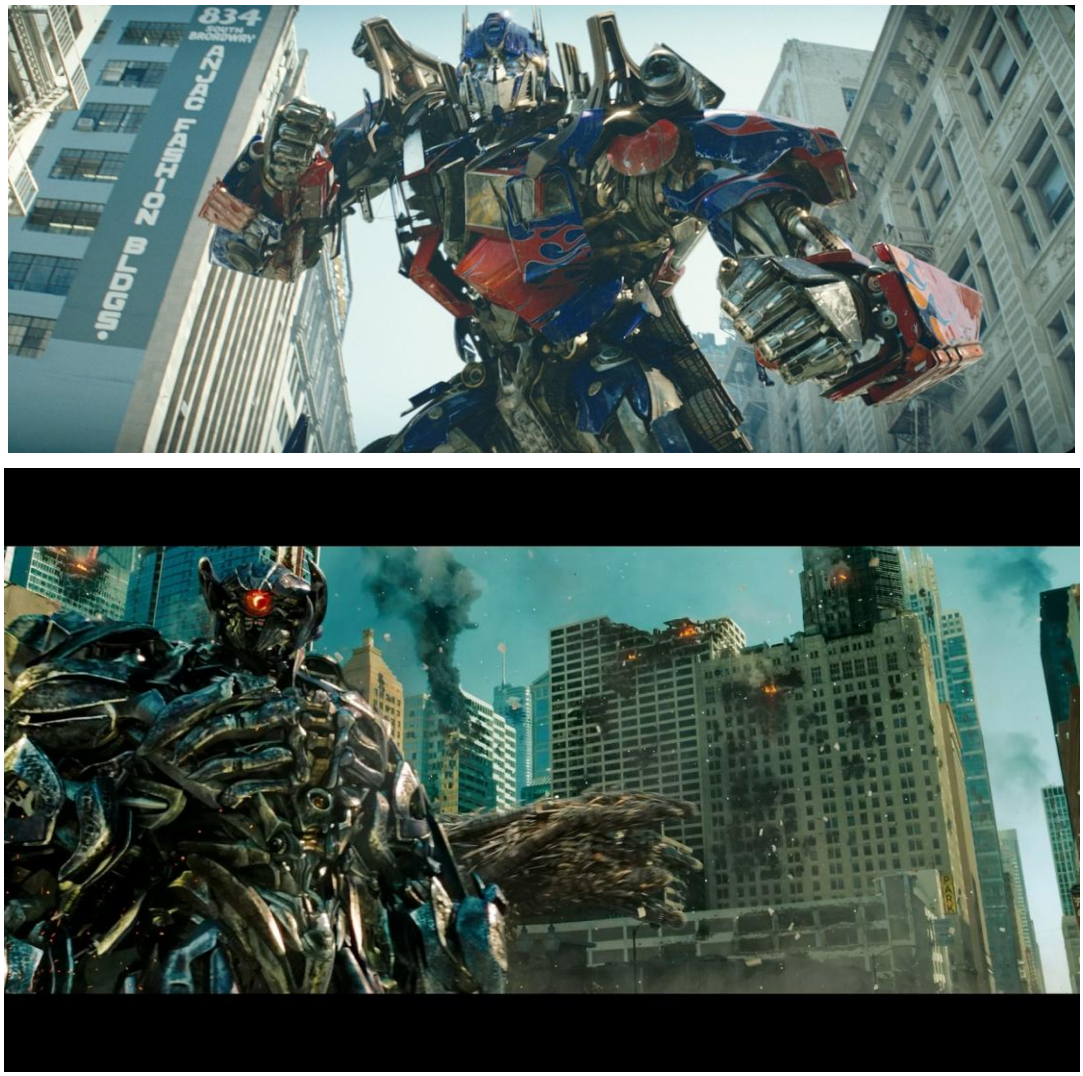


Figure 17 Transformers: Dark of the Moon © Paramount Pictures

Evidently science fiction is a broad genre entailing many different visual styles and themes. The following highlights a few elements that are put to use throughout the aforementioned examples. These will help when deciding which elements to alter for the fore- and background respectively.

- Flames/debris/cloudy/smoke
- Vivid colours/dark grey nuances
- Airborne Vehicles
- Varying sizes of Spaceships
- Buildings with large wallpapers
- Furnaces/Spiky towers/gleaming light sources
- Randomized cityscapes chaotically layout
- Machines/robots/aliens

These are all elements that could compose a set extension and be varied in terms of contents, which would be the basis for the different iterations of the product.

7.4 Analysis Summary

The final problem statement stated:

"To what extent will believability be affected when altering the elements in the foreground and background in a set extension?"

The Analysis section set out to fully investigate the problem statement. The following will conclude on the material discovered and break the problem statement into various hypotheses that in turn will motivate the design and implementation of a product to be tested.

Various theories were explored and a synthesized theory with an inherent model with respect to *Believability* and *Perception* was concretized. At the essence it was discovered that a term denoted *Semantic Inconsistency* was of particular interest to the thesis. This term in relation to the thesis' definition of Believability "*...Believability is to accept an element as being real within its own context and medium. Thus covering both unreal elements and truly realistic elements.*" proves crucial as the scene must appear semantically consistent for a spectator to fully believe in the scene. Furthermore it was discovered that if the foreground provides action or dialogue the spectator is likely to focus on this rather than a background. If a spectator focuses on a foreground is it really necessary to

spend the more work on a background set extension? In this is sense it is hypothesized that work could in fact be lowered without hurting a spectators overall level of believability for a scene if a foreground is present. Also, it was investigated how one evaluates a subject's believability for a particular scene. Very subjective in nature and particularly hard to quantify the section addressed a schema for evaluating such a situation. A full method of evaluating believability is described in section 8.2.

In addition to investigating theory regarding believability and perception the section also addressed film theory in regards to set extensions. The section defined dystopian science-fiction as being the genre of choice due to flexibility and personal interests. As with most genres, Science Fiction is a broad term and it thus had be defined within a certain context. A combination of definitions proposed by Sontag and Russ was ultimately used as the guideline for the thesis. The section went onto identifying parameters and elements which the Science Fiction genre typically supports. This proved valuable in terms of addressing which elements the Fore- and Background elements could involve.

The analysis strived to examine the problem statement extensively and it is thus at this point of interest to properly address and answer the problem statement. Broad in nature however, and inherently inclusive of a multitude of sub-topics we propose various hypotheses that in turn will constitute and address the sub-topics. Setting forward sub-hypotheses should aid us in understanding the effect of changing the relationship between fore- and background elements in relation to believability. Furthermore it will be easier to design iterations of the test product if we know exactly what is sought and why.

In order to evaluate upon the Problem Statement we outline the following hypotheses:

7.5 Hypotheses

The following will state and explain the hypotheses based on the findings from the previous chapters:

- If a foreground element is present a test subjects focus lies on that, especially if the foreground presents an action. Therefore the need for a complex background is lessened to achieve believability.
- Subsequently, if the foreground is not present the focus will lie on the background therefore a more complex background is needed.
- If the subjects fail to realise that the setting is sci-fi the execution of the set extension may be regarded as unsuccessful.

These hypotheses also create a basis for the design of both the test and video clip. There are a total of two independent variables (the fore- and background) as well as a dependent variable, being believability.

Wanting to have different iterations of both the fore- and background means that there has to be at least four different videos; two iterations of the foreground (with and without a foreground element) to test if a foreground would draw the fixation and attention away from background. Similarly the background needs to have at least two iterations in order to test if the foreground actually draws the attention away from the background. As the hypothesis states that it is an active foreground that should draw the fixation and attention (due to it affording a top down perception) it would be interesting to test if that is the actual case, or if the mare presence of a foreground object would give the same results. This gives three variations of the foreground and two of the background, for a total of six videos. Arguably several additional variations of the background could also be interesting to test upon, however a single additional level would give a total combination amount of nine videos. Therefore it has been chosen to keep the total amount at six videos, as nine could produce a rather complicated test.

7.5.1 Expected outcome

The active foreground case is the base of the initial hypothesis, stating that an active foreground will draw the viewer's attention, lessening the need for background elements to achieve believability. The difference in believability between high and low level backgrounds with an active foreground should therefore be minimal, as the attention should remain on the foreground. A similar result should emerge for the passive foreground with the difference being that there should be a slight increase in the believability difference. This should happen as the foreground is not as active, which should lead to the viewer focusing more on the background. In the final case with no foreground the believability should be entirely dependable on the background leading to the largest difference between the two background levels.

The hypothesis also leads to the assumption that in between the three foreground groups, there should be a difference in the believability. Overall the two groups with a foreground element (active and passive) should have a higher believability that the one without a foreground element. One thing would be the aforementioned theory that such an object will draw the attention lessening the need for background elements to achieve believability, provided that the foreground object is believable. Secondly the foreground object can help reinforce the context which the clip takes place within. As described previously, the context is of importance for the believability. There is also the case that even without a foreground element; the background (thus set extension) is good enough on its own to provide the correct context and a high believability. This is obviously the interesting part to investigate between the two editions of the video without a foreground.

In order to test these hypotheses the following chapter will discuss the test methodology.

8 Test Methodology

The following outlines the methodology behind evaluating upon the various hypotheses. Key points include the actual experimental design and how to evaluate the dependent variable, *Believability*.

8.1 The Experimental Design

8.1.1 Overview

The core of the experimental design is motivated entirely by the hypotheses. The main interest is centred on investigating the relationship between the fore- and background elements in a given set extension. More specifically we hypothesise that in situations where a foreground element is absent more focus will lie on the background and thus the level of detail should be the higher. Furthermore, it is a necessity that the audience realise that the clip is of a given genre as per defined by the synthesized model in 0. To realise these hypotheses and how well they correspond to reality an experimental design must be executed. Ultimately the scenario is one wherein two independent variables (that of fore- and background elements respectively) are to be investigated with respect to a dependent variable, namely *Believability*.

8.1.2 Choosing a Mixed Factorial Design

At the core of any experiment is the choice of which particular design paradigm to use for the case at hand. In many experiments a single independent variable is investigated with respect to a dependent but for the thesis the relationship between two such is of particular interest. A design method known as the *Factorial Design* proves particularly useful as it specifically evaluates upon relationships between multiple independent variables. Typically in the real world variables work together to produce a certain result and a factorial design is thus thought to produce richer and more insightful results (Cozby, 2008, p.186) as it examines exactly this. More specifically the design deals with levels of each independent variable with respect to each other formalized as e.g. 2x2 wherein each number specifies the number of conditions/levels that particular variable is divided into. This particular example has four test cases as the levels of the two variables are combined. Relative to the thesis the formalized format presents itself as a 2x3 factorial design where the prior number refers to two levels of background elements and the latter refers to 3 levels of the foreground elements.

Equivalently to other designs it must be decided whether test participants should be exposed to only single test cases or multiple during test sessions. As outlined in (Cozby, 2008, p.194) there are three main methods of going about this for a factorial design, namely that of *Independent groups, repeated measures* or *mixed*. The prior assigns each test condition with an independent group of participants whereas the repeated measures have

participants try all conditions. The latter, mixed, is in effect a combination of the two approaches where subjects will try multiple conditions within independent variable A, but are independent across levels of independent variable B. The three methods are outlined in Table 6.

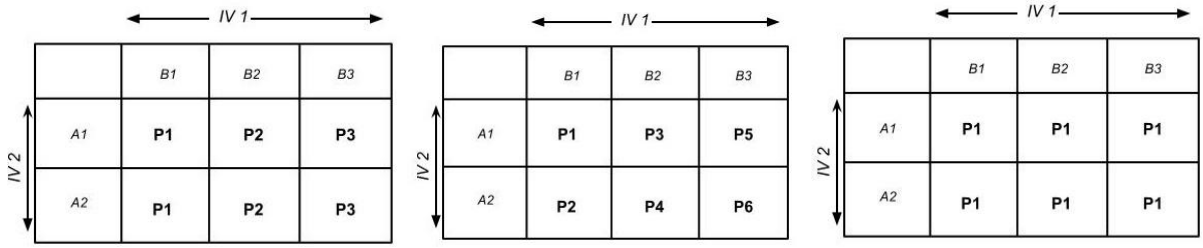


Table 6 Independent groups, Repeated Measures and Mixed designs

There are various advantages and disadvantages associated with each method. Most obviously as exemplified by the number of participants required for each method. The independent groups requires the most participants as each condition requires independent samples, whereas the repeated measures requires the fewest participants as all participants try out all conditions. The mixed design requires half of the participants. A repeated measures design furthermore has the advantage that the results will be somewhat more powerful as individual characteristics for the given participant will be eradicated. However a repeated measures design has the disadvantage that participants will be biased due to multiple viewings and might be subject to fatigue if they are to sit through many conditions and answering questions for each. One way to counter the inherent bias is make use of *counterbalancing* wherein the order at which each test condition is presented to the participants is randomized (Shuttleworth, 2009). With enough test participants the bias should thus be minimized.

With this in mind, *The Mixed Factorial Design* is chosen as the main design method for the thesis. The three levels of foreground elements will thus be independent groups whereas the 2 levels of background elements for each level of the foreground will be repeated measures. Counterbalancing will be used to minimize the inherent bias from watching two similar clips. The design as chosen for the thesis is summarized in Table 7.

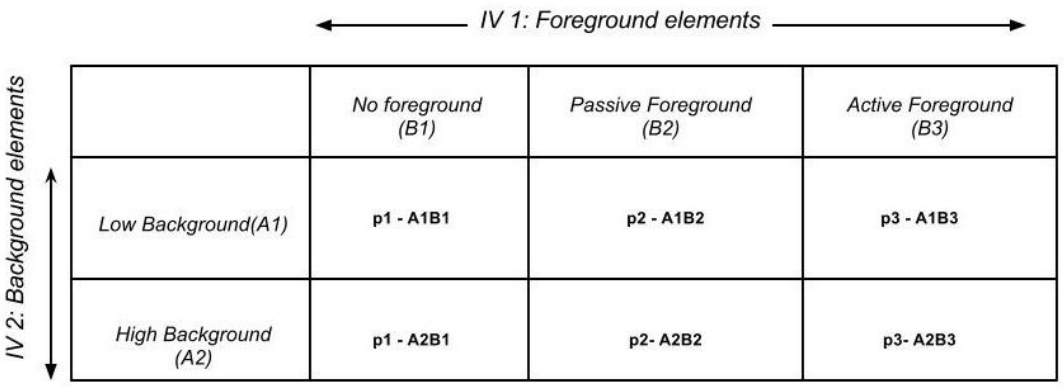


Table 7 Illustration showing the final Mixed Factorial Design in terms of Fore- and Background elements

8.1.3 Qualitative vs. Quantitative

Critical to any design is which type of format the results are and how they are gathered. Two such formats are that of quantitative and qualitative. The prior could include the use of questionnaires with numeric predefined answers whereas the latter typically includes face-to-face interviews, focus group meetings and the likes (Cozby, 2008, p.134). The main difference between the two is that quantitative results will follow a predefined format whereas qualitative results will contain much richer content but the results will not follow a predefined format. From an analytic point of view quantitative data is therefore much easier to analyze with respect to statistics and the likes whereas it is difficult to apply a statistical approach to qualitative results as the content will vary from subject to subject. To analyze qualitative results experimenters will typically focus on extracting themes and tendencies based on transcripts of the discussions.

As the dependent variable, *believability,* is highly subjective a combination of quantitative and qualitative is thought to provide for the most optimal results.

8.2 Applying the Believability Model

In order to apply the believability model in testing the problem statement and the hypotheses, it is necessary to outline how these elements compose a testable whole. The initial fixation should be established as this can be a good indication of any semantic inconsistency and/or incongruence. This depends on what is assumed that the viewer will fixate on. In the case of the active and passive foreground it would be expected that the initial fixation would be on the foreground. If it turns out that the fixation rather is on a background element, it could be a sign of semantic inconsistency and/or congruence which ultimately can hurt the believability. In

that regards it should be remembered that the foreground element also can be semantic inconsistent and or incongruent, which can be just as hurtful to the believability. The case with no foreground is a harder case to predict the initial fixation as there will be no decisive foreground object to draw the fixation. In this case, top perception (following the foreground element) could be substituted for a bottom up perception, thus leaving the viewer to gaze freely through the scene.

There are different options for determining this initial fixation. One way would be to use eye tracking and another would be self reporting from the participants.

The eye tracking could provide a rather precise image of the initial fixation as well as additional information of the view pattern through the entire duration of the clip. The downsides to this approach would be that on a logistical scale the experiment would be confined to a lab experiment but a limitation that would be possible to work around. The self reporting would require the participants to describe what they initially fixated on. This would give a greater flexibility as no special equipment would be needed, thus granting the possibility to reach a broader test group, in a shorter amount of time. Furthermore some information on the general believability could be acquired in this fashion; if the participants were to state that: ***"I noticed an object in the background, because it felt very out of place and did not fit."*** This would be very valuable information. The downside would be that there is no guarantee that what the participant report is what they actually fixated on first, and general consistency (or lack thereof) between the different participants in their way of reporting, can be a problem. While both have some distinct advantages and disadvantages it has been decided to use self reporting, as it is felt that eye tracking may limit the amount of testing possibilities on a logistic basis.

Going back to the believability model, the level of detail fixation poses an interesting element to support the hypothesis. If the viewer does not notice elements in the background this could be due to the fact they focus on another level (the foreground) making it harder to process the background.

Again this can and will be tested through self reporting; asking what the participants notice visually besides the initial fixation as well as asking them to describe what to place during the clip. This should indicate to what degree people report objects from other 'levels'. In theory if the participants focus on the foreground level throughout the duration of the clip, they should be less likely to report on every detail in the background. Asking the participants if they noticed specific things, would be an option but if the participants are to see several videos, being aware of such an question would make them more likely to notice them the second time, beyond the bias of having seen a similar video once. So once again the reliance has to be on the self reporting.

The perceived context is also important to be informed of. The thing is that if a Science Fiction shot (as it is in this case) is found believable by a viewer, but the context/genre is seen as a comedy there is a conflict with the definition of believability. The definition states that:

"Believability is to accept an element as being real within its own context and medium…"

Thus if a VFX company gets the assignment to create a believable Science Fiction scene, but if the scene is perceived as a comedy scene, yet still a believable one, the company has failed as they have not created a scene that is believable within its own context (being science fiction).

Judging overall if the shot is believable or not could be a case of directly asking if the viewer found it believable or not. However there would be several implications with such an approach: First of, it would require a unified understanding of believability. Secondly if the viewers are to watch several different videos in the test, asking directly about believability after the first video could shed too much light on the purpose of the test, thereby creating too much unwanted bias for subsequent videos.

Therefore a more subtle approach is required. Combined with all the other questions this will be approached by asking the participants to rate the given video on how much they liked it as well as explaining why they rated it as they did. This obviously implies a relation between enjoyment and the believability of the videos, but enjoyment of a video and even a VFX shot naturally depends on more than just the believability, everything from sound, mood, actors etc. has to come together. With that being said the aim is actually not just to hope that people rate the videos as being very likeable and comment that the video was believable, but rather to see why people rate the videos low, if they do so. The theory is that believability is sufficiently important, that a vague believability alone can drag down the overall impression. Therefore if the score of how well the participant liked the given video is low, it becomes a matter of identifying clues that can indicate that the believability was responsible. This means going through the aforementioned questions to get indications of semantic inconsistency or incongruence, judging how the context was perceived (did the participants experience it as a science fiction genre) and generally analyzing the responses.

8.2.1 Questionnaire and survey

As defined a combination of the quantitative and qualitative approaches is to be used. Alternatives exist but for simplicity each participant will be required to fill out a questionnaire with questions that examine their perception of a particular clip. As a means to not lead subjects onto what is being researched the questions must not be too specific even if this means that the interpretation of the results will be the more difficult. The following presents the five questions that the main questionnaire will include with justifications for each question. The following questions is based on how to apply The Believability Model.

Q1: What was the first visual thing you noticed?

This question deals with the initial fixation point as previously explained in the model. It is important to ask this first so they do not forget. It is hypothesized that they are less likely to pay attention to the background if a foreground is present. Furthermore if the foreground presents a motion or action there will be further attention paid to the foreground.

Q2: What else caught your attention visually and why?

This question contributes to determining whether the hypothesis *"If there is a foreground (with action), processing of background information is lessened, thus the need for elements in the background to create a sci-fi setting is lessened...."* is true or not.

If a participant focuses on the set extension it could indicate that it is subpar executed and stands out thus reducing a subjects believability. This however depends on the scenario - if there is no foreground chances are that they will focus on the background which should require a better set extension. This has to be correlated with their opinion of the scene overall. If they comment that the effect is 'cool' or 'very well done' then it must be reasoned that believability is still uphold. We can thus not guarantee that they won't look at background. However as mentioned, if the set extension is poorly executed there is a chance they will notice such as it will stand out.

Q3: Explain briefly what happened in the clip.

If a participant does not understand the narrative of the clip the execution of the clip must be thought to be poorly executed. In accordance to the model for believability it is important the *'Contex'* is upheld.

Q4a: Please rate how well you liked the clip. 1 being "Did not like it at all" and 5 being "liked it a lot"

It can be reasoned that if a particular test participant does not like the clip he/she will be more likely to find flaws to justify their opinion which in turn also contributes to their *suspension of disbelief*. Furthermore if believability is low it will be assumed that the score in this question will be low as well, since a low believability can ruin the shot.

Q4b. Comment on your answer.

As the actual answer to *Q4a* is numeric it is hard to know why participants gave that particular score and this question thus asks them to explain why. This will be valuable when determining whether a particular subject finds the clip believable or not.

Q5: What type of movie genre did the clip belong to? (Put one mark)

This question directly determines whether participants realise the genre of the clip. As per defined in the model it is critical that it is *"...believable within a context..."* where the context is that of a genre. If a participant does not identify the clip as science fiction it must be thought that the execution of the set extension is failed.

In addition to the aforementioned main questionnaire, participants are also required to fill out a quick survey that allows us to profile them according to film taste, sex/age/occupation and the device with which they conduct the test on.

8.3 Procedure

In order to carry out the experiment two similar approaches are conducted. The questionnaire and the six different clips will be hosted online so that a large number of people will be able to conduct the test, and to streamline the data gathering. As with any experiment the larger sample of a population the more likely one is to be able to conclude on results as big random fluctuations in answers will influence less. As there is no target group per se it is equally important to have young people partaking in the test as older people. Even if their answers will vary we are predicting that the trends in regards to the hypotheses will be equally represented no matter age, film taste, device etc. This is addressed through the survey which participants are required to fill out during the test however.

Practically participants will be friends and friends of friends, as obtained through Facebook and the likes. The questionnaire will be hosted on Google Docs which in turn will include links for YouTube where the six videos will be hosted. In addition to the 'Facebook' sample we will furthermore conduct a similar test at the AAU and ITU campuses. For these participants will be asked to take the exact same test online.

It should be noted however that we have no control whether people watch the clips multiple times at home even if the questionnaire specifically tells them not to. During the campus tests they should be more inclined to following the instructions as we will be close by and able to resolve misunderstandings.

Thus to summarise the methodology for evaluating upon the hypotheses is as follows:

- The use of a 3x2 Mixed Factorial Design wherein each participant will watch 2 clips.
- 6 different clips as combined through 3 levels of foreground elements and 2 levels of background elements
- Application of the Believability Model on each test case individually

9 Design Requirements

As motivated by the aforementioned hypotheses, the following will present the requirements for the design.

Hypotheses requirements

- An identifiable genre chosen as that of dystopian Science Fiction
- 6 versions of a similar short clip that supports the use of set extension
- 2 different levels of background details
- 3 different levels of foreground details
- Ensure that the different levels of details are distinguishable

Personal requirements

- Using readily available low budget equipment
- Obtain a result of a respective standard
- To go through the entire process of 3D compositing with respect to camera projection

10 Initial Design

The following chapter will outline the initial design considerations made in order to construct the required set extension. This will be based on the design requirements in order to create a sufficient basis for testing the proposed hypothesis. Therefore this chapter will cover the overall setting and conclude with the different design consideration for testing each hypothesis. This is only the initial considerations as the design may change during the actual implementation.

10.1 Overall setting

As stated in the design requirements the genre will be that of science fiction, with a sub-genre of dystopian science fiction. Still this leaves a broad spectrum of possible settings, however the decision has been to use a beachfront as the setting. The reasoning is the flexibility such a setting grants, in form of an open view over the ocean which allows for an unconstrained set extensions. A set extension in e.g. a city could create some strict limitations on where the extensions could be implemented. Although such limitations could serve as a framework for what and where to implement, it is opted to go for the flexibility of a beach. The theme will be

that of a science fiction invasion, in line with the dystopian sub-genre and the quote by Sontag (see section 7.3.4). This also gives some ideas about the overall colour scheme which will be the dominant colours. Based on the science fiction pictures in section 7.3.4, and our understanding of dystopian science fiction the base colour will be a grey blue tone. Figure 18 show two possible colour schemes based around such a tone.

Figure 18 Monochromatic and Complementary colour scheme based on a blue-grey colour

To go along with this dystopian setting it has been chosen to use a handheld camera movement in the shot, this accomplishes two things: First of it means that the aim is of a more gritty style, thus it should be possible to accomplish with consumer level equipment. Secondly the camera movement allows for motion parallax in the CGI elements, and provide an interesting challenge for the matchmoving process.

The most important thing to keep in mind is that the videos should fulfil the elements examined in the believability model, meaning that:

Every element should be semantic consistent and congruent, thus fit within the context of the scene as well as in relation to each other. Therefore the element should be created in a similar style to ensure congruence, and various compositing techniques such as colour correction and matchup of blur and grain will be deployed to make the elements semantically consistent within the scene.

The elements that will construct the set extension will be slightly dependable on the final shooting location by in an ideal situation the location will have a clean horizon which can be built upon. The main component that will constitute the set extension will be a variety of spaceships that should populate the sky. This also ties in with the believability, in that the genre should be clear for the viewer, and spaceships should help emphasise the scene as a sci-fi scene. These are also the background elements that will be altered as described in section 7.5.

The foreground element(s) could be taken in several directions, one would be to shoot a foreground object on set, another to shoot the foreground object on a green screen and composite it in final scene or make the object pure CGI. Here the choice has been made to go full CGI. While the initial concern can be that a fully CGI foreground object is a rather prominent element that can hurt the believability if not done properly. However as argued in section 5.1.3, the degree of realism required to achieve believability can fluctuate depending on the scene. Therefore a fully CGI foreground object provides flexibility and the possibility to

create an object that fits with the background elements adding to the semantic congruence of the scene regardless of the location, and once again emphasise the sci-fi genre.

As the foreground is going to be fully CGI it will not be a representation of anything organic in order to avoid the uncanny valley and in line with the science fiction genre the object will be that of a robot. While a robot is a rather broad term, inspiration from the dreadnought's (see Figure 19) of the Warhammer 40.000 © Universe will be used.

Figure 19 The Dreadnought from the Warhammer 40k Universe

While the shot will not have a greater narrative to follow (as it will not be a part of a longer film), there will still be simpler short narrative. The idea is to have the robot move towards the waterfront and shoot down a large spaceship that moves towards the camera.

10.2 Designing the Iterations

In considerations to the hypothesis, several iterations of the shot will be made. To summaries there are two independent variables being the foreground (with an active foreground, passive foreground and no foreground) and the level of contents in the background (high and low), thus the set extension. Beginning with the active foreground this will be the one with the walking robot, the passive will see the robot in a stationary position but still shooting at the spaceship as well. Finally there is the iteration with no foreground where the robot will be absent. Since the spaceship, upon which the robot fires, is in interaction with the robot, it will be regarded as a foreground element, thus be removed in this iteration. If the spaceship was crashing without the presence of the robot, it would most likely grab the initial fixation and act as a foreground element in that regards.

The different iterations of the background consist, of either high or low amount of background content. The exact content will vary depending on the exact shooting location, but ideally the low detailed background (thus the minimum set extension) will have a futuristic cityscape in the horizon. This will either be created as an addition to existing buildings or created from scratch as a matte painting, again depending on the shooting location. In the case with a higher detailed background additional elements will be added to the set extension, primarily spaceships hovering in the background as well as spaceships making flybys.

Eventually these two independent variables will be crossed giving a total of six different videos, each with their level of fore- and background.

11 Implementation

The following section will present the implementation as motivated by the previous design overview. As mentioned the implementation will vary as dictated by the actual shooting location. The execution is divided into various sections that are organised after context rather than chronological order as most elements overlap as to when they were implemented.

11.1 Overall process

As the design proposed using a beach for the setting we set out to scout for suitable locations. Various alternatives were visited and by the end it was concluded that *Amager Strand* was closest to the ideal scenario. Following shooting various takes on location a single shot was chosen as the hero plate (the shot that is chosen as the one to work with). Central to any VFX shot, especially ones that involve inserted CGI, it is critical to compute a match move to get camera movement and a pseudo 3D representation of the real footage. Based on this it was thus possible to start mocking up the overall structure of the scene. Within 3D compositing software we setup an environment that enabled the positioning of elements with respect to a moving camera. Concurrently, various versions of animations and 3D models were created and refined as needed. The final scene was thus stitched together iteratively. The version with which we finalized at was the most complex one involving high background details and an active CGI foreground. To produce the five other versions it was therefore only a matter of removing elements from this version as per required.

11.2 Acquisition

11.2.1 Shooting

As briefly mentioned various locations were scouted before deciding upon the final. The ideal scenario was thought to be a beach setting overlooking an otherwise clean horizon. As proposed in the design the short

narrative of the final scene would portray an invasion of sorts and it was thus important that the location supported this. A few of the alternative beach settings are shown in Figure 20, and the final setting is shown in Figure 21.

Figure 20 Various Shooting locations

Figure 21 Final Shooting Location

The final location was suited for the narrative and the concrete pier (which had not been proposed in the design) was furthermore aesthetically in accordance with the overall aim and possible to integrate into a military type setting. A concrete ground would also be easier to deal with in terms of CGI insertion as compared to a sandy beach that would have to act according to the Dreadnoughts steps etc.

11.2.2 Technicalities and Panoramic HDRI

As part of personal requirements for the implementation was the notion of using low budget equipment wherever possible. The camera was that of the *Canon 550D DSLR* medium range readily available for the

consumer market. Even if DSLR cameras have known issues with regards to artefacts(Lancaster, 2010, p.158) and the rolling shutter issue (Lancaster, 2010, p.210) it was deemed interesting to see whether it was technically possible to produce a result of a somewhat high standard only using such equipment.

The Canon *550D* is able to shoot *bracketed* images which refer to shooting images at different exposures. Combining these 3 images to a single will result in a *High Dynamic Range, HDR, image* which is ideal for lighting CGI realistically. In order to obtain a full 360 degree panorama of the scene we positioned the camera on a tripod at roughly the position of the robot. Shooting 12 sets of bracketed images rotating the camera respectively each time would provide us with a panoramic HDR from the actual scene. As will be described later this was used to light the inserted Dreadnought realistically.

The film was shot on a 18-55mm standard Canon Kit Lens at 18 mm with f/8 to minimise the DoF.

11.3 Initial processing

11.3.1 Match moving

As soon as the footage was finalized upon we commenced *match moving* the shot. Match-moving is a process wherein the characteristics and motion of a real world video camera are computed. This process is fundamental to inserting CGI into live footage. The output from the match-moving software is a null object that comprises the motion of the camera and other computed properties, e.g. the horizontal and vertical Field of Views. Using the information in a 3D application allows for the CGI objects of interest to behave correctly according to perspective shifts and movement of the camera. Such software will furthermore compute the relative 3D positioning of points to one another which composes the *Point Cloud* of a particular shot. This point cloud refers to an estimation of the 3D space and may be directly used to position 3D objects such as the Dreadnought.

Typically such software will allow the artist to combine an automatic track with a supervised track. Supervised tracking refers to the artist manually selecting points to track and is useful in difficult sequences. The final sequence was difficult in nature as it was a single take, 40 seconds long and very shaky due to being handheld. Furthermore the shot was not stationary, but rather covered a good deal of physical ground. The camera was thought to follow the robot in position whilst walking behind it. For these reasons the sequence was difficult to properly track, but as it was the building block for the entire set extension, the necessary time was spent. Figure 22 shows the final setup within *Syntheyes* (the software used for the matchmoving) where each point corresponds to a location in the real world.

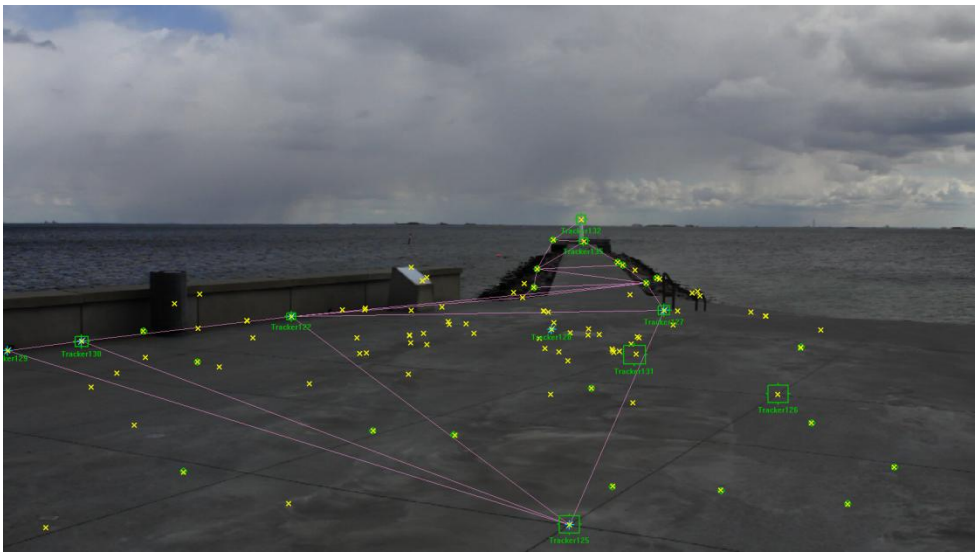


Figure 22 The Matchmoved camera. Each point represents a location in real space

Supervised tracking was used to ensure that the *Point Cloud* includes points located on the ground which eases the task of properly positioning the Dreadnought. If a plane is created that corresponds exactly to the real world ground the inserted robot will remain in position and not exhibit any sliding. A poor matchmove will exhibit exactly this and even small errors are apparent to an audience. This relates to Biedermans Relational Violations (see section 7.1.1). As according to the proposed Believability Model such inconsistencies are important to reduce if the scene is to be believable.

11.3.2 HDR processing

An essential part of any CGI live action integration is that the CGI lighting corresponds to the live action plate. One way to partly ensure that at least the base ambient light is correct is to make use of a HDR image. As briefly mentioned it involves shooting a sufficient number of images ideally around the nodal point of the camera and then combining these to produce one large 360 degree panorama. To produce the HDR's Photoshop was initially used to combine the three exposures of each (one proper exposed, one a single stop underexposed and one a single stop overexposed) of the 12 images to produce 12 HDR images. *Autopan Giga, software* specialized for Panoramas was then input the series of HDR's and seamlessly stitched them together to one whole. The thing that makes HDR panoramas highly usable in CGI integration is that general light direction, ambient light and reflections are present 360 degrees around the object in question. The panoramic image is shown in Figure 23.

Figure 23 Panoramic image of the actual shooting location used for lighting the CGI.

11.4 CGI

Central to the shot was the requirement for CGI to be inserted. This involves two categories namely that of spaceships hanging in the sky and the Dreadnought to be inserted in the foreground. As the workflow varies for the two the following is divided for simplicity.

11.4.1 Spaceship models

The scene includes various renditions of spaceships - those in the far background and those in the middleground doing flybys. The former will not require any CGI integration per se, as a simple 2D image will suffice (see later section on camera projection) whereas the latter requires a CGI pipeline. The actual pipeline varies from a traditional CGI pipeline wherein the renderer outputs an image sequence of the CGI that corresponds to camera movement as it makes extensive use of camera projections. It is effectively setup within the compositing software rather than the 3D software (see later section on camera projection for details).

However in order to make use of the camera projections the 3D package *Cinema4D* was used extensively to texture and alter models as obtained at www.turbosquid.com. Modelling was not at focus for the production at hand and it was therefore deemed suitable to use already existing models. They were however combined and altered in various fashions to produce a small arsenal of different ships of different sizes which was then followed by creating suitable textures. The panoramic HDR was setup as the environmental light source for the 3D scene and what followed were six orthographic renderings (see Figure 24) at a 4K resolution from each side of the spaceship in question.

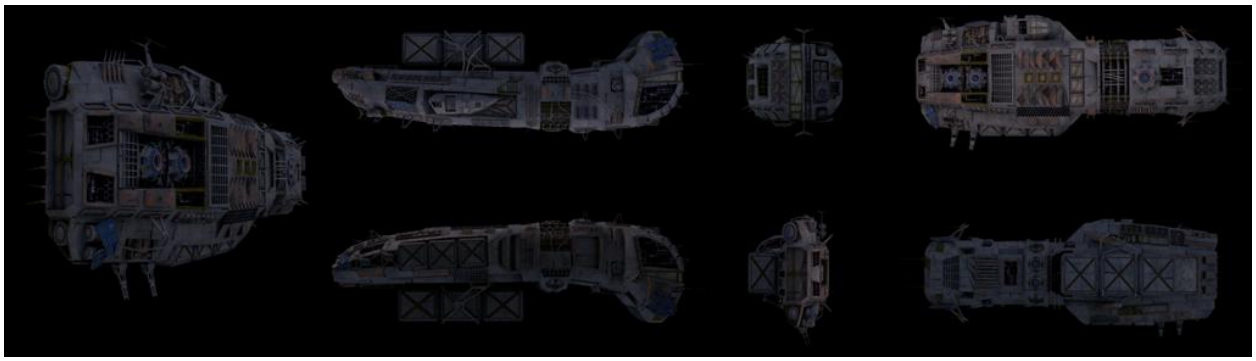


Figure 24 The spaceship and its orthographic renderings

11.4.2 The Dreadnought

The implementation of the Dreadnought followed a more regular CGI pipeline. Similarly to the spaceships, www.turbosquid.com was used to find a starting point with respect to the getting the basic geometry and the general 3D scene was setup within *3ds Max 2012*. To get the Dreadnought positioned in space and correspond correctly to the real world camera movements the match move solution was imported. This ultimately setup a scene wherein a virtual camera with parameters equivalent to the real world camera is positioned relative to the point cloud and ground plane as previously defined.

Various early tests with respect to general positioning and size of the Dreadnought were carried out to get an idea of what worked. The overall process from that point was a series of iterations including various versions of animations, textures and renderings.

The design proposed two sets of animations for the Dreadnought; the Aim/Shooting sequence and a walk cycle. As with animations for most CGI pipelines it is necessary to *rig* a character prior to animation. Rigging is the process of creating a skeleton for a geometry mesh followed by binding parts of the mesh to their respective skeletal counterparts. The result of this process is a mesh that will move accordingly to properties of the skeletal. Figure 25 shows the bone system that enabled carrying out the required animations.

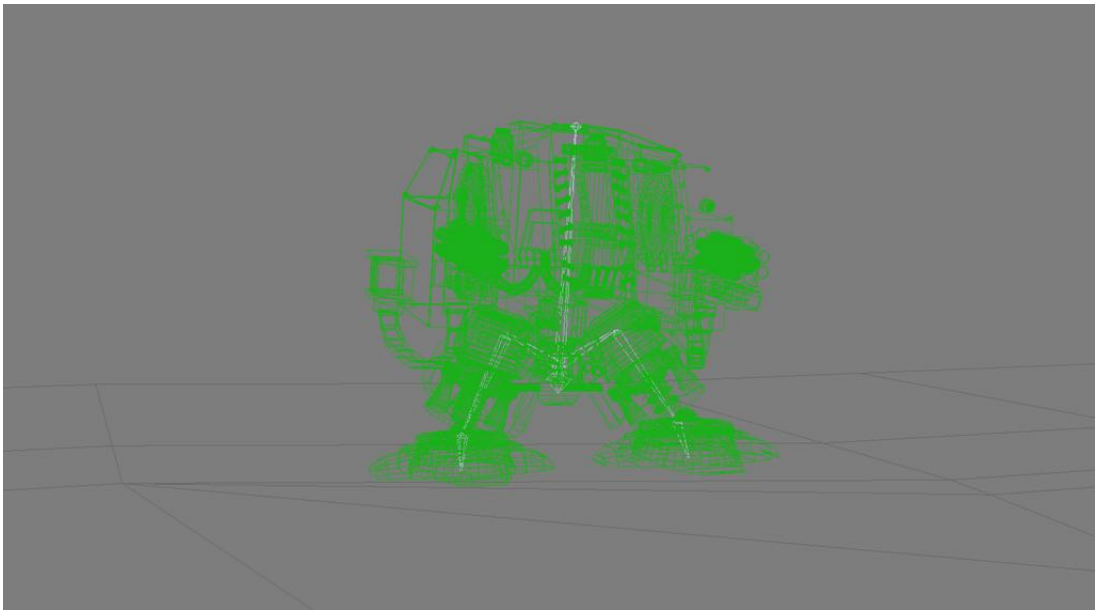


Figure 25 The Skeletal rig for the Dreadnought. The rig is composed of the grey bones visble.

The four distinct poses of a walk cycle, *Contact, Recoil, Passing* and *High-Point* were used to ensure that the flow of the movement was as natural as possible. One may decide to just start animating a character but making sure that the rhythm and timing is consistent throughout the cycle will make things much more efficient (O'Conner, n.d.)An example of this is shown in Figure 26 which show the position curves for the left and right feet. When one foot rests the other will travel and vice versa.

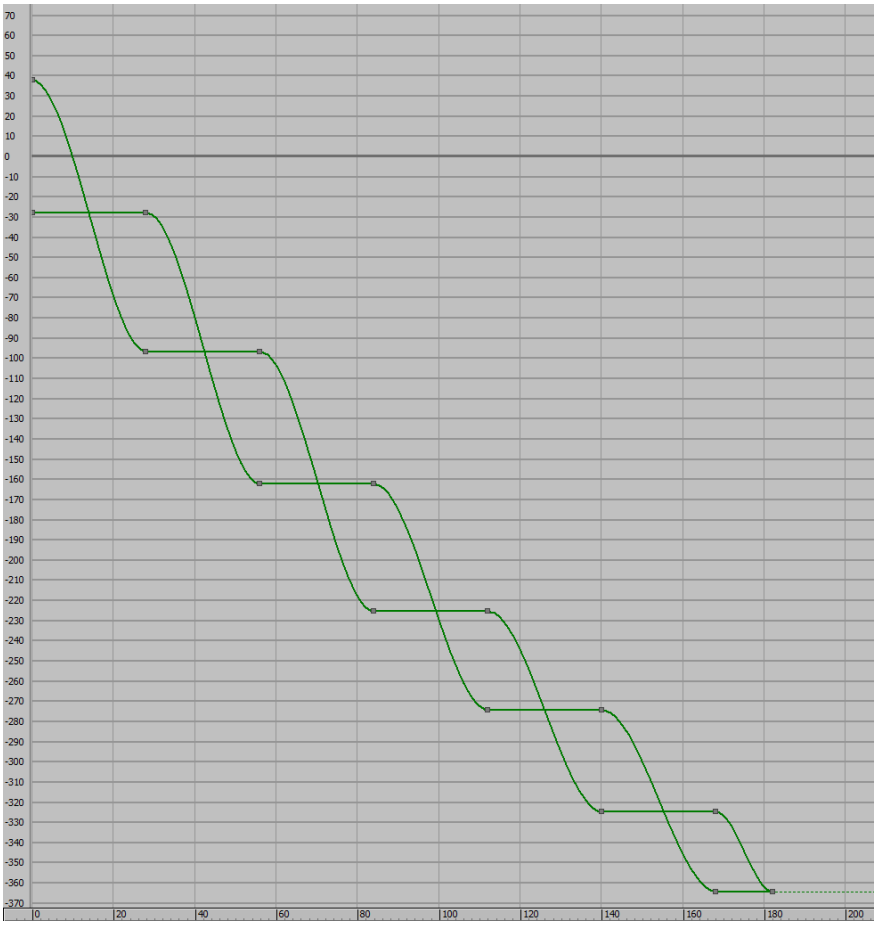


Figure 26 Position curves for the Dreadnoughts left and right feet in the animation walk cycle. When one foot rests the other will travel and vice versa.

Particular time was spent creating textures and realistic lighting for the Dreadnought. The textures involved combining a series of procedural materials to obtain a rusty type surface and bump maps were put to use to give a sense of volume to the actual metal. The panoramic HDR was put to full use to light up the scene and an extra directorial light source was added to mimic the sun and to cast appropriate shadows. The last step of the pipeline was to render out the appropriate image sequence to be composited with the live action sequence. Careful attention was paid to tweaking the Global Illumination in order to properly mimic how colours bleed onto one another and for this the *Vray* renderer was chosen due to its high quality output. However, as with all CGI that involves Global Illumination and animated sequences it is notoriously difficult to obtain results that are noiseless/grainless in lesser lit areas. Throughout the production period this proved to the biggest time sink of all. It was of critical importance that the CGI objects would not stand out in terms of graphical errors as this would contribute to inconsistencies with the overall scene and thus reduce the level of believability. The actual renders were computed as multi-channel EXR files including individual shadow/diffuse colours/global illumination/ambient occlusion passes. Separating the channels that make up the render in such a fashion make for facilitated compositing as each channel may be corrected individually to

ensure maximum flexibility for integration into the live action sequence. The various passes are shown in Figure 27 and the final render result for the Dreadnought is shown in Figure 28.

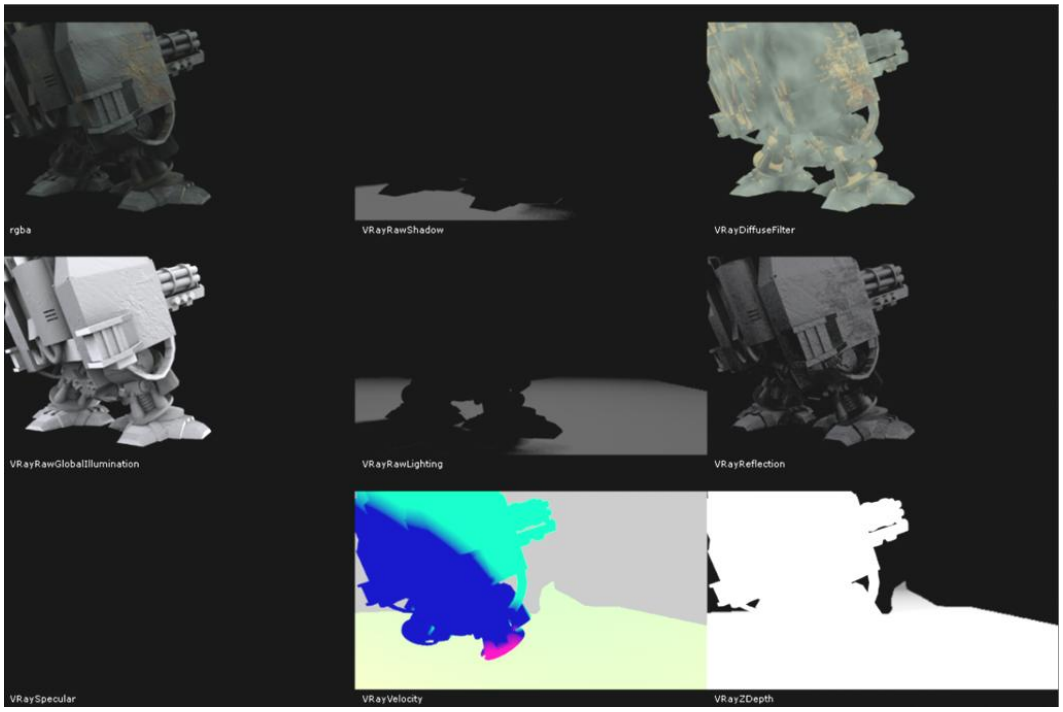


Figure 27 The multichannel EXR render passes for the Dreadnought

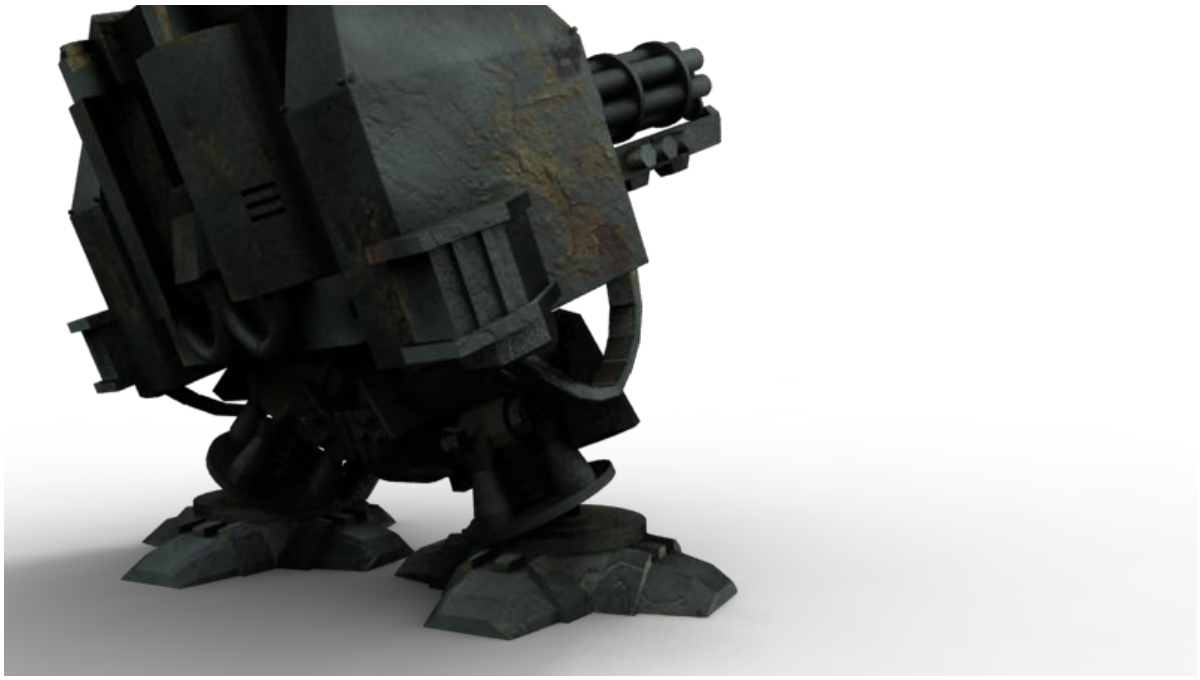


Figure 28 The Final Render

11.5 Matte painting

In order to create the background a traditional digital matte painting was put to use. As mentioned in the pre-analysis (section 5) objects in the far distance will not exhibit any parallax and it therefore suffices using mere 2D even if the scene is 3D. In order to emphasize the narrative in which an imminent invasion is about to take place it was deemed appropriate to replace the sky with a more dramatic one. As we filmed additional footage on the other locations we found a different sky that would work well. Apart from the sky replacement, mountains were added on either side and a cityscape was added to the left side. The individual elements were comprised from a wide variety of sources and they were effectively meant to make for a more interesting scene the final background as shown in Figure 30. The final result of the previous steps was a 2D still image that in turn was to be projected in compositing. See section for *camera projection* for details.

Figure 29 The final matte painting

11.6 Camera projection

One of the key interests for the thesis was the concept of camera projection. As previously addressed (section5.3) camera projection allows for highly flexible workflows and as such is used in a large variety of situations. A typical use is in 3D compositing applications where individual elements are positioned with respect to a virtual shooting camera. The process of using a camera projection setup typically involves the actual projection camera (includes lens properties etc.), source footage, and rough geometry onto which the source footage is projected at. Depending on where the element is to be positioned the accuracy of the geometry differs greatly; in the far distance a 2D plane suffices whereas the middle- to foreground will require more accurate geometry.

11.6.1 Projection of the Matte Painting onto Sphere

Camera projection was amongst a multitude of things used to project the matte painting onto an elongated sphere in the scene. An analogy of the far background of a scene may be regarded as that of a curved surface and it thus makes sense to use a wide sphere. When viewed at an angle that differs from the shooting angle of the source footage the end result will gradually become distorted due to the perspective change but in most situations it proves surprisingly flexible. Figure 30 shows the background setup in place.

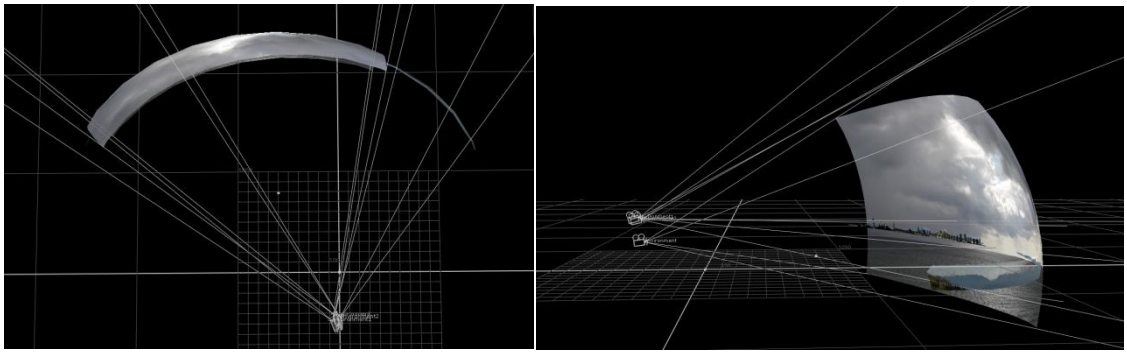


Figure 30 Matte Painting Projection

11.6.2 Multi-camera Projection of the Spaceship Flyby's

Where the matte painting projection and other individual projections only involve a single projection camera there are situations that call for more complex setups. The High Background scene involves a fleet of spaceships that cross the field of view at a relatively close distance to the shooting camera and a simple one-projection camera does therefore not suffice. The multi-camera setup involves 6 cameras that each projects an orthographic render onto geometry that represents the spaceship. The end result is in effect a 3D object with high resolution renders projected onto its geometry which can be positioned anywhere within the 3D environment. The multi-camera setup is shown in Figure 31.

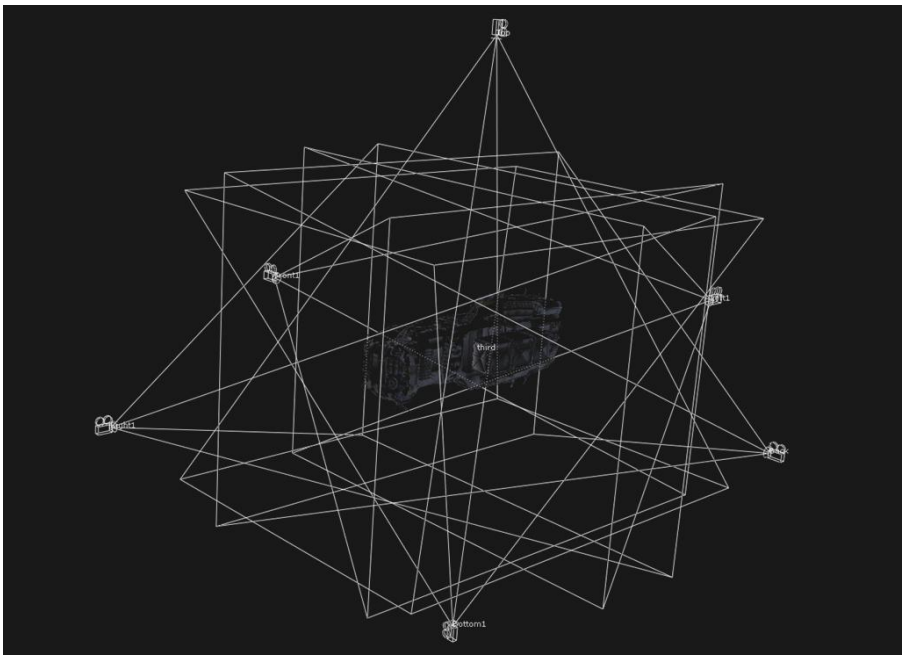


Figure 31 The Multi-camera used for projecting orthographic textures onto appropriate geometry.

The main advantage of this approach is that the object itself can be animated and positioned as desired whilst not requiring the hugely time consuming 3D rendering process as exemplified by the Dreadnought. If

possible the same approach would have been applied to the Dreadnought but as it was very close to the camera, required within-mesh animations, and minute details were essential the approach was not suitable.

11.7 Compositing

This final section will go through the process of compositing all the elements arriving from the various sources into one whole. The software NukeX 6.3 was used for all compositing related tasks. Compared to other tools (e.g. After Effects) that also allows for compositing, NukeX includes the opportunity to composite elements in a real 3D environment as compared to regular 2D and 2.5D.

11.7.1 The 3D Setup

Central to the compositing of the various elements a rather extensive setup was used. The overall process involves importing the matchmoved camera and then using this as the main camera for the scene. All the individual elements (matte painting background, far distance spaceships, middleground flyby etc.) are then positioned appropriately in the 3D environment. What follows is a re-photographing of all elements back to the 2D. The virtual matchmoved camera moves exactly as the real world camera and it is therefore not necessary to actually use the original source footage at all from that point on. If desired the camera could be positioned in any other scene and it will thus still look as if a person walks forwards in a shaky fashion. The key advantage of this approach is that all objects do not move in space (except for animated ones) but due to the moving camera the result will exhibit appropriate perspective changes. Figure 32 shows the overall setup seen from a few different perspectives. As the camera moves parallax will be inherently present and the objects at the far distance will thus move less than objects closer to the camera.

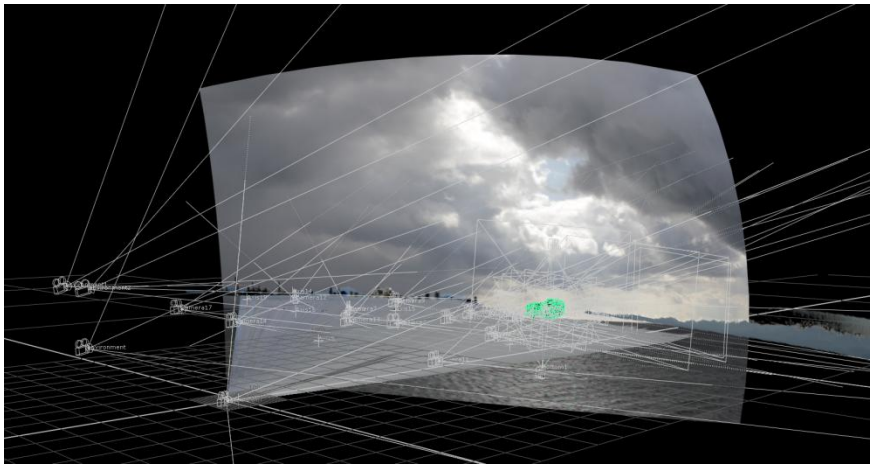


Figure 32 The 3D setup showing how elements are positioned in relation to each other

11.7.2 Back to 2D Compositing

As mentioned, the matchmoved camera is used to re-photograph the entire scene which brings back the 3D elements back into a 2D image. Where 3D Compositing is mainly about setting 3D elements up in terms of parallax etc. 2D compositing makes sure that all the various elements are well integrated and that the 'whole' carries the desired mood and ambience.

General workflow involves going back and forth from 3D to 2D repeatable to see how elements are positioned and timed in regards to one another. When that is in order all elements must be colour graded to match the background plate. Such procedures involve changing the gamma/grade values to match the overall tone of the image and *lightwrapping* to use any bright background colours to aid the integration. This whole process is a lengthy one but very important as it will ensure that nothing 'sticks' out.

When integrating CGI renders into live footage it is often a good idea to use renders at a lower resolution compared to the resolution of the live footage. The reason is simply that the imagery tends to be too 'perfect' and smooth where the live footage will have flaws. The actual workflow for the Dreadnought was rendering at 800x450, reformatting/resampling that to 1280x720 and downsizing the live footage from 1920x1080 to 1280x720.

When dealing with CGI renders there is always the unfortunate chance of noise in the imagery due to the way Global Illumination approximations work. After 4 days non-stop rendering of the Dreadnought and its movements the result had unfortunate noise in its dark areas. Noise is very apparent and a spectator will immediately notice and as such must be avoided. As time was a constraint it was no option to do a second batch and an alternative approach had to be sought. NukeX has various degrain/denoise and a combination of these was thus used to get rid of most of the noise. One downside with such techniques is that they often blur out the imagery to counter the noise and much of the 'lushness' of the original renders is thus lost. What followed was a *regraining* of the entire image using filmgrains. Apart from giving the film a certain aesthetic look it also helped the integration of the Dreadnought with the rest of the elements. Conclusively, time was spent grading the final image which involved bringing out more contrast and thus giving the image a rougher more dramatic look. Figure 33 and Figure 34 show the final image before and after regrained grading.

Figure 33 Before grading

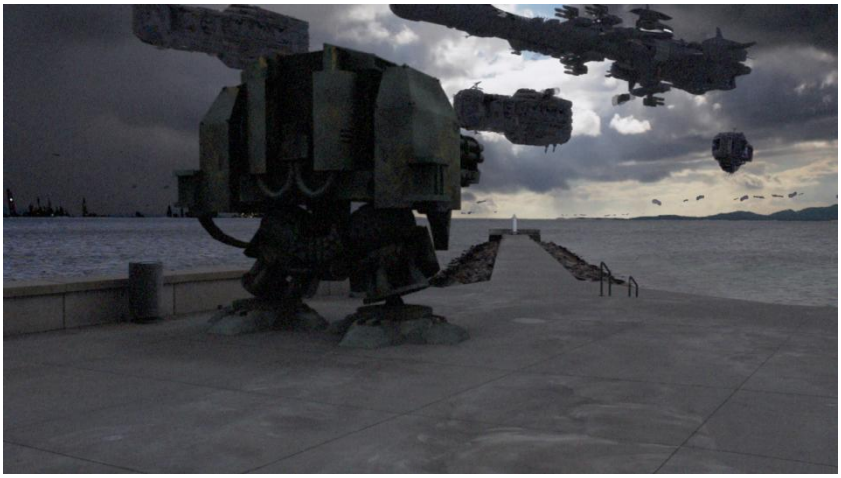


Figure 34 The final re-grained and graded image

12 Test Results

The purpose of the following section is to present the results that were obtained throughout the experiment in an objective manner. Analysis of the results with respect to the problem statement and respective hypotheses will follow in the *Discussion* section.

12.1 General overview of the obtained Data

After the 6 iterations to be tested had been designed and implemented we carried out two similar tests over the span of three days. The main test was conducted online through friends and friends of friends and shall be referred to as the *Peer Group Test*. The secondary test was conducted at the AAU and ITU campuses on students and this shall be referred to as the *Campus Test*. In total 103 people partook in the *Peer Group Test* and 41 people partook in the *Campus Test* totalling at 144 people as illustrated in Figure 35. It should be noted that

each person would watch 2 clips in 1 single test session resulting in 2 sets of main questionnaires and 1 survey questionnaire per person.

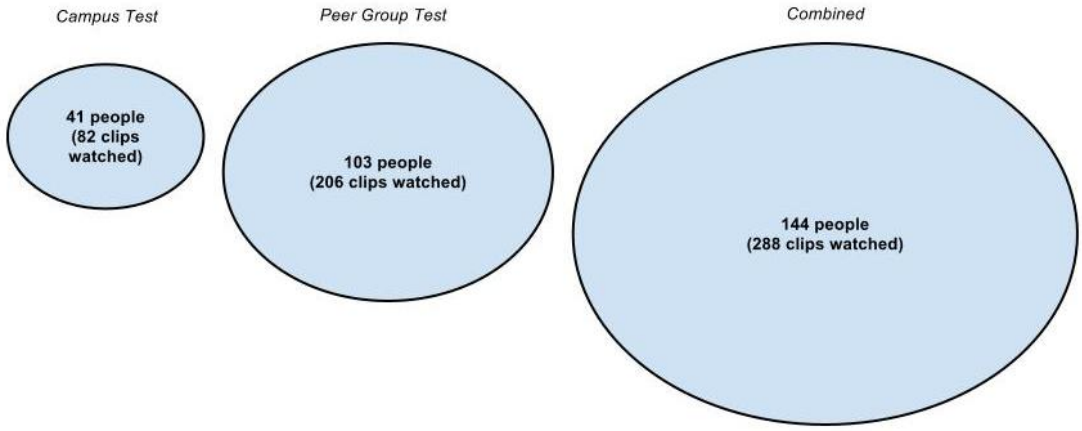


Figure 35 Overview of how many people were in each Test, as well as the number of video clips watched in total

Even if the execution of the two groups differ slightly, one conducted in people's homes and the other conducted at campuses we argue that as we do not have an actual target group we may combine the two samples. It may also be argued that the campus results are biased as we as test conductors were practically close to the participants but we made sure that they were at ease and did not feel pressured. Also, it should be mentioned that the campus tests were not conducted in a controlled laboratory but were rather done wherever the students happened to be. As the tests were not conducted in a strictly controlled environment there might be external stimuli that affect the results as exemplified by people not necessarily only watching each clip only once. However, we judged that obtaining the larger sample (possible through call-to-actions on Facebook etc.) was more important than restraining such issues.

Referring back to the Experimental Methodology section we carried out a *Mixed Factorial Design* which involved a combination of independent groups and repeated measures. The first independent variable, (going forth shortened to IV) (foreground elements), involved 3 separate main groups(No Foreground, Passive Foreground and Active Foreground) which followed an independent groups design. The second IV (background elements) involved repeated measures for each of the 3 iterations of foreground in terms of High and Low backgrounds. An illustration showing the overall structure of how this relates is shown in Table 8.

IV 1: Foreground elements

IV 2: Background elements

	No foreground (B1)	Passive Foreground (B2)	Active Foreground (B3)
Low Background(A1)	p1 - A1B1	p2 - A1B2	p3 - A1B3
High Background (A2)	p1 - A2B1	p2- A2B2	p3- A2B3

Table 8 The relationship between the two independent variables

P1 refers to one group of people that watches one level of Foreground and two Levels of Background in one test session. Whether they watch Low- or High Background first is randomized to reduce the bias of watching close to the same thing multiple times. B1, B2 and B3 refer to the 3 independent groups and A1, A2 refer to the repeated measures design. We made sure the distribution was as equally spread as possible in terms of test participants.

Each cell in the matrix, and thus the dependent variable, was thought to include a single number which effectively was how many found a particular clip believable.

12.2 Changes made to the Experimental Methodology

Prior to carrying out the experiment we researched how to validate the results in terms of inferential significances and it seemed there was a multitude of options. However what we did not realise at this point was how one would go about significance testing when dealing with other variables than continuous. In the case of our experiment, the two independent variables are categorical in nature and the dependent is in effect binary. Recall that according to the believability model a person can either find a clip 'believable' or 'unbelievable' which is equivalently to 1 or 0 and thus binary (Michael, 2002). The outcome for our test is thus effectively frequencies of a binary variable as divided into the six cells in the matrix in Table 8. Traditional significance tests such as T-Tests, ANOVA's and so forth assume continuous variables(e.g. measurable heart rates) and therefore cannot be applied to the experiment with the exception of mean values for how participants 'liked' the clip. (StatSoft, Inc, n.d.).

It was found that the main issue at hand was not as much the trouble of having a dependent variable comprising binary frequencies, but rather the very nature of a *Factorial Mixed Design* due to the use of repeated measures. If one is to compute statistical significances for continuous data arriving from *Factorial Mixed*

Designs, alternatives exist to the traditional ANOVA approach namely that of the *Factorial ANOVA* as outlined later. However this is not as readily possible for data that is binary in nature. Even if approaches such as the *Chi-square* exist to look at relationships between two or more independent binary variables a common critical assumption for most is the prerequisite of *independent samples* (Borkowski, 2010) which rules out procedures that involve repeated measures. It is however possible to test for significance for *paired-samples* and thus within each main group through the use of the *McNemar* test as will be described later.

As such we were forced to find an alternative route. We chose to divide our original pool of 144 participants(288 clips) into 6 *independent* groups of participants as based on which particular version the participant watched to start with thus totalling at 144 clips instead. Recall that each person would watch two clips in a test session namely a Low- and High background version and that we used counterbalancing(randomisation of which version a participant watches first) to reduce the bias of watching two similar clips. We were thus forced to subdivide the pool in order to eliminate the need for repeated measures. The original pool is shown in Table 9 and the final pool is shown in Table 10.

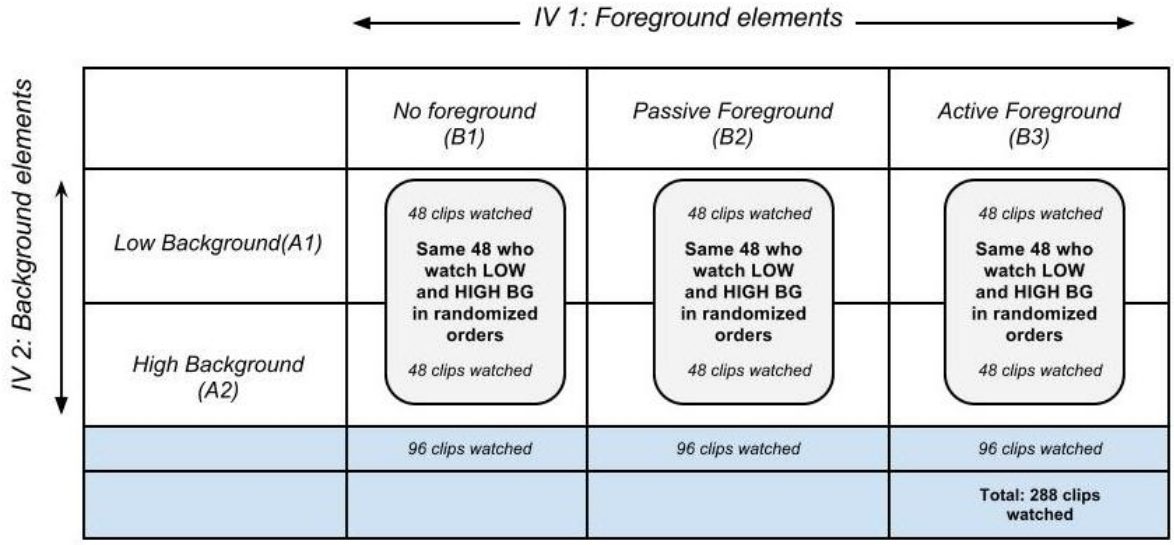


Table 9 Shows the original spread of participants for the 3 main groups - 48 participants in each main group watching High and Low backgrounds respectively in randomized orders thus totaling at 288 clips watched

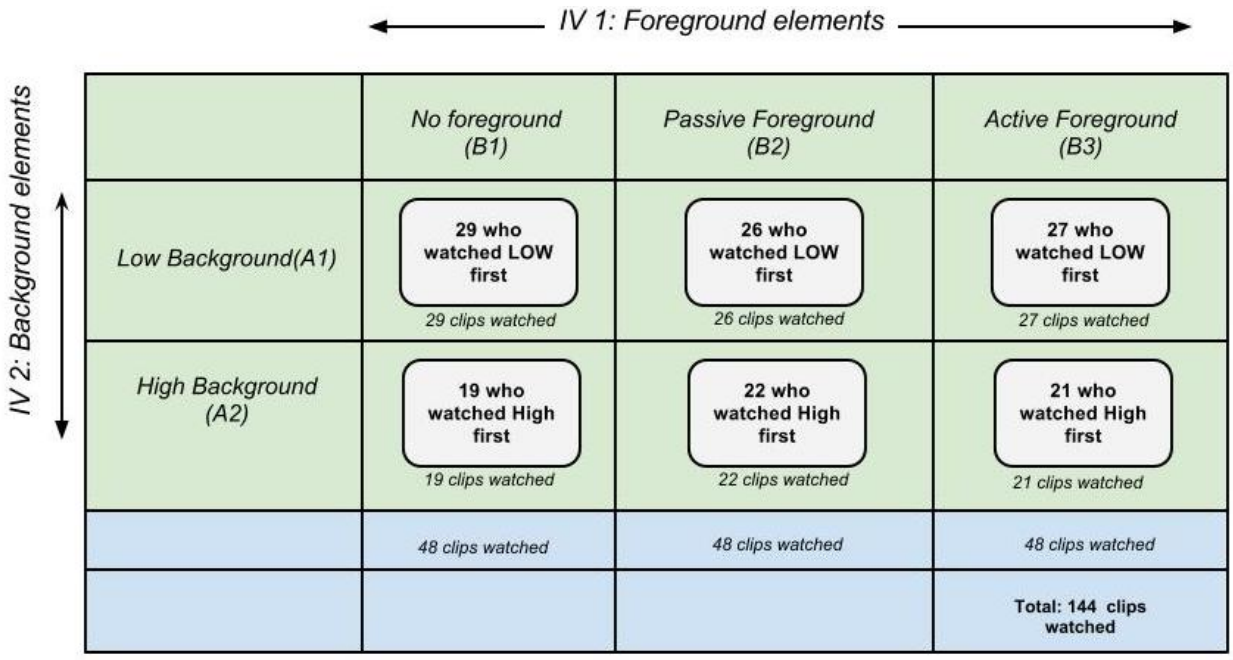


Table 10 The final spread of effectively separating the original single sample into 6 independent groups of participants who watched a single clip first to eliminate the need for repeated measures

By halving the original sample we effectively lost half of the samples but as we were fortunate to have a relatively large sample to begin with it was not thought to be of critical importance. Most importantly; the resulting set of data was six independent groups of results for each of the 6 conditions thus allowing for significant tests across the variables.

12.3 Significance of results

As already mentioned we used three different approaches to test for significant differences and relationships.

12.3.1 Chi-square

For testing the significance of the overall relationship across the two independent variables and thus covering the factorial nature of the experiment the *Chi-square* was identified as suitable. Recall that the results were halved due to the assumption of independent samples. The test essentially compares *observed frequencies* of a dependent variable with computed *expected frequencies* and outputs a Chi-square value(Burnham, 2011, p.Chapter 22). As with most other significance tests the value is compared to a *critical* value as identified in a statistical table, with respect to the number of degrees of freedom and alpha value of choice. The alpha value is typically at 0.05 in most research and refers to the researcher's certainty of rejecting the null hypothesis and accepting the alternative hypothesis. Put alternatively, the researcher accepts that there is a 5% chance the results could be due random chance.

12.3.2 McNemar

To test for significance within each group the *McNemar Test* was found to be suitable as it compares *paired samples*, e.g. repeated measures and thus before/after type tests. It is similar to the *Chi-square* but assumes the samples are paired(Newsom, 2009). For these reasons we may use the entirety of the original test data as each person would watch two clips and thus followed a repeated measures design.

12.3.3 ANOVA for Mixed Designs

In order to test for significance for the obtained mean values of how participants liked the clips a Factorial ANOVA is put to use. Furthermore a variation of this exists, namely the *Factorial ANOVA for Mixed Designs* which computes significances for factorial designs involving repeated measures(Newsom, 2009). This in turn suggests that it is possible to use all the original samples, compared to only using half.

12.4 Overall Believability

Table 11 shows the relative proportions of *independent* participants who found a particular clip *believable* thus 1/29 refers to 1 out of 29 participant who are rated as 'believable' with respect to the single clip they watched. In the case of *No Foreground* the believability is much higher whereas *Passive Foreground* and *Active Foreground* present smaller differences in believability. Overall, *Passive Foreground* has the highest believability of the 3 levels of foreground. Whether the results are significant will be determined later.

IV 1: Foreground elements →

IV 2: Background elements ↕

	No foreground (B1)	*Passive Foreground (B2)*	*Active Foreground (B3)*
Low Background(A1)	1/29 = 3.45%	24/26 = 92.31%	24/27 = 88.89%
High Background (A2)	15/19 = 78.95%	18/22 = 81.82%	16/21 = 76.19%
	16/48 = 33.33%	42/48 = 87.5%	40/48 = 83.33%

Table 11 The relative proportions of independent participants who found a particular clip believable. 1/29 refers to 1 out of 29 participants who are rated as 'believable' with respect to the single clip they watched

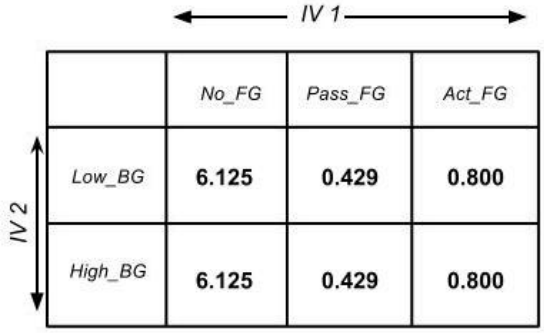
12.5 Factorial Relationship between Fore- and Background Elements

One of the main interests of the thesis was to investigate whether a relationship between the two independent variables existed. In order to compute if such a relationship exists, then the *Chi-square* is put to use. An alpha value of 5% is chosen, the null hypothesis is defined as no relationship between the two independent variables and the alternative hypothesis is defined as there is a relationship. Statistical software such as *XLSTAT*(Addinsoft, 2012)assumes a contingency table when computing the *Chi-square* which is very similar to Table 12. The contingency table essentially requires the number of positive observations of the dependent variable which in this case refers to the number of believability observations. A simplified table is shown in

		IV 1		
		No_FG	Pass_FG	Act_FG
IV 2	Low_BG	1	24	24
	High_BG	15	18	16

Table 12 Contingency table for the observed frequencies of clips that were believable for use in the Chi-square test

The outcome of the *Chi-square* test is a *p* value < 0.0006 which is less than the alpha value of 0.05 and we thus reject the null hypothesis and accept the alternative hypothesis. There is thus sufficient evidence that we may say there is a significant relationship between the Fore- and Background elements with the samples at hand. With a relationship established we proceed onto determining which variables contribute the most to the significant relationship. One such *Post-hoc* test is done by describing which cells in the table contribute the most to the overall *Chi value*(DeVries, 2007). Table 13 shows the individual *Chi-values* for each combination of the 2 independent variables.



	No_FG	Pass_FG	Act_FG
Low_BG	6.125	0.429	0.800
High_BG	6.125	0.429	0.800

Table 13 Chi value contributions per individual cell to determine which cell contribute the most to the overall significant relationship

Evidently and also to be expected the main contribution is found in the *No Foreground* group due to the large difference in observed believability. The second largest contribution is found in the *Active Foreground.*
We thus conclude there is a significant relationship between Fore- and Background elements in a set extension with the samples at hand.

12.6 Believability within each group

In order to test whether the observed differences between Low and High backgrounds are significant the *McNemar* test is put to use. This particular test assumes that samples are paired meaning that in our case a participant is exposed to 2 clips in one session. Although similar to the *Chi-square* the table that it uses internally to compute whether the difference is significant is slightly different. The table may only be 2x2 and each cell refers to the number of participants that found a combination of the two clips either believable or unbelievable. The test uses both the frequency of positive outcomes (believable) but also the number of negative outcomes (unbelievable).

For each of the three levels of Foreground we investigate whether there is a significant difference in the believability of each group in terms of Background elements. Furthermore we investigate whether the order in which they watch the two levels of Background matters. Recall that for this particular test all samples are used as the McNemar assumes repeated measures. The procedure will only be described for the *No Foreground* group as the method for the *Passive* and *Active* groups is similar.

12.6.1 No Foreground - Participants who watch Low Background First

Table 14 shows a *paired* sample table of whether participants found the two clips they watched in a session believable or not. The '26' indicates that 26 participants found the first clip (Low background) *Unbelievable* and their second clip (High Background) *Believable*. Similarly the '2' represents two participants that found both clips they watched for *Unbelievable*.

No Foreground - Low BG first		Low_BG	
		Believable	Unbelievable
High_BG	Believable	1	26
	Unbelievable	0	2

Table 14 McNemar table that shows whether the participants found the first and second clips believable. This particular example had the participants watch the Low background first

The calculated *McNemar p* value < 0.0001 is less than alpha=0.05 and the null hypothesis is thus rejected. The evidence suggests there is a significant difference between Low and High background in terms of believability.

12.6.2 No Foreground - Participants who watch High Background First

No Foreground - High BG first		High_BG	
		Believable	Unbelievable
Low_BG	Believable	3	1
	Unbelievable	12	3

Table 15 McNemar table that shows whether the participants found the first and second clips believable. This particular example had the participants watch the High background first

Similarly a test is carried out that determines whether the difference is significant when participants alternatively watch the High background first. The calculated *McNemar p* value < 0.0034 is less than the alpha value at 0.05 and the null hypothesis is thus rejected. The evidence suggests there is also a significant difference between Low and High background in terms of believability.

12.6.3 No Foreground - Believability Summary

Thus we conclude as suggested by the evidence that there is a significant difference between watching Low and High background versions of the No Foreground group regardless of which order the clips were presented to the participants.

12.6.4 Passive and Active foregrounds

Table 16 shows the same method applied to the Passive and Active foreground in regards to Low and High backgrounds.

		P-value	Significant
Pass FG	Low BG First	0,219	No
	High BG First	0,688	No
Act FG	Low BG First	0,625	No
	High BG First	0,688	No

Table 16 The p-values as computed for the Low and High backgrounds for Passive and Active foregrounds. None of the differences are significant

12.6.5 Passive and Active Foreground - Believability Summary

The evidence suggests there was no significant difference in believability for Low and High Backgrounds for the Active Foreground group regardless of which order the clips were presented to the participants.

12.7 Passive and Active Foregrounds compared

A one-way *Chi-square* test was used to examine whether the difference in believability for the Passive and Active Foregrounds was significant. It should be noted that the only two clips of interest in this context are those of High Backgrounds, as it is hypothesized that an Active Foreground should afford less background content, whilst maintaining believability.

The one-way Chi-square was computed to a chi value of 0.471 which is less than the critical value of 3.841 (alpha at 0.05 and 1 degree of freedom) and the observed difference in believability is thus insignificant.

12.8 How participants rated the clips

Question 4 dealt with how a particular participant rated a clip in terms of sheer likability. In comparison with the other variables these results were not binary in nature and a different way of significance testing was thus used. The *Factorial ANOVA for Mixed Designs* was used to test whether there was a significant relationship between Fore- and Background elements. Using 288 samples of scores (1-5) spread across the 6 clips, acknowledging that a participant sees two clips in random order, the following results were obtained.

12.8.1 In general

Table 17 shows the mean scores with respective standard deviation for each clip. As it may be witnessed the Active Foreground-High Background had the highest mean rating. Furthermore the Active Foreground has the highest overall rating and the No Foreground has the overall lowest. Figure 36 the overall ratings for the 3 main groups. The actual ratings are used in the per-case evaluation for Believability and are as such not used to evaluate upon the hypotheses. As we are only interested in an overall view of the means significance tests were not performed.

IV 1: Foreground elements →

IV 2: Background elements ↕

	No foreground (B1)	*Passive Foreground (B2)*	*Active Foreground (B3)*
Low Background(A1)	2.81 *0.87*	3.65 *0.84*	3.58 *0.79*
High Background (A2)	3.69 *0.8*	3.69 *0.8*	3.81 *0.79*
Overall Mean **Overall STD**	3.25 *0.94*	3.67 *0.82*	3.7 *0.8*

Table 17 Mean values and standard deviations for each clip

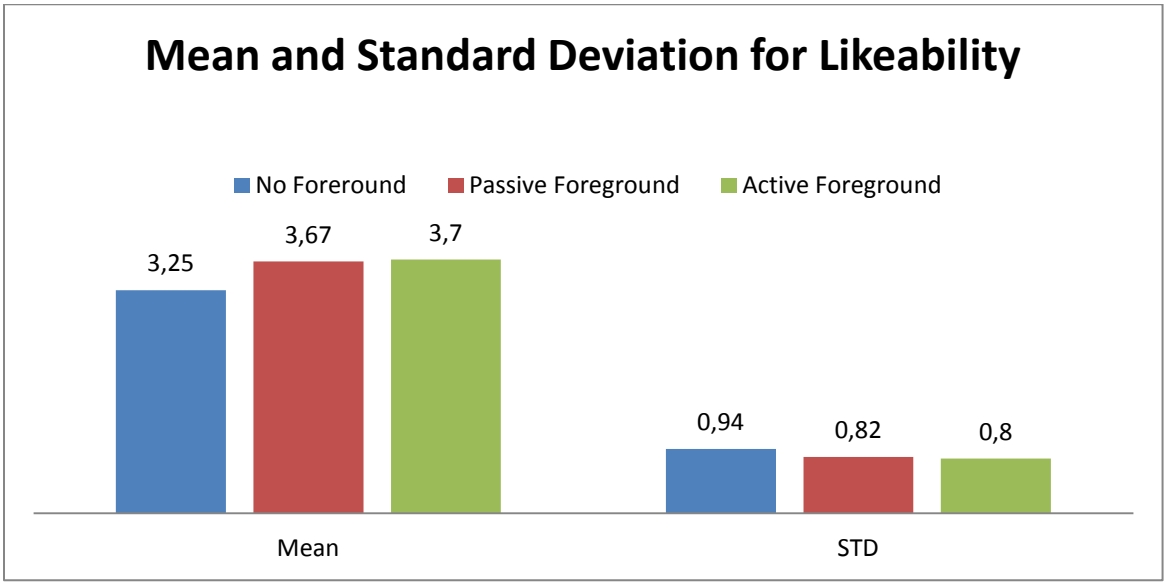


Figure 36 Mean value and standard deviation for each of the 3 main groups

12.8.2 The Factorial Relationship between Fore- and Background Elements

As mentioned the *Factorial ANOVA for Mixed Designs* was carried out to investigate whether a significant relationship existed between the two independent variables of interest. Using XLSTAT and an alpha=0.05 a significant relationship between the two variables was identified. Furthermore it was identified that both variables had an impact on the believability and that the *High Background* had the largest overall impact. This in turn corresponds to the significance relationship as identified through the *Chi-square* believability tests.

12.9 Element Identifications

As part of the experiment it was of interest to see patterns of what participants noticed during the various clips. Specifically it was of interest to see whether they noticed the *Dreadnought, Middleground Flyby's* and the *far-distance ships*. These observations are used per-case in evaluating whether a clip is found to be believable in relation to initial fixation and semantic inconsistency and incongruence, thus they are not used directly to evaluate upon the thesis. The following will thus only serve as tendencies and significance tests are not deemed necessary to be performed. Figure 37 through Figure 39 show the percentages of the three identification categories as obtained through the use of all the sample data (counter-balancing is used to reduce the bias). It should be noted that Low Backgrounds do not contain *Middleground Flyby's* or *Far-distance ships*.

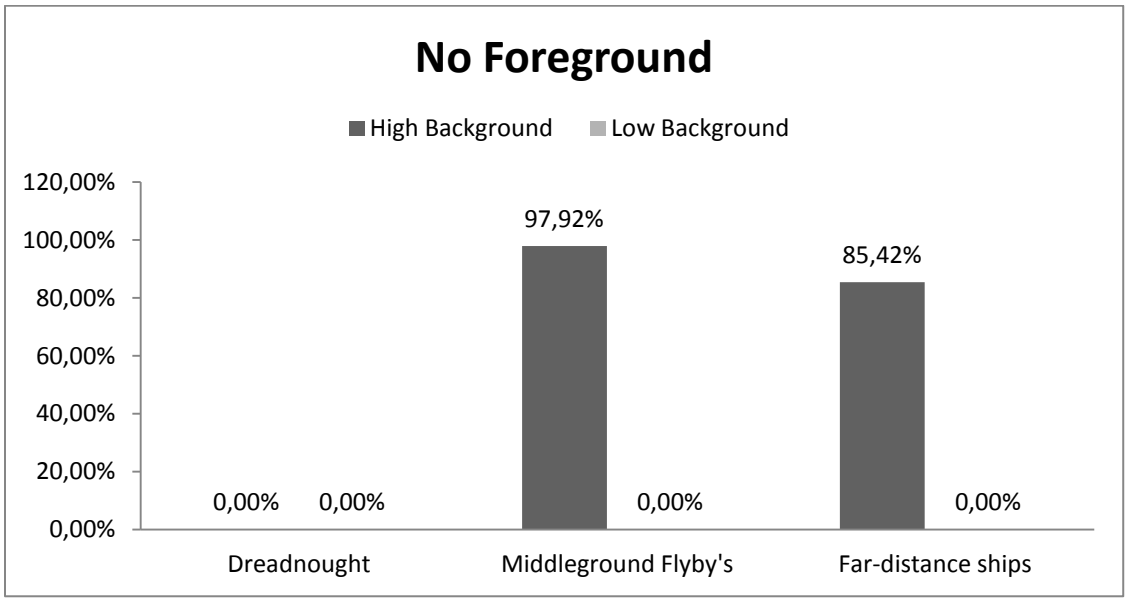


Figure 37 Identification of Background elements with no foreground

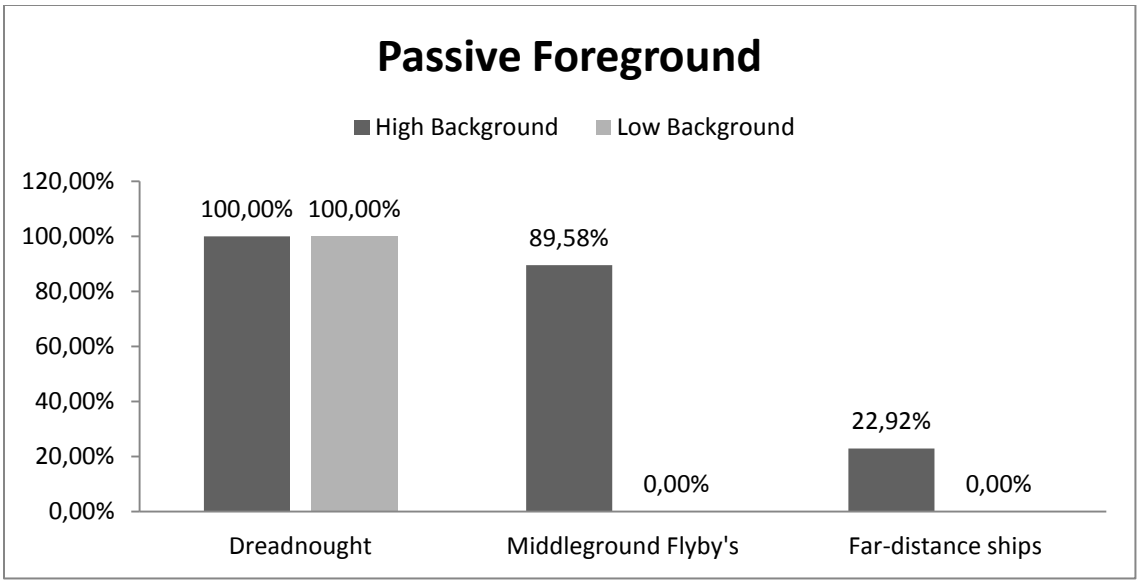


Figure 38 Identification of Background elements with passive foreground

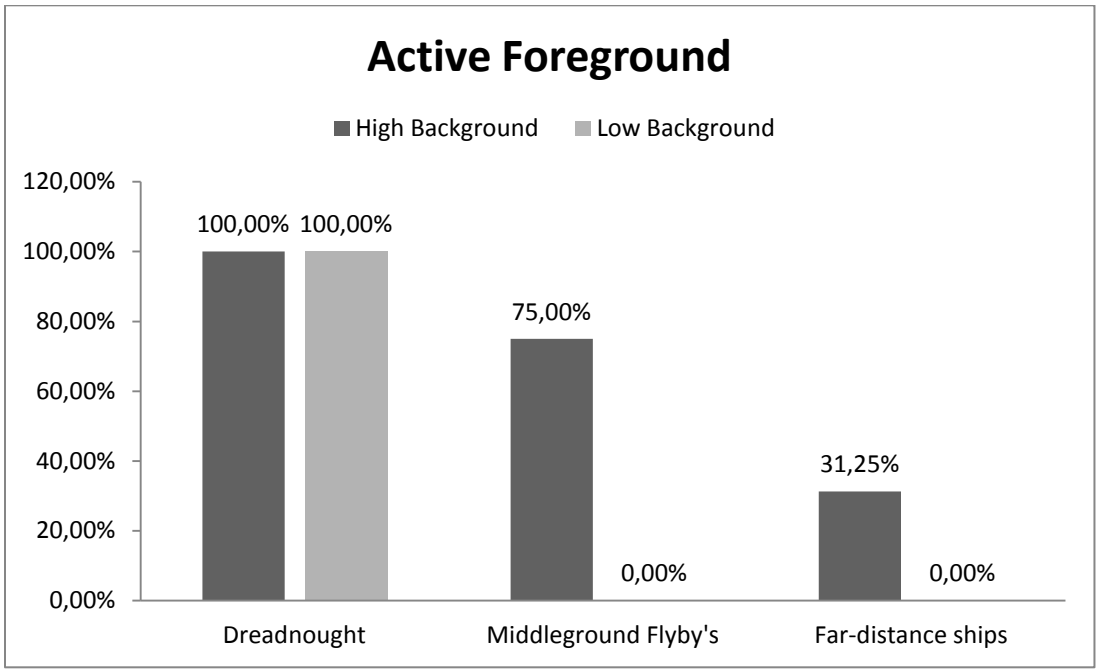


Figure 39 Identification of Background elements with Active foreground

The three graphs show tendencies of how participants notice *Middleground Flyby's* and *Far-distance ships,* to a much greater degree than under the No Foreground compared to Passive and Active.

12.10 Science Fiction Identifications

Furthermore the experiment examined whether participants were able to identify the genre of a particular clip. As motivated by the Believability Model a clip may only be regarded as believable if the participant is able to identify it correctly as science fiction as the model identifies believability only within a certain context. Similarly to the element identifications this was identified on a per-case basis when examining for believability and the following is thus only meant to show tendencies and not tested for significance. Figure 40 through Figure 42 show the percentages of Science Fiction identifications as obtained through the use of all the sample data (counter-balancing is used to reduce the bias).

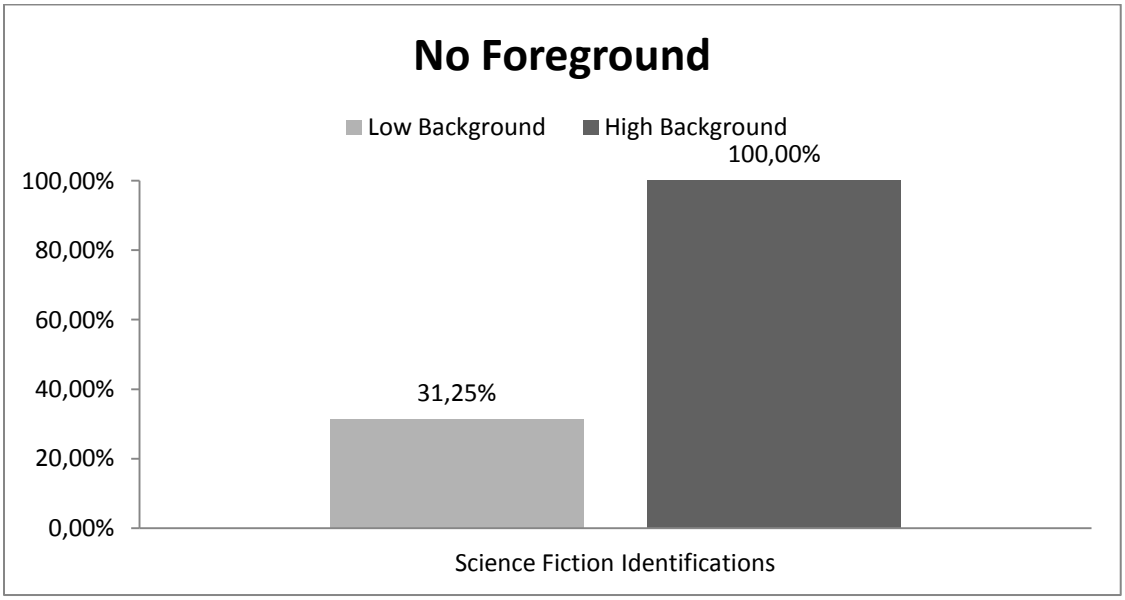


Figure 40 Science Fiction identification with No Foreground

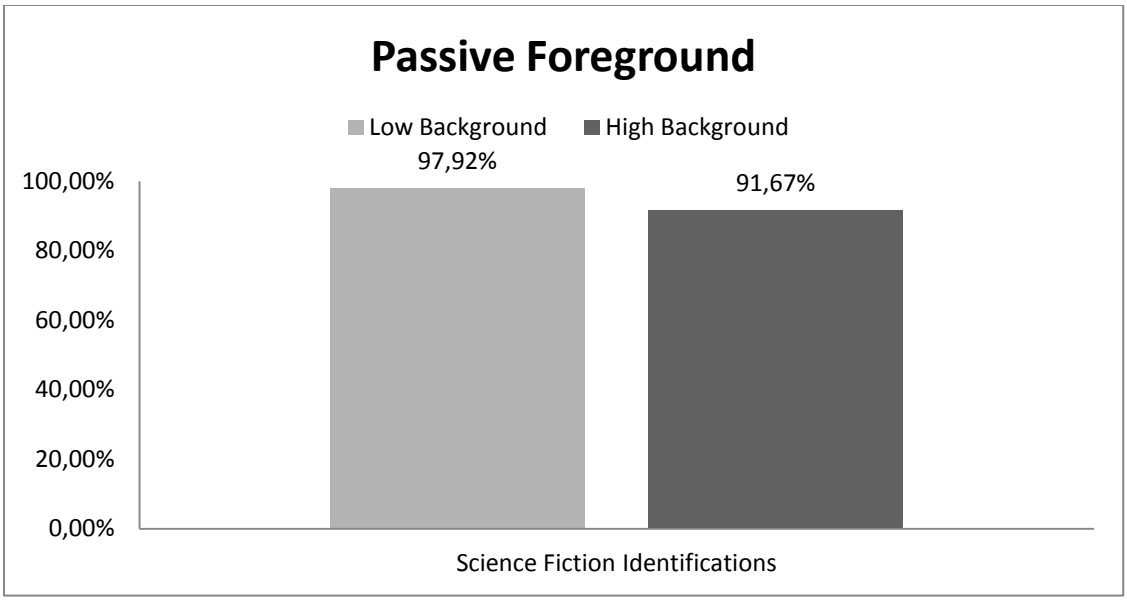


Figure 41 Science Fiction identification with Passive Foreground

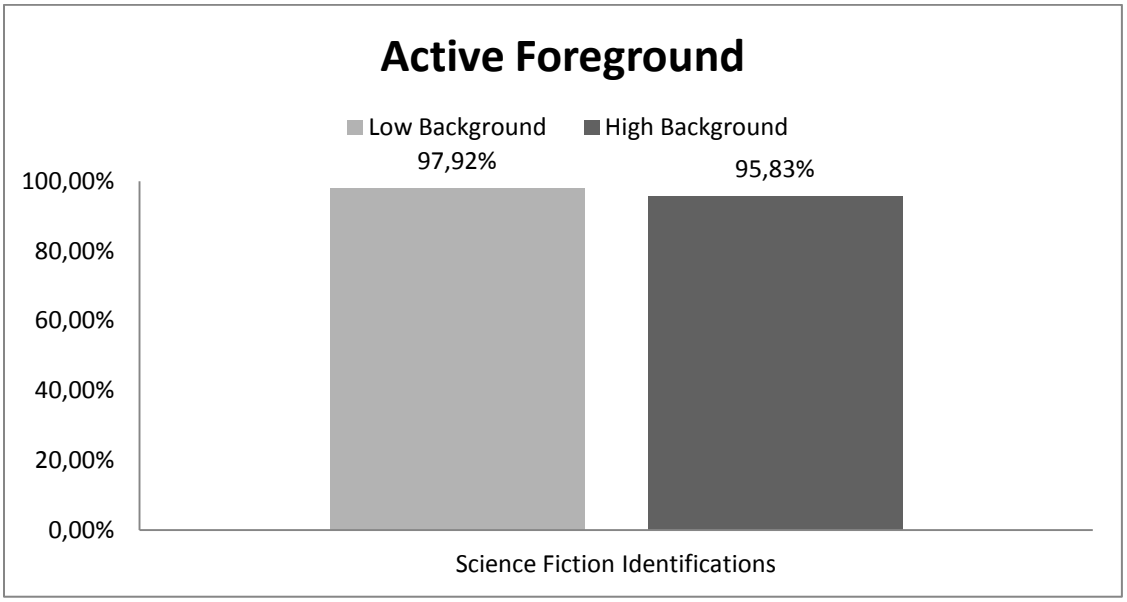


Figure 42 Science Fiction identification with Active Foreground

The graphs show tendencies of participants having trouble with identifying the genre when there is no foreground.

12.11 Results Summary

To summarise the following was covered:

- Chi square, McNemar and Factorial ANOVA for Mixed Design were used for significance tests

- The original pool of samples had to be halved to accommodate the *Chi-square's* assumption of independent samples for testing whether the relationship across the two independent variables was significant
- The Passive foreground presented the highest overall Believability
- Active Foreground-High Background was found to have the highest overall 'Likeability' rating

Overall believability relationship between foreground and background elements across the 3 groups

- There is a significant relationship between the foreground and background using 3 groups of data(3 levels of foreground) in terms of believability
- This is mainly attributed to the difference in the No Foreground group for Low and High Backgrounds
- Thus, there is a relationship between Foreground and Background elements in a set extension. This is in particular due to the massive difference found in No Foreground version.

Believability difference within the 3 main groups (between background levels)

- The No Foreground group presents a significant difference in terms of believability. When presented with a High Background people find it much more believable than compared with a Low Background. This confirms our hypothesis.
- The Act and Pass groups present no significant differences in terms of believability within the groups. This also confirms our hypothesis.
- These findings are true regardless of whether people watch Low or High versions first.

Believability between Active and Passive with High background clips

- The difference in believability between the two clips is insignificant

13 Discussion

The following chapter will be an evaluation of the results and how different decisions made throughout the project may have influenced these. To provide an overview the following topics will be discussed:

- Results Discussion
- Defining and Testing Believability
- The Videos
- Influential Factors on the Videos
- Testing Method

13.1 Results Discussion

The hypothesis dealt with two major points of interest, in regards to believability as a result of the relation between fore- and background elements in a scene. The hypotheses were as follows:

> ***"If a foreground element is present a test subjects focus lies on that, especially if the foreground presents an action. Therefore the need for a complex background is lessened to achieve believability.... Subsequently, if the foreground is not present the focus will lie on the background therefore a more complex background is needed."***

The test set out to examine this, and according to the hypothesis this meant that there should be a significant difference in believability between the groups with a foreground element and the group without. Furthermore for the case of background difference, there should be no difference in the active and passive foreground groups, while the one with no foreground should see a difference in believability between low and high background. The results as shown in section support these hypotheses as there were the presumed significant differences as well as the lack of difference where presumed. This then gives the indication that the presence of a foreground element indeed reduces the need of elements in a set extension in order to achieve believability.

These results should however be taken with some initial reservations as will be evident after the following discussion.

13.2 Defining and Testing Believability

As believability is one of the key elements in this project, the strong and weak points of the definition and model will be evaluated. The main challenge in regards to the definition of believability is the highly subjective and rather broad nature of it. The subjectivity was one of the reasons why it was questionable if it

was possible to directly address if the test participants felt the videos were believable or not. Because how one defines what believability, and on which basis one judges that, may vary a great deal from person to person. Naturally that was also the reason for formulating the definition and constructing the model, to get a uniform way of evaluating the believability. The question that has to be raised is whether or not the approach became that uniform in the end. Some of the participants responses were fairly easy to judge (in accordance with the procedure described in section 8.3) whether they found the video believable or not, as an answer such as;

"It is really well made - almost can't tell it isn't real" -Facebook Participant (Active foreground, Low background)

makes it pretty straightforward to interpret, as the person almost literally states that the clip was believable. Not every case was that simple to interpret however and it often became a matter of a very subjective interpretation. While subjective interpretations cannot be completely avoided when dealing with qualitative questions, they may be minimised however as exemplified by follow up questions that can serve as cross checks.

In this particular test such cross checks could have, in retrospective, been valuable to serve as general validation, and evaluate the more ambiguous case such as:

"Nothing much happened. I liked the effects, but there wasn't much going on in the scene." -Facebook Participant (Passive foreground, Low background)

Every question was quantified to some extent during the result processing, and it thus worth considering whether each question should been asked as a quantitative question from the beginning. As argued in section 8.1.3 quantitative questions would be tricky to use in the case of a repeated measure experiment. Since asking specific question in the likes of "Did you notice this particular object in the video?" could very well make the participant be on the lookout when watching any subsequent video. So if quantitative questions were to be applied it would require a very elaborate system of asking around a subject without mentioning it. This is especially present, considering that the proposed believability model deals with things such as initial fixation which would be convoluted to deal with as a quantitative question.

While all the elements that compose the believability model are based on previous research from various sources and from various disciplines, in order to improve on the model it would be interesting to test on each element separately within this context. This could help find weak points, missing links and overall strengthen the model. As it stands now it is hard to pinpoint a single element in the model that is severely flawed, and overall it does provide a solid theoretical basis for further research in the field.

The way that the model and questions have been applied, means that a substantial part of the interpretation on whether or not the participant found it believable or not, relied on question four. "Please rate how well

you liked the clip". The assumption was that if the believability in a video was low then the rating of how the participants liked the scene would be low as well, since believability should be a vital part of how well liked a VFX shot is (with some special exceptions such as if people like a film for being bad). The implications is, that how well a clip is liked is influenced by far more factors than just the believability which became apparent with the responses to the video with no foreground element, and low background level. Initially it seemed that the responses supported the hypothesis that with no foreground the fixation would lay on the background, and with the lack of background elements the scene (and set extension) would fail as one within the science fiction context. Most did not identify the scene as science fiction (thus failing as a believable science fiction scene) and the 'like' score was low as well (a mean of 2.7).

While the video was not found believable as a science fiction scene, in theory the scene would actually have a higher chance of appearing believable as there is far less VFX elements that potentially can break the believability. The qualitative responses indicate that the low 'like' score were more due to the fact that were the video clip was found boring:

"Nothing happened. The music was very dramatic but there was no visual effect moving in the picture." -Facebook Participant (No foreground, Low background)

This only reinforces the need for additional questions to get a better evaluation of believability.

Context has been quite a keyword throughout the thesis seeing how it has been stressed as a point that believability should be evaluated within a given context (in this case a science fiction VFX shot). Context in its broad definition also plays a significant role on how believability can and should be tested in regards to a VFX shot. Because the perhaps biggest challenge in this believability model is to be more generalised, is that a VFX shot varies on many levels from shot to shot. One factor is the length of a shot, is it possible to measure believability on a shot that last 5 or 10 seconds? And the fact that such a shot normally would be a part of a full film also raises some issues. Some participants directly mentioned that they did not understand purpose of the videos out of context;

"I still do not understand the purpose of the clip out of its context." - Facebook Participant (Active foreground, High background)

and a full film naturally only helps define the genre (the context according to the model) which has been highlighted as an important element.

13.3 The Videos

The six videos provided some interesting findings that went slightly against the hypothesis, but gives some interesting considerations. One of the sub-hypotheses and the reason why the passive foreground was

included was to see if the foreground had to be active to divert and maintain the fixation of the viewer. In essence it was expected to see that in the case with a passive foreground people would notice slightly more of the background, as the active would afford more of a top down perception. While the results for this were not tested for a significant difference the results did however show an interesting tendency. Apparently the participants noticed the background more often in the active case than the passive, opposite of expected. Considering this and looking at the videos, the explanation might be that the active foreground consists of a walkcycle added to the dreadnought, which means that it starts mainly out of frame and walks slowly into the frame. This could eventually give more time for the viewer to fixate on the background initially, even though it is only a brief moment. Furthermore the passive version is not truly passive as the dreadnought still goes through the shooting animation. Therefore more extreme cases e.g. having the dreadnought make bigger and more spectacular movements, would have to be tested in order to see if the addition of an action to the foreground object, makes a difference.

The thesis dealt with semantic inconsistency and incongruence to some extent and it is thus interesting to consider if these occurred to some extent in the videos. Going through the results, it was mentioned by a couple of participants that the aim of the dreadnought was off target compared to where it shot. While not directly listed this could be considered something very close to Biederman's relational violations (see section 7.1.1) and thus be considered as a semantic inconsistency. Generally those who stated this mentioned it in relation to why they rated the video slightly lower than they otherwise would have done:

"I would have given it a 5, but the turret looked like it was aiming too high" - Facebook Participant (Active foreground, Low background)

While this does not show to what exact extent such a semantic inconsistent element affects the overall believability, it is still an indication that semantic inconsistency (and incongruence for that matter) may indeed be an influence on believability.

13.4 Influential Factors on the Videos

In the summary of the pre-analysis (section 5.5) a series of factors that influences the VFX production were identified. Factors such as camera movement, lighting, the use of depth of field etc. are all things that have a great deal of influence on the shot, thus most likely also on the believability. As the focus was on the VFX the idea was to use the factors to optimise the VFX workflow, but also had to be done within the limits of our practical possibilities.

Two of these factors showed up in the results, one of the main factors that turned out to influence the videos was sound. This is not a huge surprise as sound is a main component of films and thereby also VFX shots,

and sound may even add a great deal to the believability of a VFX shot. From the start of the project it was decided to focus only on the visual aspect as it was felt that sound would require just as much research and work to get to the same level. The problem as discussed previously was that if sound was totally omitted all attention would on the visuals and small mistakes more likely to be noticed. The choice then became to use a piece of classical music (a section of Symphony No. 9 by Beethoven) which mostly by chance turned out to fit the videos, as the shots from the dreadnought aligned perfectly with the time of the music. But none the less the music added a factor that could influence the outcome of the participants' answers. While not all mentioned the music, there were some divided responses that seemed to influence how well liked the video in question were, thus possible also affecting the believability.

"I think it was pretty nice produced, and the music was great." - Facebook Participant (Active foreground High background)

"I didn't rate it higher, because the camera is too static, the music is too dramatic compared to the scene..." - Facebook Participant (Active foreground High background)

But besides influencing the test to some degree, it also provides some interesting considerations concerning the believability model. As the believability model is structured in this project, it is based mainly on visual perception theory and sound is just as likely as big a subject that could be implemented in future work on the model, and it thus provides some interesting aspects.

13.5 Testing Method

All of the investigated theory contributed to creating various iterations of two independent variables of interest. With two such variables identified the thesis went onto examining whether those variables were related in any fashion and to what extent. A significant relationship was in fact found to exist between the Fore- and Background of a given shot but as the following addresses the result should only be regarded as tendencies.

The original experiment methodology turned out to be unsuitable for the main relationship test as the dependent variable (believability) comprised binary frequencies and most significance tests for such data assume independent samples. We were thus forced to half the samples in order to obtain 6 independent groups for each of the IV combinations. As the original sample was relatively large it was not regarded critical but with all tests a larger sample will yield more representative results and this could thus invalidate the results to some extent. As we ended up combining the samples from the Campus- and Peer Groups the need for a Mixed design in the first place might not have been necessary after all. The two sample groups did differ slightly but as we had no target group per se they were regarded as belonging to the same overall group and the combination was thus deemed appropriate. We did however not have any means to verify that

participants at home did not watch clips multiple times as compared to the ones on campus but this was regarded less important compared to obtaining a large sample.

Conclusively it should be mentioned that the device with which participants watch a given clip may have an influence on their answers and this would be substance for further research. As the thesis had multiple other interests this was however not at focus. Furthermore per each case we attempted to quantify otherwise qualitative answers and as this is highly subjective in nature this could invalidate further the results to some extent.

Conclusively the test methodology did however create a means of examining highly subjective and intangible topics and within the given context the methodology with respect to the model must be regarded as successful.

14 Conclusion

The following chapter will conclude on the process and findings to determine to what extent the problem statement was answered.

The project was motivated by the increasing use of VFX in both film and TV, and in particular the increased use of digital set extensions. This increasing use leads to a similar increasing workload on the artists, which lead to the speculation of what extent this workload could be reduced without hurting the believability of the set extension. The preliminary analysis was based around analysing VFX and set extensions to find factors which could reduce the workload. This was done through interviews with the industry as well as researching contemporary VFX. Furthermore defining believability became a focus of the preliminary analysis. This gave room for the idea that there was a relationship between fore - and background in regards to that the presence of a foreground element might lessen the need for a highly detailed background (thus reducing the workload on the set extension).

This materialised in the development of the first iteration of a believability model which was meant as a basis for the testing of believability. The model was based around perception theory and primarily on visual perception as the interest lay within VFX. Therefore it subsequently became apparent that a whole parallel model on auditory believability could be made and intertwined with the proposed model, to make a make more cohesive unit.

Regardless of the shortcomings of the model, it still provided a basis for testing believability and a starting point for further research into the field.

The problem statement was as follows:

To what extent will believability be affected when altering the elements in the foreground and background in a set extension?

In order to investigate and answer this statement the following hypotheses were constructed:

- If a foreground element is present a test subjects focus lies on that, especially if the foreground presents an action. Therefore the need for a complex background is lessened to achieve believability.
- Subsequently, if the foreground is not present the focus will lie on the background therefore a more complex background is needed.
- If the subjects fail to realise that the setting is sci-fi the execution of the set extension may be regarded as unsuccessful.

The test results supported these hypotheses as the participants generally found the videos with a foreground element more believable than the videos without a foreground.

Furthermore it was found that the case with no foreground presented results where participants had more focus on the background and a more complex background was thus needed in order for participants to identify the video as science fiction.

In regards to the actual relationship between Fore- and Background elements we conducted a series of significance tests including the Chi-square test and the Factorial ANOVA for Mixed Designs. The prior made use of per-case frequencies of believability whereas the latter used the more concrete likeability scores. Both indicated that there was a significant relationship between the two variables which confirmed our estimations.

In part we also hypothesized that there would not be any significant difference within the Active and Passive groups with respect to Low and High backgrounds and this was also confirmed by the tests. The same results were also obtained when altering the order at which the participants saw the two clips. Alternatively it was found that there was no significant difference between Passive and Active with High Backgrounds which went against our hypothesis of an audience having higher believability when presented with an Active foreground.

The Discussion section presented a selection of things that should be considered when looking at the results and they should therefore not be regarded as definitive but rather as tendencies. It should however be

emphasized that the topics that were dealt with for the thesis were of a very subjective nature and through the synthesising of a model for believability the results must thus be regarded successful to some extent.

In the end this project showed indications of the relation between fore- and background elements in regards to believability. Through the thesis we undertook a long process of analysing the topic of believability extensively. The end result was a theoretical model and relevant test methodology based on various theories and methodologies. The model is by no means flawless but given its intangible and subjective nature we hope that the knowledge obtained may be used in further research within the field.

Such further research should consider the following:

Further research into the connections between the different elements would be needed.

- Expanding the model to include elements such as sound.
- An even more quantifiable way of testing believability.
- Follow up questions to cross check the participants' responses.
- A refined test methodology.

15 References

Addinsoft, 2012. *Xlstat*. [Online] Available at: http://www.xlstat.com/en/.

Avatar. 2009. [Film] Directed by James Cameron. US: 20th Century Fox.

Babcock, J.S., Lipps, M. & Pelz, J.B., 2002. How people look at pictures before, during, and after scene capture: Buswell revisited. Rochester, 2002. Chester F. Carlson Center for Imaging Science.

Bickerton, J., 2010. Budget Effects. *Televisual*, pp.37-44.

Bickerton, J., 2010. Television visual effects on the rise, budgets not. *Televisual*, (01), p.p. 21.

Biederman, I., Mezzanotte, R.J. & Rabinowitz, J.C., 1979. Scene Perception" Detecting and Judging Objects Undergoing Relational Violations. *Cognitive Psychology*, 14(2), pp.143-77.

Blade Runner. 1982. [Film] Directed by Ridley Scott. USA: Warner Bros. Pictures.

Boardwalk Empire. 2010. [Film] US: Leverage, Closest to the Hole Productions, Sikelia Productions, Cold Front Productions.

Borkowski, J., 2010. *Chi-Squared Tess for Homogeneity*. [Online] Department of Mathematical Sciences - Montana State University Available at: http://www.math.montana.edu/~jobo/st431/ho3.pdf.

Burnham, B.R., 2011. *Fundamental Statistics for the Behavioral Sciences*. [Online] Available at: http://humanattention.org/Textbook/FundamentalsofStatisticsTextbook.html.

Buswell, G.T., 1935. *How people look at pictures: a study of the psychology and perception in art*. Chicago: The University of Chicago Press.

Cater, K., Alan, C. & Ward, G., 2003. Detail to Attention: Exploiting Visual Tasks for Selective Rendering., 2003. EGRW '03 Proceedings of the 14th Eurographics workshop on Rendering.

Children of Men. 2006. [Film] Directed by Alfonso Cuarón. USA, UK: Universal Pictures.

Coren, S., Ward, L.M. & Enns, J.T., 2004. *Sensation and Perception*. 6th ed. Hoboken, NJ: John Wiley & Sons.

Cozby, P.C., 2008. *Methods in Behavioral Research*. 10th ed. McGraw-Hill.

Davenport, J.L. & Potter, M.C., 2004. Scene Consistency in Object and Background Perception. *Psychological Science*, 1 August. pp.559-64.

DeVries, J., 2007. *About Chi Squares*. [Online] University of Guelph Available at: https://dspace.lib.uoguelph.ca/xmlui/bitstream/handle/10214/1863/A_About%20Chi%20Squares.pdf?sequence=2.

District 9. 2009. [Film] Directed by Neill Blomkamp. South Africa, USA: TriStar Pictures.

Ezra, E., 2000. *George Méliès (French Film Directors)*. 1st ed. Manchester: Manchester University Press.

Failes, I., 2011. *Building a Boardwalk Empire*. [Online] Available at: http://www.fxguide.com/featured/building-a-boardwalk-empire/ [Accessed March 2012].

Failes, I., 2011. *Pan Am: retro green screen world*. [Online] Available at: http://www.fxguide.com/featured/pan-am-retro-green-screen-world/ [Accessed March 2012].

Ferwerda, J.A., 2003. Three Varieties of Realism in Computer Graphics. *Human Vision and Electronic Imaging*, 17 June. pp.290-97.

Final Fantasy: The Spirits Within. 2001. [Film] Directed by Hironobu Sakaguchi & Moto Sakakibara. USA, Japan: Columbia Pictures.

Finance, C. & Zwerman, S., 2010. *The Visual Effects Producer: Understanding the Art and Business of VFX*. 1st ed. Oxford: Elservier.

Friedman, A., 1979. Framing Pictures: The Role of Knowledge in Automatized Encoding and Memory for Gist. *Journal of Experimental Psychology: General*, 108(3), pp.316-55.

Game of Thrones (Tv Series). 2011. [Film] US: HBO, Created By, Management 360.

Green Lantern. 2011. [Film] Directed by Martin Campbell. US: Warner Bros.

Grodal, T., 2006. The PECMA Flow: A General Model. *Film Studies*, 8, pp.1-11.

Hasic, J. & Chalmers, A., 2007. *Visual Attention for Significantly Influencing the Perception of Virtual Environments*. Plzen: Proceedings of the Winter School of Computer Graphics. http://wscg.zcu.cz/WSCG2007/wscg2007.htm.

Henderson, J.M. & Hollingworth, A., 1999. High-level scene perception. *Annual Review of Psychology*, 50(1), pp.243-71.

Henderson, J.M., Weeks, J..P.A. & Hollingworth, A., 1999. The Effects of Semantic Consistency on Eye Movements During Complex Scene Viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 25(1), pp.210-28.

Hoffman, J.E., 1980. Interaction Between Global and Local Levels of a Form. *Journal of Experimental Psychology: Human Perception and Performance*, 6(2), pp.222-34.

Hooks, E., 2011. *The Willing Suspension of Disbelief*. [Online] Available at: http://www.awn.com/blogs/ed-hooks-acting-animators/willing-suspension-disbelief [Accessed 2012].

Hugo. 2011. [Film] Directed by Martin Scorsese. USA: Paramount Pictures.

Industrial Light & Magic, 2012. *The Visual Effects of "Transformers: Dark of the Moon" Part 2*. [Online] Available at: http://www.youtube.com/watch?v=L7r5mz9aTlA [Accessed March 2012].

Johnston, O. & Thomas, F., 1981. *The Illusion of Life - Disney Animation*. New York: Disney Editions.

Jurassic Park. 1993. [Film] Directed by Steven Spielberg. US: Amblin Entertainment.

Khouri, A., 2011. *'Green Lantern' Increases VFX Budget as Hollywood Trends Towards Deadline Disasters*. [Online] Available at: http://www.comicsalliance.com/2011/04/27/green-lantern-increases-vfx-budget-as-hollywood-trends-towards-deadline-disasters/ [Accessed March 2012].

Lancaster, K., 2010. *DSLR Cinema - Crafting the Film Look with Video*. 1st ed. Focal Press.

Le Voyage dans la lune. 1902. [Film] Directed by George Mèliés. France: Gaston Méliès Films.

Loftus, G.R. & Mackworth, N.H., 1978. Cognitive Determinants of Fixation Location. *Journal of Experimental Psychology: Human Perception and Performance*, 4(4), pp.565-72.

Mattingly, D.B., 2011. *The Digital Matte Painting Handbook*. 1st ed. Sybex.

Michael, R.S., 2002. *Types of Variables*. [Online] Indiana University Available at: http://www.indiana.edu/~educy520/sec5982/week_2/variable_types.pdf.

Mitchell, R.L., 2004. *WETA: Rendering Lord of the Rings*. [Online] Available at: http://www.digitalartsonline.co.uk/features/?featureid=1162.

Monnet, L., 2004. A-Life and the Uncanny in "Final Fantasy: The Spirits Within". *Science Fiction Studies*, 31(1), pp.97-121.

Mori, M., 1970. The Uncanny Valley. *Energy*, 7(4), pp.33-35.

Møller, M. & Pellengahr, C.E., 2011. *Less Definition*. Master Thesis. Ballerup: Aalborg University Copenhagen.

Neisser, U., 1979. The control of information pickup in selective looking. *Perception and its*, pp.201-19.

Newsom, J.T., 2009. *Chi-square for within-subjects: McNemar's test*. [Online] Portland State University Available at: http://www.upa.pdx.edu/IOA/newsom/ho_chisq%20within.pdf.

Newsom, J.T., 2009. *Factorial ANOVA for Mixed Designs*. [Online] Portland State University Available at: http://www.upa.pdx.edu/IOA/newsom/da1/ho_mixed.pdf.

O'Conner, D., n.d. *The Walk Cycle*. [Online] Available at: http://www.idleworm.com/how/anm/02w/walk1.shtml.

Okun, A.J. & Zwerman, S., 2010. *The VES Handbook of Visual Effects*. 1st ed. Waltham, Massachusetts: Elsevier.

Oxford University Press, 2012. *Oxford Dictionaries - Definition of congruence*. [Online] Available at: http://oxforddictionaries.com/definition/congruence?q=congruence [Accessed 20 April 2012].

Pan Am (TV Series). 2011. [TV] USA: ABC.

Prince, S., 1996. True Lies: Perceptual Realism, Digital Images, and Film Theory. *Film Quarterly*, pp.27-37.

Rademacher, P., Lengyel, J., Cutrell, E. & Whitted, T., 2001. *Measuring the Perception of Visual Realism in Images*. London: Springer-Verlag.

Rise of the Planet of the Apes. 2011. [Film] Directed by Rupert Wyatt. US: 20th Century Fox.

Russ, J., 1995. *To Write Like a Woman: Essays in Feminism and Science Fiction*. Indiana : Indiana University Press.

Sandal, J.S., 2010. *Fra fraktaler på Commodore 64 til huden på Avatars Na'vi*. [Online] Available at: http://www.version2.dk/artikel/fra-fraktaler-paa-commodore-64-til-huden-paa-avatars-navi-14843 [Accessed March 2012].

Schaper, E., 1978. Fiction and the Suspension of Disbelief. *The British Journal of Aesthetics*, pp.31-44.

Seymour, M., 2011. *Hugo: a study of modern inventive visual effects*. [Online] Available at: http://www.fxguide.com/featured/hugo-a-study-of-modern-inventive-visual-effects/ [Accessed March 2012].

Shuttleworth, M., 2009. *Counterbalanced measures design*. [Online] Available at: http://www.experiment-resources.com/counterbalanced-measures-design.html.

Simons, D.J. & Chabris, C.F., 1999. Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception*, 28(9), pp.1059-74.

Sontag, S., 1976. The Imagination of Disaster. In M. Rose, ed. *Science Fiction, A Collection of Critical Essays*. Prentice Hall. pp.113-27.

Spilotros, F.L. & Parzy, E., 2010. *Semantic Congruence and Perception : An Experimental Investigation*. Ballerup: Aalborg University Copenhagen.

StatSoft, Inc, n.d. *Introduction to ANOVA / MANOVA*. [Online] Available at: http://www.statsoft.com/textbook/anova-manova/.

The Twilight Saga: Breaking Dawn Part 1. 2010. [Film] Directed by Bill Condon. US: Summit Entertainment.

Thomas, A., 2004. *Anatomy of a blockbuster*. [Online] Available at: http://www.guardian.co.uk/film/2004/jun/11/3 [Accessed March 2012].

Transformers: Dark of the Moon. 2011. [Film] Directed by Michael Bay. US: Paramount Pictures.

VES, 2012. *Visual Effects Society*. [Online] Available at: http://www.visualeffectssociety.com/ [Accessed 18 Febuary 2012].

Yantis, S. & Jonides, J., 1984. Abrupt Visual Onsets and Selective Attention: Evidence From Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, 10(5), pp.601-21.

16 Figure References

Figure 2 from http://www.thegate.ca/reviews/010123/film-friday-rise-of-the-planet-of-the-apes-and-the-change-up/

Figure 3 from http://www.fxguide.com/featured/more-incredible-invisible-effects/

Figure 5 and Figure 11 from http://www.youtube.com/watch?v=L7r5mz9aTlA

Figure 7 from http://www.fxguide.com/featured/building-a-boardwalk-empire/

Figure 8 from http://www.fxguide.com/featured/pan-am-retro-green-screen-world/

Figure 9 and Figure 10 http://www.fxguide.com/featured/hugo-a-study-of-modern-inventive-visual-effects/

Figure 19 from
http://www.games-workshop.com/MEDIA_CustomProductCatalog/m1252429_99120101014_SMDreadnoughtmain_873x627.jpg

17 Appendix

17.1 Appendix A: Interview with Slowmotion FX

Interview with Ian Bach, founder Slowmotion FX.

Q1. How many people you have working at slowmotion fx?
Ans: Between 2 and 8 people

Q2. Which vfx elements do you specialize in?
Ans: Photorealism, repairs/fixes, quick solutions that do not require CGI and solved as analogous as possible. We also focus on the artistics behind a certain shot. Besides this we have large experience with general compositing.

Q3. Are you primarily focused on feature films or do you do work on commercials also?
Ans: Worked on all kinds of projects but primarily feature films. Thats where our strength lies.

Q4. Which softwares do you use on an everyday basis?
Ans: Nuke for compositing, Da Vinci for colour grading and Syntheyes for match moving.

Q5. How valuable is communication in VFX work?
Ans: Invaluable! The most important element of them all. You can have the most talented people working on a shot but the overall result may be nice because the communication lacked short. Very often if you have a good communication you might actually come to a result that cost half the price. For example: if the shot requires a person to jump down from a roof and the studio doesn't want the actor to actually do it. This could require all kinds of plates and animations etc but if the director says exactly what he wants done one could perhaps cut off entire sections and thus save valuable time.

Q6. What conflicts are typical in a large production in terms of producer/director/etc.?
Ans: The most classical is that the director would like something that the producer does not have money for. It involves legalities involving who has the last say in terms of decision making etc. If the deal is that the scene has to be solved with a certain narrative its hard to concretize which VFX elements are actually required to solve the given problem. Its a matter of when the scene is actually regarded as 'solved' and this depends on the opinions of the director and producer. A typical issue is when the editing department makes

several cuts of a scene on a predefined budget thus makes even more work for the VFX department.

Q7. Are CGI effects generally sought in VFX work in Denmark?
Ans: No not generally. The most sought effects are invisible effects where the given effect could be filmed on set but logistics dictate that it is impractical. We have had only few requests for CGI. People do like CGI effects but the budgets are very limited in Denmark and big large CGI shots are therefore typically erased from scripts due to the nature. It is quick to make quick CGI but in order to make photoreal CGI it takes four-fold the time.

Q8. What should we pay attention when dealing with CGI effects?
Ans: Real good planning and make very careful pre-visualizations. When dealing with CGI everything is possible but keep it basic. Even if everything is possible in the CGI world try and simulate how a real shot would be done. E.g. don't make a super complicated camera move when in reality this would be impractical.

Q9. Any example of using alternative means for a shot that otherwise would require CGI?
Ans: A good example was a shot where I used coffee steam instead of particle simulations. Another example was a shot where people were meant to implode. Instead of using particle volumetrics we used animated still images of cigarette smoke. It should be remembered that you can control it less but the result will look more photoreal in a quicker time.

Q10. In terms of set extensions are there different scenarios where traditional matte paintings should be used as compared to full 3D camera projections?
Ans: Depends on the camera movement. If the camera is moved there will be parallax and a matte painting will thus be suitable and the 3D camera projection should be put to use.

Q11. In terms of VFX, what should we pay attention to when shooting on set?
Ans: Always note down camera optics and the type of camera. If you'll be shooting on other sets you need where roughly the camera is positioned and the shooting angle of the camera. This is particularly true when shooting individual elements.