

TRINITY COLLEGE DUBLIN, THE UNIVERSITY OF DUBLIN

---

**Exploring Parameters of Virtual Character Lighting  
Through Perceptual Evaluation and Psychophysical  
Modelling**

---

*by*

**Pisut Wisessing B.A., M.F.A.**

Supervisor: Dr. Rachel McDonnell

*A thesis submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy*

*in the*

School of Computer Science and Statistics

January, 2021



## Declaration

I, Pisut Wisessing, declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I, Pisut Wisessing, agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I, Pisut Wisessing, consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).



---

Pisut Wisessing

January, 2021



TRINITY COLLEGE DUBLIN, THE UNIVERSITY OF DUBLIN

## *Abstract*

### **Exploring Parameters of Virtual Character Lighting Through Perceptual Evaluation and Psychophysical Modelling**

by Pisut Wisessing

This thesis explored the parameters of virtual character lighting and their connections to the perceived emotion and appeal of the character. Our main interest is to empirically evaluate various common practices of setting up these parameters in traditional art forms, such as painting, theatre and cinematography, and their psychological effects on the perception of the character according to artistic conventions. We also aimed to standardise a general guideline for lighting design that will enhance the inner states of virtual avatars for maximum audience engagement.

We conducted an extensive set of novel psychophysical experiments attempting to assess the links between the physical properties of lighting and the responses of the audience. The results were discussed in relation to theories found in the literature of visual perception, psychology and anthropology. We adapted classic research methodologies such as the multidimensional scaling analysis, the method of constant stimuli and the method of adjustment to the modern research question of how we perceive virtual characters and what makes them engaging for various applications, for example, self-avatars on social media platforms that drew massive interest from professional developers and casual makers alike.

Some of our findings agreed and some disagreed with certain codes in cinematic lighting. Based on these newfound insights, we derived a set of lighting guidelines that can be used to enhance the emotion and appeal of digital characters and demonstrated a use case of a perceptual lighting tool. Moreover, our experiment designs, particularly the method of adjustment with real-time graphics, broke new ground for future research in virtual avatars. In summary, our contributions found applications in both industry practice and academic research.



## *Acknowledgements*

First, I would like to express my greatest gratitude to Dr. Rachel McDonnell, my supervisor, for the trust and creative freedom she has given me since the beginning of my Ph.D. journey. I also would like to thank Dr. John Dingliana, my co-supervisor while Rachel being on parental leave. I still remembered our early conversation when you two interviewed me for the position via Skype. Thank you for being patient and supportive all the way till the end.

Another person that was instrumental to my academic and professional endeavors was Dr. Timothy A. Davis, my M.F.A. adviser who was also the interim Director of the Digital Production Arts Program at Clemson University. I appreciate that you waited twelve days for my late arrival caused by a visa issue. You and the rest of the Clemson faculty taught me so much.

In the summer of 2005, Professor Steve Marschner accepted a clueless intern to his group. The experience was so eye-opening that I switched my interest from Physics to Computer Graphics. Two years later, Professor Lynn Tomlinson introduced me to the world of traditional and experimental animations. Her animation workshop altered the course of my life. I am indebted to both of you.

I am grateful to all the healthcare workers at the Counseling & Psychological Services, Cornell Heath, particularly, Dr. Wai-Kwong Wong my first counsellor who always believed that I would become an animator one day. I also extend my appreciation to the staff members at the Student Counselling and the Student Health Services at Trinity College. I am also thankful for all the support and friendship from SIGGRAPH, SIGGRAPH Asia, Cornell, Clemson, Trinity, and Thai-Dublin communities that made my life bearable during dark moments.

I would like to thank Siva Kumar Kasetty and Dagon Potter for giving me my first job at DreamWorks Animation India, as well as Nishant Khanna and John Huikku for brining me to Brown Bag Films.

This research was funded by the Science Foundation Ireland under the ADAPT Centre for Digital Content Technology (Grant 13/RC/2016) and the Game Face Project (Grant 13/CDA/2135). My thanks also go to the Development and Promotion of Science and Technology Program of Thailand and the Clemson's Digital Production Arts Program for their financial support in the past.

Most importantly, I would not be here without my family. Words may not describe my feeling right now. I love you all and I am coming home very soon.





# Relevant Publications

## Journal Paper

1. Wisessing, P., Zibrek, K., Cunningham, D. W., Dingliana, J., & McDonnell, R. (2020). "Enlighten Me: Importance of Brightness and Shadow for Character Emotion and Appeal". *ACM Transactions on Graphics (TOG)*, 39.3.

## Conference Papers

2. Wisessing, P., & McDonnell, R. (2020). "Evalu-light: A practical approach for evaluating character lighting in real-time". In *ACM Symposium on Applied Perception 2020 (SAP'20)*.
3. Wisessing, P., Zibrek, K., W. Cunningham, D., & McDonnell, R. (2019). "A psychophysical model to control the brightness and key-to-fill ratio in CG cartoon character lighting". In *ACM Symposium on Applied Perception 2019 (SAP'19)*.
4. Wisessing, P., Dingliana, J., & McDonnell, R. (2016). "Perception of lighting and shading for animated virtual characters". In *ACM Symposium on Applied Perception 2016 (SAP'16)*.



# Contents

<b>Declaration</b>	<b>1</b>
<b>Abstract</b>	<b>3</b>
<b>Acknowledgements</b>	<b>5</b>
<b>1 Introduction</b>	<b>21</b>
1.1 Motivation . . . . .	22
1.2 Methodology . . . . .	23
1.3 Scope . . . . .	24
1.4 Contributions . . . . .	25
1.5 Thesis Overview . . . . .	26
<b>2 Related Work</b>	<b>29</b>
2.1 Character Lighting . . . . .	29
2.1.1 Early History . . . . .	30
2.1.2 Character Lighting in Computer Graphics . . . . .	32
2.2 3D Animation Production Pipeline . . . . .	33
2.2.1 Modelling and Surfacing . . . . .	34
2.2.2 Rigging and Animating . . . . .	35
2.2.3 Lighting and Rendering . . . . .	37
2.2.4 Visual Perception of CG Scenes . . . . .	38
2.3 Visual Perception . . . . .	39
2.3.1 Psychophysics of Lighting . . . . .	40
2.3.2 Shading, Shadows and Light Direction . . . . .	42
2.3.3 Movement . . . . .	44
2.3.4 Emotions . . . . .	44
2.3.5 Appeal . . . . .	45

2.3.6	Uncanny Valley . . . . .	47
2.3.7	Language and Culture . . . . .	48
2.4	Perceptual Experiment Design and Analysis . . . . .	49
2.4.1	Multidimensional Scaling (MDS) . . . . .	50
2.5	Conclusion . . . . .	50
<b>3</b>	<b>Lighting Parameter Investigation</b>	<b>53</b>
3.1	Introduction . . . . .	53
3.2	Stimuli . . . . .	54
3.2.1	Character . . . . .	55
3.2.2	Emotions . . . . .	55
3.2.3	Lighting Parameters . . . . .	56
3.2.4	Movies . . . . .	59
3.3	Experiment . . . . .	59
3.3.1	Participants . . . . .	60
3.4	Results . . . . .	60
3.4.1	Recognition Accuracy . . . . .	61
3.4.2	Intensity . . . . .	61
3.4.3	Appeal . . . . .	62
3.5	Discussion . . . . .	63
<b>4</b>	<b>Parameter Selection</b>	<b>65</b>
4.1	Introduction . . . . .	65
4.2	Stimuli . . . . .	66
4.2.1	Three-point Lighting . . . . .	66
4.2.2	Light Intensities and Key-to-Fill Ratio . . . . .	67
4.2.3	Movies . . . . .	67
4.2.4	Environment . . . . .	69
4.3	Experiment 1A - Dissimilarity . . . . .	69
4.3.1	Experiment . . . . .	69
4.3.2	Participants . . . . .	70
4.3.3	Multidimensional Scaling (MDS) . . . . .	70
4.3.4	Results . . . . .	71
4.4	Experiment 1B - Un-anchored MDS . . . . .	72
4.4.1	Experiment . . . . .	72

<i>Contents</i>	11
<hr/>	
4.4.2	Participants . . . . . 73
4.4.3	Results . . . . . 73
4.5	Parametric Model . . . . . 73
4.6	Experiment 2 - Model Evaluation . . . . . 75
4.6.1	Stimuli . . . . . 76
4.6.2	Experiment . . . . . 76
4.6.3	Participants . . . . . 76
4.6.4	Results . . . . . 77
4.7	Discussion . . . . . 78
<b>5</b>	<b>Perception of Appeal and Emotion</b> . . . . . <b>81</b>
5.1	Introduction . . . . . 81
5.2	Cartoon Experiments . . . . . 81
5.2.1	Stimuli . . . . . 82
Recordings	. . . . . 83
Lighting	. . . . . 84
Movies and presentation	. . . . . 84
5.2.2	Laboratory Experiment (Baseline) . . . . . 86
Participants	. . . . . 87
5.2.3	Online Experiments . . . . . 89
5.2.4	Audio, Background and Movement Experiments . . . . . 93
Audio	. . . . . 94
Background	. . . . . 95
Movement	. . . . . 95
5.2.5	Cartoon Experiments - Discussion . . . . . 95
5.3	Realism Experiments . . . . . 96
5.3.1	Stimuli . . . . . 98
5.3.2	Experiment Design . . . . . 98
5.3.3	Results . . . . . 98
5.3.4	Realism Experiments - Discussion . . . . . 100
5.4	General Discussion . . . . . 100
<b>6</b>	<b>Perceptual Lighting Tool</b> . . . . . <b>105</b>
6.1	Introduction . . . . . 105
6.2	Tool Development . . . . . 106

6.2.1	Real-time Lighting & Rendering . . . . .	106
6.2.2	Lighting Control & Interface . . . . .	106
	Tool A . . . . .	107
	Tool B . . . . .	108
6.3	Experiment . . . . .	108
6.3.1	Character . . . . .	108
6.3.2	Presentation . . . . .	109
6.3.3	Design . . . . .	109
6.3.4	Intensity Task . . . . .	109
6.3.5	Appeal Task . . . . .	110
6.3.6	Speed & Accuracy Task . . . . .	110
6.3.7	Usability Questionnaire . . . . .	110
6.3.8	Participants . . . . .	111
6.4	Results . . . . .	111
6.4.1	Intensity . . . . .	112
6.4.2	Appeal . . . . .	113
6.4.3	Speed . . . . .	114
6.4.4	Accuracy . . . . .	114
6.4.5	Usability . . . . .	114
6.5	Validation . . . . .	115
6.5.1	Brightness . . . . .	115
6.5.2	KTFR . . . . .	115
6.5.3	Time . . . . .	115
6.6	Discussion . . . . .	117
<b>7</b>	<b>Conclusion</b>	<b>119</b>
7.1	Contributions . . . . .	119
7.1.1	The Character Lighting Guideline . . . . .	119
7.1.2	The Psychophysics of Character Lighting . . . . .	120
7.1.3	Method of Adjustment with Real-time Graphics . . . . .	121
7.2	Future Work . . . . .	121
<b>A</b>	<b>Summary of Main Effects and Interactions</b>	<b>125</b>
	<b>Bibliography</b>	<b>133</b>

# List of Figures

2.1	A close-up of sfumato technique in <i>Mona Lisa</i> by Leonardo da Vinci (1506) . . .	30
2.2	A close-up of tenebrism technique in <i>John the Baptist</i> by Caravaggio (1604) . . .	31
2.3	An abstraction of 3D animation production pipeline . . . . .	33
2.4	Facial shapes illuminated from above and below used in Hill and Bruce [1996]’s experiments . . . . .	43
2.5	Characters from Pixar’s <i>Inside Out</i> representing basic emotions categorised by Ekman: disgust, fear, joy (happiness), sadness and anger. Note that surprise was not included in the movie. . . . .	45
2.6	A graph from [Mori et al. 2012] illustrating the Uncanny Valley effect of still and moving stimuli. . . . .	47
3.1	Mery character . . . . .	55
3.2	Frames from <i>Alfred Hitchcock’s Rear Window</i> (1954) showing high-key (left) and low-key (right) lighting . . . . .	56
3.3	Frames showing Mery emotions rendered in the combination of “light from above” (Ab), “high contrast” (Hi) / “low contrast” (Lo), and CG-shaded / Toon-shaded conditions . . . . .	57
3.4	Frames showing Mery emotions rendered in the combination of “light from below” (Be), “high contrast” (Hi) / “low contrast” (Lo), and CG-shaded / Toon-shaded conditions . . . . .	58
3.5	Frames showing Mery emotions rendered in the “no” directional light condition	59
3.6	Main effect of Emotion on recognition accuracy . . . . .	61
3.7	Averaged ratings of some of the main effects and interactions for intensity (top) and appeal ratings (bottom). Hi: high contrast, Lo: low contrast, Ab: above light, Be: below light, No: no directional light. . . . .	62
4.1	Three-point lighting setup with detailed light directions . . . . .	66

4.2	Contribution from key light and fill light were rendered separately and later combined and manipulated to create different levels of brightness and KTFR. The picture shows adding 100% key light brightness and 100% fill light brightness to make 100% brightness at 1:1 KTFR. . . . .	67
4.3	Still images taken from the 16 movies, rendered in different key light intensities (displayed in percentages) and key-to-fill ratios (displayed as ratios) . . . . .	68
4.4	A screenshot of a trial in the anchored MDS experiment (Experiment 1A) . . . . .	69
4.5	The MDS stress plot suggests that two dimensions are sufficient to explain the dissimilarity as the stress is reduced more than half from one to two dimensions but does not decrease much from two to three dimensions. . . . .	71
4.6	The blue data-points, representing the different levels of key intensity and KTFR are shown in the MDS perceptual space along the two dimensions. The yellow lines represent the proposed parametric log-polar model that best fits the data. . . . .	72
4.7	The parametric model of Logvinenko and Maloney [2006] that best fits the observers' data. The points (light-surface pair $ij$ ) lie along the radii of concentric ellipses. . . . .	74
4.8	The red lines show the location in the perceptual space of the stimuli with 'adjusted' brightness. . . . .	76
4.9	perceptually 'adjusted' stimuli set. Notice how the overall brightness is more consistent across the rows than in Figure 4.3. . . . .	77
4.10	A screenshot of a trial in the validation experiment . . . . .	78
4.11	A subset of stimuli (top) representing the perceptual space of brightness and shadow intensity (bottom) that could be used in future experiments. . . . .	80
5.1	The experiments were divided into cartoon and realism experiments. . . . .	82
5.2	(left to right) Mery, Jasmine, Franklin and Malcolm characters rendered in their original shaders . . . . .	82
5.3	(left to right) Mery, Jasmine, Franklin and Malcolm characters rendered in our experiment shaders . . . . .	83
5.4	(left to right) Mery, Jasmine, Franklin and Malcolm portraying anger . . . . .	84
5.5	Nine lighting conditions used in the Cartoon Experiments, which were rendered with different (key light) Brightness and key-to-fill ratios (KTFR). . . . .	85
5.6	Franklin and Mery with neutral expression in front of two complex background scenes. . . . .	86
5.7	Interaction between Brightness and Emotion for ratings on Intensity in the baseline experiment. Star labelled lines point to significantly different means according to the post-hoc test. Error bars show standard error of the means. . . . .	89



5.8	Interaction between Brightness and Emotion for ratings on Appeal in the baseline experiment. Star labelled lines point to significantly different means according to the post-hoc test. Error bars show standard error of the means. . .	90
5.9	Interaction between Brightness and Emotion for ratings on Intensity in the online experiment. Star labelled lines point to significantly different means according to the post-hoc test. Error bars show standard error of the means. . .	91
5.10	Interaction between Brightness and Emotion for ratings on Appeal in the online experiment. Star labelled lines point to significantly different means according to the post-hoc test. Error bars show standard error of the means. . .	92
5.11	Samples of stimuli used in the Realism Experiments. From left to right shows stylisation levels: Realistic, Middle and Toon. Row 1: sadness male-100% brightness-1:1 KTFR, Row 2: sadness male-100% brightness-16:1 KTFR, Row 3: happiness female-100% brightness-1:1 KTFR, Row 4: happiness female-100% brightness-16:1 KTFR. . . . .	97
5.12	Interaction between KTFR and Shape for ratings on Appeal. Star labelled lines point to significantly different means according to the post-hoc test. Error bars show standard error of the means. . . . .	99
6.1	Experiment setup in Unreal Engine 4. Top: digital lighting Tool A being used in the Speed Task. Bottom: perceptually-based Lighting Tool B being used in the Appeal Task. . . . .	107
6.2	Usability Questionnaire . . . . .	111
6.3	Main effects and interactions of intensity task. . . . .	112
6.4	Main effects and interactions of appeal task. . . . .	112
6.5	Screenshots from our real-time system showing the averaged key light and KTFR values chosen by the participants for the intensity task. . . . .	113
6.6	Screenshots from our real-time system showing the averaged key light and KTFR values chosen by the participants for the appeal task. . . . .	113
6.7	Usability questionnaire ratings . . . . .	114
6.8	Correlation between Tool A and B of the individual parameter values of the appeal task selected by the participants for the intensity task. . . . .	116
6.9	Correlation between Tool A and B of the individual parameter values of the appeal task selected by the participants for the appeal. . . . .	116
6.10	Mean and individual parameter values of the appeal task selected by the participants for the intensity task. . . . .	118
6.11	Mean and individual parameter values of the appeal task selected by the participants for the appeal task. . . . .	118



# List of Tables

2.1	Terminology of similar lighting design in different fields of visual arts . . . . .	32
3.1	Validated affective sentences for anger, sadness, disgust, fear and happiness . . .	55
4.1	The coordinates of the blue data-points in Figure 4.6, representing the locations of the stimuli the MDS perceptual space. . . . .	73
4.2	Weight and constant values of the parametric model (Equation 4.3) that best fits the MDS perceptual space—the yellow lines in Figure 4.6. . . . .	75
4.3	The coordinates of the red data-points in Figure 4.8, showing the location in the perceptual space of the stimuli with “adjusted” brightness. . . . .	79
5.1	Validated affective sentences for neutral, anger, sadness, fear and happiness . . .	83
5.2	The confusion matrix of the recognition rating of Mery character . . . . .	88
5.3	The confusion matrix of the recognition rating of Franklin character . . . . .	88
5.4	Summary of conditions and participants in Cartoon Lab Experiments. . . . .	93
A.1	Chapter 5 - Baseline Experiment: main effects and interactions with post-hoc analysis. . . . .	126
A.2	Chapter 5 - Online Experiment: main effects and interactions with post-hoc analysis. . . . .	126
A.3	Chapter 5 - Effect of Audio: main effects and interactions with post-hoc analysis. . . . .	127
A.4	Chapter 5 - Effect of Background: main effects and interactions with post-hoc analysis. . . . .	128
A.5	Chapter 5 - Effect of Movement: main effects and interactions with post-hoc analysis. . . . .	129
A.6	Chapter 5 - Realism: Brightness: main effects and interactions with post-hoc analysis. . . . .	130
A.7	Chapter 5 - Realism: KTFR: main effects and interactions with post-hoc analysis. . . . .	131

A.8 Chapter 5 - Realism: KTFR (continued): main effects and interactions with post-hoc analysis. . . . .	132
--	-----

*For Mom and Dad*



## Chapter 1

# Introduction

The demand for cartoon animation content is at an all time high, and on the rise, driven largely by social media (e.g., Apple's Memoji and Facebook Horizon), and magnified by the availability of powerful mobile devices (e.g., the depth sensor in the iPhone X). Content creators are no longer limited to big-budget movie studios with teams of world-leading artists. There are thousands of animation studios with diverse budgets and expertise worldwide creating content for movies, advertising, TV, games, etc. Additionally, animation content is increasingly being generated by technical developers and non-professionals for virtual-assistants, apps, social-media and VR/AR. Character lighting is one aspect of content creation that is particularly important for establishing the look and feel of a character [Lowell 1992]. This study was motivated by identifying a clear need for standardised guidelines on lighting virtual characters for non-experts who wish to enhance emotion and increase the appeal of their characters.

Disney's classic principles of animation [Thomas and Johnston 1995], a set of guidelines to assist animators, do not incorporate lighting and how to use it to alter the appeal and intensity of cartoon characters. However, lighting is often used in everyday language and is something people relate to, for example, "she brightened up", and "he told me his darkest fears" [Barchard et al. 2017]. The fact that we have so many metaphors relating brightness to good and darkness to bad has been shown in Psychology research to be due to developmental experiences which pair these factors, for example social and physical rewards being more prevalent during the daytime while the darkness of night hides potential dangers [Landau 2014; Meier et al. 2004]. The effect is also common in popular culture (e.g., evil is paired with darkness and good with light). This suggests that lighter scenes would be perceived as more inviting, and friendly and darker situations are more sinister and gloomy.

Artists have been using lighting in this way to create mood and to influence the appeal of a character or object [Gurney 2010; Brown 2016]. Artists want the audience to feel a connection with the character and use light to support and strengthen what they want to communicate. One often used measure to control lighting is the key-to-fill ratio (KTFR) which compares the illumination due to the primary lighting source, i.e., key light; and the secondary light that brightens the shadows, i.e., fill-light. Artists and cinematographers use high-key lighting (with light shadows) to create a hopeful mood, or low-key lighting (with dark shadows) to add a sense of gloom [Pramaggiore and Wallis 2005]. Scenes lit with dark shadows rarely appear happy [Landau 2014]. Our objective is to examine how viewers perceive different illumination levels in computer animation and what effect this may have on their perception of virtual characters.

On the other hand, there is mounting evidence that people do not fully discount illumination when perceiving surface reflectance. Logvinenko and Maloney [2006], for example, investigated the relationship between illumination and perceived surface lightness. They found that while albedo (surface reflectance) was the primary determinant of surface lightness, the intensity of the illumination (i.e., the shading) also had an effect. That is, people do not fully discount the effect of shading in lightness perception. There has been a long history of investigation in perceptual psychology, psychophysics, and vision science, focusing on the relationship between the physical changes in illumination intensity one side and the perception of either surface reflectance (lightness) or surface illumination (brightness) on the other side. However, there are not many studies striving to explain the artistic lighting conventions, based on centuries of observations, with perception theories and carefully designed psychophysics experiments.

## 1.1 Motivation

The motivation of this research was twofold:

Visual constancy is a principle asserting that the perceived appearance of objects by the human eye remains relatively constant even under large variations in the lighting conditions. This may imply that the perception of a character's appearance, and, as a result, its perceived emotional states, would be minimally affected by illumination conditions, contrasting to the established conventions of artists using lighting to create mood and to influence the appeal of a character or object and the convention in cinematography, which



has largely been adopted in 3D computer animation. The discrepancy between the psychophysics theory and traditional practice of artists motivates the need of carefully designed perceptual experiments aimed at evaluating and quantifying the effect of lighting on emotional character, for particularly its role in the recognition of emotion, emotion intensity and the overall appeal, as these are the most important factors for audience engagement.

Secondly, there is currently a high demand in CG cartoon characters from the Disney's blockbusters such as *Frozen* and *Moana* to Netflix's streamed content, and from AAA titles, e.g. *The Legend of Zelda*, to casual mobile games like *Pokemon Go*. as evidenced by the recent estimated value of global industry has doubled in a decade to \$264 billions) [Research and Markets 2011]. Content creators are no longer limited to big-budget Hollywood studios with teams of world-leading artists. There are thousands of animation studios with ranging budgets and expertise worldwide creating content for movies, advertising, TV, games, etc. Additionally, animation content is increasingly being generated by technical developers and non-professionals for virtual-assistants, apps, social-media and VR/AR. The lack of artistic training in the later group of content creators identifies a clear need for standardised guidelines on lighting. The empirical study of CG character lighting grounded in psychophysics and perception can produce a set of data-driven protocols that anyone can follow.

This study will not only quantify the conventional thought process in lighting design for a better and more effective of command of the light, but also fill some of the experience gaps among casual lighters and elevate the quality of animation across the various emerging content markets.

## 1.2 Methodology

The design of lighting in Computer Graphics is directly derived from cinematography, and many digital artists follow the conventional wisdom on how lighting is set up to convey drama, appeal, or emotion. In this thesis, we are interested in investigating the most commonly used lighting techniques to more formally determine their effect on our perception of animated virtual characters, using CGI stimuli together with a mix of traditional and novel methods in psychophysics.

Psychophysics is instrumental for perceptually adaptive graphics, as it explores the relationship between the objective stimuli and subjective responses. The majority of this thesis employed the method of constant stimuli. For the stimuli, we obtained industry-standard character models and commissioned professional animators to create a corpus of animations of the characters enacting several key emotions. The animated characters were then rendered using a range of shading and lighting conditions. For the responses, Likert's scale is a psychometric test that was used to collect the subjective ratings.

A common challenge with lighting studies is the explosion of the vast number of the parameters. The multidimensional scaling (MDS) analysis was a useful tool for capturing the connections between the physical parameter space and perceptual space which has been proven to be robust with just a small number of trained participants in low-level perception studies [Logvinenko and Maloney 2006]. The resulting MDS plots were used to determine the threshold of the stimuli, evenly sampled from the perceptual space and derive a manageable subset of differing illumination conditions that were perceptually distinguishable from each other and adequately represent the space as a whole. As a result, our research will be simplified and yield concise results by exploring reduced but representative parameter spaces and number of samples.

From the MDS plots, we also proposed a parametric model describing the perceptual space and developed a prototype of a new perceptual lighting tool based on the mathematics of the model. With the new tool and the latest graphics hardware and game engine technology, we developed a real-time method-of-adjustment experiment design for exploring parameters of lighting in real-time. The new paradigm delivered interactive high-fidelity renders and offered new possibilities in virtual character research. Overall, this thesis extends the theories and methodology of much previous work in the field of psychophysics in an attempt to empirically explain the perceptual effects of CG character lighting on the audience.

### **1.3 Scope**

This thesis presented a suite of perceptual studies on CG character lighting which includes proposing a psychophysical model of the lighting perceptual space as a parametric function, showing the effect of certain lighting conditions on recognition of emotions, emotion intensity and the overall appeal of a character, and prototyping a new lighting tool concept.

Due to the overwhelming number of parameters, we mainly focused on the most commonly used lighting design in portraiture, the three-point lighting setup. The preliminary experiment (Chapter 3) eliminated the further testing of light direction and shading technique. The psychological effect of colour in lighting is another important and complex issue that could by itself constitute a full other thesis. However, it was not the primary aim of this study and lied outside of the scope. To avoid the influence of light and material colours, we set our light colour temperature to the standardised 6500K white point, and tried kept all the materials in the scene in the neutral grey.

Despite the goal to study lighting effects in the context of computer animation content and most of our stimuli are animated stylised characters, we attempted to generalise our results over realistic characters by including a set of characters with different levels of stylisation in one of our experiments (Chapter 5, Section 5.3).

To reduce other interfering circumstances, the majority of the experiments were conducted in a completely dark room, and the stimuli were displayed on a colour-calibrated monitor. Nonetheless, some of the our key results were later confirmed by online experiments with participants using different display device types and sizes.

## 1.4 Contributions

We conducted a series of psychophysics experiments and, as a result, gained new insights into the effects of different lighting conditions on the perception of CG character lighting. To the best of our knowledge, this thesis was one of the few studies that empirically assess the connections between lighting and the subjective responses from the audience.

In particular, we looked at the problem with a new experiment paradigm. Applying a technique from a low-level perception study to a more complex but practical stimuli, we have developed and validated a parametric model describing a proximity structure (relative locations of different lighting conditions in perceptual space) of perceived CG characters lighting. The mathematical model was later observed and utilised to effectively and systematically sample a character lighting perceptual space to produce a representative set of stimuli for later experiments that investigated the effect of brightness and the strength of shadow on the perception of higher level factors such as emotion and appeal. We have carefully designed our experiments to determine the influences of isolated factors such as the intensity of the light sources directly illuminating the character, as well as the

modulations by interactions with other elements in CG production such as motion, audio and background.

Through extensive analysis, we derived a concise set of lighting guidelines that could be used by either professionals in well-established studios needing to optimise their lighting processes, or a small team of developers with little artistic training wanting fast production of appealing virtual characters.

Finally, we also proposed a perceptual lighting tool design based on our results with comparable usability but higher efficiency than traditional tools in repetitive lighting tasks, as well as demonstrated that the new tool, implemented in a high-fidelity game engine, could also be used to explore the lighting parameters and produce results with smaller thresholds and in a shorter amount of time, compared to the method of constant stimuli employed in most experiments of this thesis.

## 1.5 Thesis Overview

The rest of the thesis has been divided into the following chapters:

- **Chapter 2** provides an overview of the previous background and related work on character lighting, computer animation production and visual perception.
- **Chapter 3** presents the result of the a preliminary study on the perception of CG lighting. This chapter particularly examines the effects of lighting direction (light from above vs. light from below), and contrast (low vs. high and shading techniques (CG shading vs toon shading)).
- **Chapter 4** explores the perceptual space of lightness and shadow discrimination in CG character lighting by employing a multidimensional scaling analysis, typically used in low-level perceptual experiments. A psychophysical model was proposed based on the result.
- **Chapter 5** delivers a series of perceptual experiments on CG character lighting, determining the effect of brightness and shadows on recognition of emotions, emotion intensity and the overall appeal of cartoon and realistic characters.
- **Chapter 6** uses the method of adjustment to develop a new perceptually-based tool for CG character lighting, proposes a new real-time experiment design, and validates the previous results.

- **Chapter 7** summarises our contributions, and provides a discussion of future work.



## Chapter 2

# Related Work

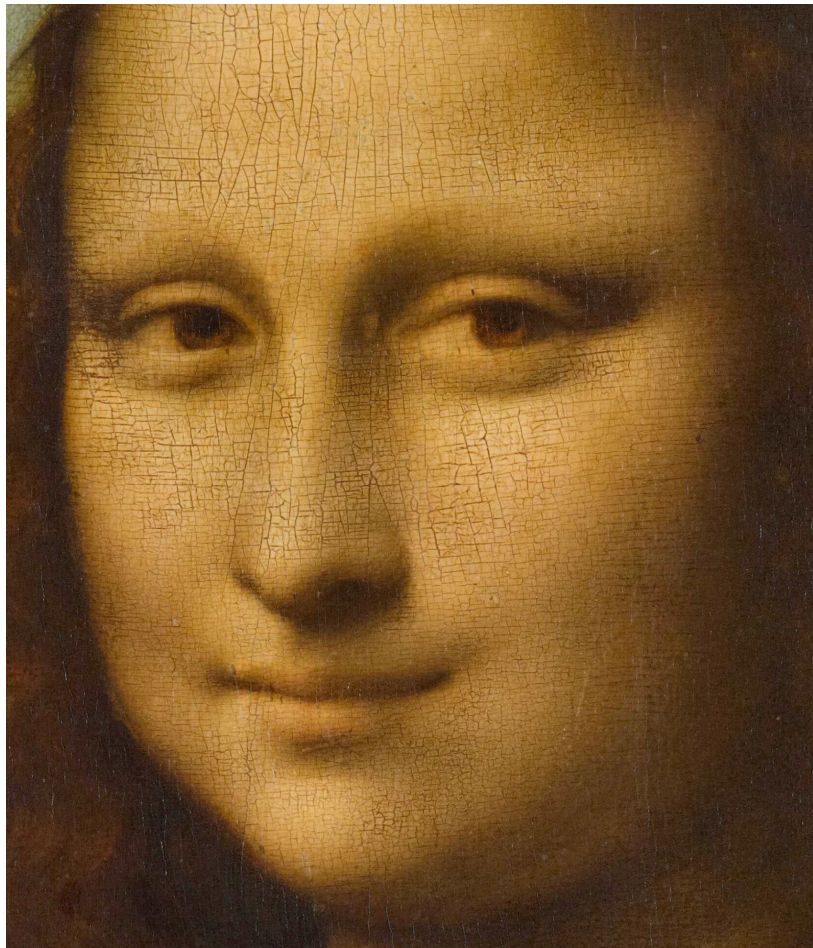
Lighting has been used to convey emotions and enhance appeal in all forms of art for centuries, from painting to film. However, there have not been many studies that empirically examine the psychological effects of lighting design, especially when employed in a virtual environment with a virtual character such as in a computer animation scene. There are many intervening factors from the perception of the physical world such as light and shadows to the illusion of a computer graphics scene imitating the real world, or from the cultural influences to the artistic conventions. In an attempt to truly understand the connections between long-established lighting techniques and their applications to modern computer-generated characters, some important pieces of the related literature in the history of lighting, computer graphics, and visual perception will be examined and presented in this chapter.

### 2.1 Character Lighting

In computer graphics, there are three main components in creating an image from a virtual 3D scene: a camera, an object and a light. All objects in the scene need to be illuminated to be visible but in this thesis, we will focus our attention on just illuminating a character. In computer animation and other digital art forms, **character lighting** is a common term referring to *portrait lighting* when it is applied to a virtual computer-generated character. Before the advent of computer graphics, the craft of lighting design was practised in the fields of painting, theatre, photography, and cinematography. It is a standard practice in art education to study and draw inspiration from the old masters, and so will our study.

### 2.1.1 Early History

Lighting in arts has been developed as early as in the Renaissance. Leonardo da Vinci (1452-1519) was credited for the development of **sfumato** (Figure 2.1), meaning to smoke or to soften, was one of the canonical techniques in portrait paintings during the period [McIver 2017; Hall 1992]. Sfumato can be achieved by using the overlay of fading transparent oil paints to produce soft colour transitions, especially in the skin tones. Highlights and shadows create the sense of volume in paintings, at the same time, sfumato softens the gradation in the skin tone, making them look more natural and appealing. The modern version of sfumato in photographic portraiture is using large softboxes (large-area diffuse light source) to produce soft lighting. This technique smoothly blends the bright area and dark area with no visible boundary, similar to the concept of **low contrast** in photography and computer graphics.



**Figure 2.1:** A close-up of sfumato technique in *Mona Lisa* by Leonardo da Vinci (1506)

On the other hand, Caravaggio (1571 - 1610) was generally respected as the inventor of



**tenebrism**, meaning darkness or gloom, which is the use of dramatic light, or the extreme rendition of **chiaroscuro** (light and dark), which became popular in the Baroque movement [Moffitt 2004]. Tenebrism focuses on the stark light-shadow differences while chiaroscuro covers a broader range of contrast that enhances the perception of three-dimensionality (Figure 2.2). Tenebrism could be thought as the spotlight effects in the cinematography, or low-key lighting popular in *Film Noir* visual style, or the **high contrast** lighting, compared to sfumato [McIver 2017]. See Table 2.1 for a summary of terminology.



**Figure 2.2:** A close-up of tenebrism technique in *John the Baptist* by Caravaggio (1604)

Fast forward to the contemporary lighting in theatre, photography and cinematography, the two lighting approaches are believed to evoke different emotions from the viewers, and lighting, in general, has been widely used to create moods in fields of visual arts—including computer animation. However, the association between lighting styles and emotional

responses, particularly in film [Grodal 2005], are mostly conventional practices based on observations. Many filmmakers experimented with different types of lighting and concluded their assumptions. To test the pragmatism of film theory, a recent empirical study by Poland [2015] found “low-key” (high contrast) stimuli produced reports of lightheartedness contrary to the beliefs of many theorists and cinematographers. However, the study used short films with storylines as stimuli and the results were concluded to be partially influenced by the narrative.

Field	Term	
Painting	<i>sfumato</i>	<i>tenebrism</i>
Film	<i>high key</i>	<i>low key</i>
Computer Graphics	<i>low contrast</i>	<i>high construct</i>

**Table 2.1:** Terminology of similar lighting design in different fields of visual arts

### 2.1.2 Character Lighting in Computer Graphics

Lighting in computer graphics (CG) is mostly derived from traditional cinematography, and artists follow film *conventions* on how lighting is set up to convey drama, emotion and appeal. There were only a few studies that empirically investigated the influence of lighting in storytelling. De Melo et al. [2007] and Seif El-Nasr et al. [2006] studied effects of lighting in creating emotion in virtual characters and video-games respectively, but both studies focused mainly on the implementation and did not conduct any perceptual experiments.

To better understand CG lighting and its perception, first, we will examine how a CG scene is put together from the ground up in the next section. Although explained as an abstract high-level process, it should give enough details for considering the perceptual effects of each element. A CG scene description could be broadly divided into the static description such as shape (geometry) and material, created during the *modelling* and *surfacing* processes in the creation pipeline, and the dynamic description which are generated by *rigging* and *animating*. Finally, with *lighting* bridging between abstract and visual representations, *rendering* interprets scene descriptions into visible images of the 3D scene (Figure 2.3).

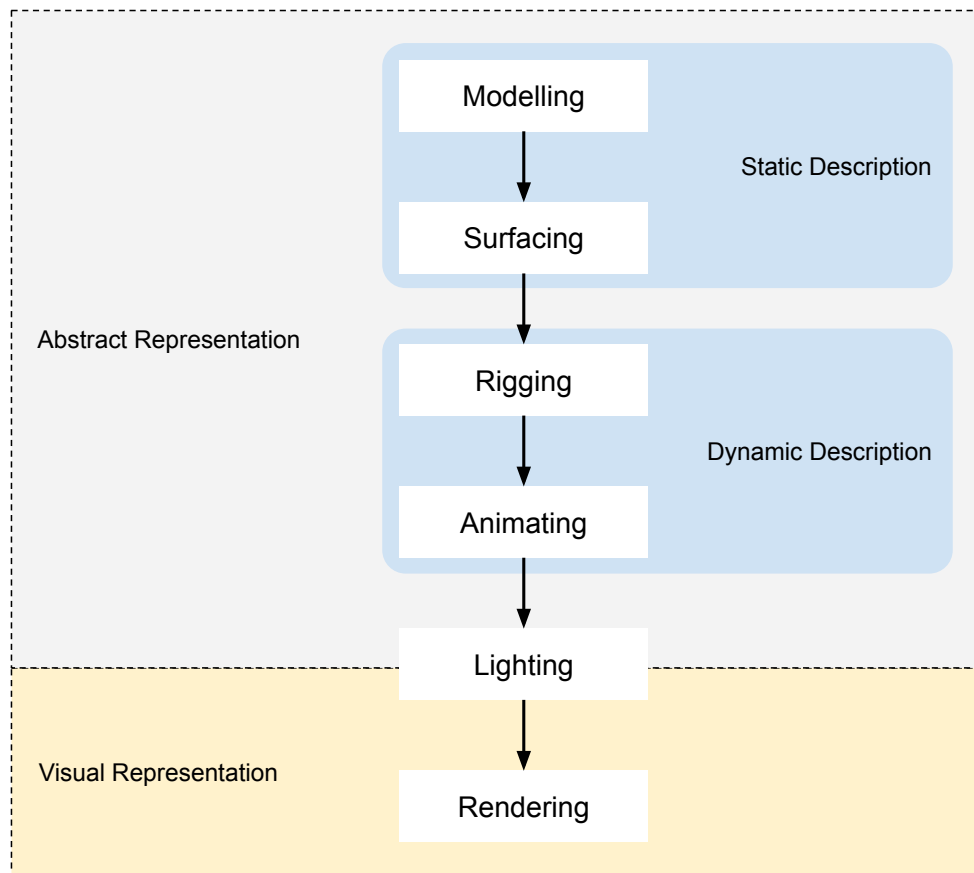


Figure 2.3: An abstraction of 3D animation production pipeline

## 2.2 3D Animation Production Pipeline

Providing both objective and subjective visual information, lighting design is a crucial stage in the 3D animation production workflow that is often explained as an assembly line. A new scene starts from one end (modelling) of the line and finishes at the other end (rendering), and in the middle, there are several other departments to carry out specific tasks. An artist receives a piece of work from a (virtual) pipe and performs a task on it. Upon finishing, the artist will then send it down another pipe to a different artist, hence the term *pipeline*. The production pipeline is dynamic, modular, and flexible, and often different from one company to another, depending on the available technologies and specific needs of each type of work (animation, visual effects, games, virtual reality, etc.). Although this thesis only concerns character lighting, other processes in the pipeline also dictate how a character interacts with the light. The following sections will provide a brief introduction to the most common stages in a 3D production pipeline and a discussion on how they can affect the perception of character lighting.

### 2.2.1 Modelling and Surfacing

As mentioned earlier, to render a CG image, one of the three required elements is the object of interest, and a process to create it called modelling. Modelling is constructing a mathematical representation of a physical object in the three-dimensional virtual space. The simplest basis of modelling is points or vertices carrying information such as position, normal or colour. Points can be connected to create lines or edges (two points), triangles (three points), or quadrangles (four points), and so on. A generic word "polygon," can refer to any closed shape with three or more vertices [Masson 2007].

A character's face can be modelled or represented by a polygonal mesh which is a set of connected polygons forming a complex shape. The rule of thumb in modelling is to create a representative (e.g. a polygon face) that best resembles the reference (e.g. a real face or concept drawing of a face) with the lowest number of polygons as the computational cost of rendering is directly proportional to the number of vertices. This practice leads to quality and performance trade-off, especially for mobile games with limited computing resources that often settle with low-poly modelling.

There is another method to define surfaces using piece-wise polynomial functions such as non-uniform rational basis spline (NURBS) surfaces [Rogers 2001], which was superior to polygonal modelling in creating smooth surfaces with the same number of vertices despite having few caveats of its own. However, since the introduction of subdivision surfaces [Catmull and Clark 1978], NURBS and other spline-based modelling have become less popular and will not be discussed further.

There is a wide range of modelling styles of virtual characters, and it is difficult to compare one to another. There are many definitions of different levels of geometric stylisation. However, the two most common scales are iconic/stylised and photorealistic [McCloud 1993; Ritchie et al. 2005]. Furthermore, modelling can also be subdivided into technical elements such as shape and levels of tessellation. These categorisations allow us to track their contribution to the overall perception. For example, the influence of shape stylisations on the perceived appeal and expressivity [Zell et al. 2015]. As different levels of stylisation and details interact with lighting differently, we also investigate the shape effect on the perception of character lighting in Chapter 5 (Figure 5.11).

Modelling only builds the geometric boundaries of an object. The description dictating surface-light interactions is constructed in the next stage called *surfacing*. Also known as

*shading*, this process generates materials and defines their optical properties such as colour, reflectance or translucency [Masson 2007]. In the early development of computer graphics, surfacing was often referred to as texturing, as texture images are simply mapped onto the polygon to determine its appearance [Catmull 1975]. However, in modern computer graphics, surface properties are governed by *shaders*, the code snippets describing how a surface should react to incident light, and hence the name shading. A simple shader could display a single solid colour or sample predefined colour from a texture image. Shaders, depending on the rendering algorithms, could also dictate how shadows are depicted. In the past decades, many physically-based shaders have been developed to produce realistic materials. *Bidirectional reflectance distribution function (BRDF)* [Nicolodemus et al. 1992]—the relation of the incident and reflected light—and *sub-surface scattering (SSS)* [Jensen et al. 2001]—light enters the material, bounces inside and reflects to the surface—are the current industry-standard techniques in creating complex material appearances such as metal, skin and hair.

CG material representations have been perceptually investigated extensively in the attempt to relate mathematical models such as the BRDF to human perception [Anderson 2011; Fleming 2014], or assess the accuracy compared to perceived real-world materials [Filip et al. 2018]. Although the perceptual process of material discrimination is complex and not well understood as evidenced by the pattern of errors often found in appearance studies, humans are still adept at perceiving different materials [Fleming 2014; Maloney and Brainard 2010]. Our brain can reliably determine the similarity or dissimilarity of shared visual features, and a model of material appearance that aligns with human perceptual judgements has been proposed recently by Lagunas et al. [2019]. Another significant ability of our vision system is spotting small imperfections in modelling or shading that can trigger negative reactions [Seyama and Nagayama 2007], as well as mismatching the style of shape and material could result in undesired psychological effects [Zell et al. 2015]. We will discuss this negative response, commonly referred to as the *Uncanny Valley Effect* in more detail in Section 2.3.6

### 2.2.2 Rigging and Animating

So far, we have been discussing the creation of a static CG scene. There are several techniques an artist can deploy to add movement, such as using a physics simulation, but in the context of character lighting, the models are rigged and predominantly animated by

hands or a performance capture system. While each vertex of a polygon mesh can be moved individually, such procedure is hugely cumbersome and time-consuming. A practical approach is to use a set of simplified representations such as line segments commonly called “joints” or “bones” to influence more extensive and complex geometry [Magnenat-Thalmann et al. 1988], similar to the actual bones in a body controlling the movement of the muscle and everything else around them. The process of setting up the virtual bone hierarchy and how it influences each part of the polygon mesh is called *rigging* [O’Hailey 2018], and the skeleton structure is commonly referred to as a *rig*.

Skeletons are suitable for controlling most body parts except the face where skin deformations are more delicate, particularly when emoting. For high-quality facial expressions, using *blendshapes* can provide the animators far greater command and higher fidelity. Blendshapes are copies of the original polygon mesh but have been altered to a specific—and often intricate—deformation [Parke 1972]. A linear (or non-linear) interpolation function is set up to blend these shapes. This technique is favourable for facial animation as it reduces the number of controllers yet strikes precise poses [Orvalho et al. 2012]; for example, an animator can have a single control to shift between a neutral and smiling mouth. Both the skeleton and blendshape can be driven by an artist or motion capture data. The former can achieve a broader range of emotion and is often exaggerated, whereas the latter is realistic and true to the motion of the actor whose performance is being captured [Ruhland et al. 2017]. This thesis focussed on the stylised characters, and hence we commissioned animators to hand-craft our animations to match the cartoony style of the characters.

Most dictionary definitions of animation involve movements or making something appear to move; however, the word *animation* comes from Latin, meaning *from of life* or *imparting life live* [Dictionary 1989]. To many animators, making it move is not animation, but the mechanics of giving life [Thomas and Johnston 1995]. Thomas and Johnston [1995] introduced the Disney’s twelve basic principles of animation in their book, *The Illusion of Life: Disney Animation*, as a guideline for producing more realistic animations, which have been adopted by both 2D and 3D animators, more specifically, John Lasseter [1987] formally introduced the principles to the computer graphics community at the 1987 SIGGRAPH conference. The principles do not only deal with realism complying the law of physics (squash and stretch, slow in and slow out), but also the abstraction of emotion and attractiveness (exaggeration, appeal).

### 2.2.3 Lighting and Rendering

Up until this point in the pipeline, everything in the scene is well-defined. However, the image is still total darkness until it undergoes the process of lighting and rendering which transform mathematical representations to visible objects using various statistical light sampling techniques. Since the beginning of CG development, lighting and rendering have been the focal point of researchers [Kajiya 1986; Cook et al. 1987] who have been trying to accurately estimate the behaviour of an infinite number of light rays being emitted from light sources, interacting with the environment, and eventually, some will be reflected into the camera.

As computer animation is essentially a virtual film production, the light models usually employed in a CG scene are emulations of the physical cine lights [Birn 2000] such as spotlights and area lights (the equivalent of softboxes in photography). The number of lights and their placements or a lighting design is inevitably influenced by the artistry of past cinematographers. The adaptation of cinematic lighting in CG character lighting will be discussed in detail throughout the thesis. The quality of light and shadow or its realism depends on the sampling technique and rendering methods. There are two distinct approaches of rendering developed side-by-side for different applications and requirements, ray-tracer for slow but realistic renders, and rasteriser for fast but less-realistic renders [Marschner and Shirley 2016].

A *ray-tracer*, in an overly simplified explanation, shoots a ray from each pixel in the direction derived from the specification of the camera. Each ray travels through the scene—the collection of virtual objects hits an object, bounces to a light source or other objects, picks up the colour calculated by a surface shader at each hit location, and returns the colour (averaged colour in the case of multiple rays) to the source pixel. This process is the reverse of how a camera works in the real world—light travelling from a light source to an object and reflecting into the camera. *Global illumination* is a generic term describing a render that considers both direct and indirect illumination from all the light transporting through the scene [Dutre et al. 2018]. A physically-based global illumination can be achieved by a complex *path-tracer*, a multi-ray-multi-bounce ray-tracer that produces subtle lighting effects, such as soft shadow and colour bleeding (a colour of an object reflects onto other objects nearby). Rendering a high-quality photorealistic image with global illumination using a path-tracer often requires a large amount of computational power and time such that a frame in an animated film could take up to several hours to render.

On the other hand, A *rasterizer* is the other major render scheme popular in games and interactive applications. Each object in the scene is broken down into triangles and squashed into the screen with transformation functions. The triangles closer to the screen will be fully visible, and the triangles behind will be blocked or partially visible. Carrying the information (positions, normals, colours, etc.) necessary for shading computation, the flattened triangles are later disintegrated into small fragments around the size of one pixel (triangle rasterisation) and fed into an associated shader to compute the final pixel colour. This technique could achieve incredibly high frame rates with modern graphics cards designed specifically for triangle processing [Akenine-Möller et al. 2019]. However, the images rendered this way are not physically corrected, especially in terms of lighting, for example, accurate light bounces and area lights are difficult or too expensive computationally for a rasteriser to perform and often cheated with tricks to attain artificial global illumination.

In this study, we rendered our stimuli with a path-tracer to obtain the state of the art image quality, close to the animation industry standards for animated films, except for the last chapter, in which we experimented with an advanced real-time rasteriser for faster turnaround time while still producing good quality renders.

#### 2.2.4 Visual Perception of CG Scenes

Most of the recent research in computer graphics and animation focuses on the development of realistic models of the world, such as global illumination [Ritschel et al. 2012], unbiased rendering [Jensen 2001], and physically-based dynamics [Baraff and Witkin 1997]. However, these are often simplified computational models of physical laws that do not consider the human perception of the world. The parameters of the computer graphics algorithm and the produced perceptual effects are tricky to form a meaning relationship and make use in graphics, visualisation, or art. Cunningham et al. [2007] attempted to reparametrize complex reflectance models into a perceptually uniform parameter space. They rendered images of an object with different reflectance parameters and collected pairwise similarity ratings among the renders, and applied the multidimensional scaling (MDS) analysis to arrive at the fundamental perceptual dimensions and the location of each image in the perceptual space. Another extreme example of the difference between "physically-based" and "perceptually-based" models was provided by [Khan et al. 2006] in which the authors demonstrated the so-called *bas-relief ambiguity* [Belhumeur et al. 1999], showing that in



order to change the appearance of the material dramatically, one does not need to re-render three-dimensional model with a new shader (with a difference reflectance function), but instead, simply manipulate the image statistics using a 2D filter. Other work in the field of computer graphics has used psychophysics to propose novel perceptual models for computer generated objects. Examples include material reflectance [Vangorp et al. 2007], local adaptation [Vangorp et al. 2015], and more recently material appearance [Lagunas et al. 2019]. The next section went into more details of many virtual perception studies directly related to this thesis.

## 2.3 Visual Perception

In the previous section, we discussed the 3D animation production pipeline, the method we employed to synthesise a stimulus and explored the past literature related to our research in the aspect of stimuli creation. In this section, we provided a brief overview of the mechanism of how these stimuli are perceived by the human visual system.

When we move about in the world, our brain constructs a mental model of how the world should work. The sensation stimulated by the physical world (sight, touch, sound, smell, and taste), through a sense organ, gets interpreted and given a meaning that fits in the mental world by our brain. The process of organizing, identifying, and interpreting the sensation to create a mental representation is called *perception* [Schacter et al. 2011], which comes from Latin meaning apprehension with the mind [Dictionary 1989].

Perception is a vast interdisciplinary subject, but only a small subset of related topics in visual perception will be presented in this chapter. As perception is understanding the world via the interpretation of the sensory information, *visual perception* is acquiring the knowledge of the surroundings through the visible light. The main objective of visual perception is to allow one to plan and act appropriately [Yantis 2001]. The various components of visual perception, from the eyes to the brain, are referred to as the *visual system*.

The visual system is the information pathway from our eyes to our brain that enables us to understand the physical environment. It takes visible light as the input and returns experience or actions as the output. Seeing begins when light travels from the outside world through the eye lens, and then gets focused on to the retina, a light-sensitive receptor, in the back of the eye. Different types of photoreceptive cells of retina are responsive to different light information, for example, *cones* are adapted for colour vision, daytime vision, and

detailed vision, while *rods* are adapted for vision in dim light. [Kalat 2007]. These photoreceptors transduce photons into distinct neural signals that get transmitted to and processed by different parts of the neural networks in the primary and secondary visual cortex. There are approximately 30 to 50 areas in the brain dedicated to vision [Schacter et al. 2011].

The visual process uses a set of *visual cues* to extract the properties of surfaces or environments. Each cue can infer a property by recognising a particular visual pattern [Thompson et al. 2016]. Consider watching a movie on a screen, although the moving picture of an actor is two-dimensional as the screen is flat, the movement of various facial muscle groups (motion cues) and the head itself, and the shade of the skin (pictorial cues) due to illumination are processed to construct the three-dimensional representation of the actor's face and expression in our brain. One of the challenges in recovering surface properties is the information imaged on the retina is a compound of multiple physical properties. For instance, brightness is a function of both reflectance and illumination, and it is also influenced by the relative brightness of the environment.

However, the human visual system has developed an incredible ability to account for contextual information, illumination, object geometry, material properties, and other characteristics, as well as the sophisticated capability to isolate specific visual cues of interest, such as the reflectance or albedo of a surface [Gilchrist 2013]. One key finding demonstrating this ability to separate sources that cause changes in the brightness, as mentioned above, is *lightness constancy*. In the previous section, we discussed the 3D animation production pipeline, the method we employed to synthesise a stimulus and explored the past literature related to our research in the aspect of stimuli creation. In this section, we will provide a brief overview of the mechanism of how these stimuli are perceived by the human visual system.

### 2.3.1 Psychophysics of Lighting

Light is physical, and the law of physics is universally deterministic, for example, the light speed in a vacuum is a constant of 300,000 metres per second, anywhere in the universe. However, the basic perception of sight may vary from one person to another. Changes in the properties of light, as an electromagnetic wave, such as amplitude (intensity) and frequency (colour), can be measured precisely, even for a small amount. However, our brain tends to

ignore subtle differences in the sensory information, and the perception remains consistent. This perceptual principle is known as *constancy* [Schacter et al. 2011].

A black cat always looks the same shade of black, either it is lying in a room illuminated by a 100-watt incandescence light, or walking outside under the bright noon sunlight, even though in the latter case more light is reflected from the cat to our eyes. The *lightness* or the perception of achromatic surface (white, grey and black) remains the same, and *lightness constancy* is when we tend to perceive the lightness of an object is unchanged under different illuminations [Goldstein 2009]. The explanation of this phenomenon has been a major challenge in visual science [Brainard 2003; Gilchrist et al. 1999].

In order to scientifically investigate the human visual system and lightness constancy, objective measurement is required. Measuring the properties of light such as intensity and colour is straight forward, but quantifying a person's subjective perception of that light can be problematic. In the mid-19th century, a German scientist, Gustav Fechner, pioneered a framework to measure sensation and perception called *psychophysics*. In a typical psychophysical experiment, a researcher measures the strength of the stimulus and the observer's sensitivity to that stimulus. The psychophysicist later derives the relationship between the stimulus and the observer's responses [Fechner et al. 1966].

In psychophysics, *lightness* is the achromatic perception of an object's albedo, and *brightness* is an achromatic perception of the strength of light reaching our eyes, the combination of the illumination, the light *intensity* reaching the object surface and the object's *reflectance*, the proportion of light intensity allowed to bounce off the object's surface.

Aiming to understand how humans perceive lightness, there have been many attempts to map the ability of viewers to discriminate lightness levels in the empirical work in psychophysics [Stevens 1957; Fechner et al. 1966]. Although humans are relatively good at discriminating between different lighting intensities, the perception of a surface's brightness is thought to be dependent almost exclusively on the surface properties of an illuminated object. Indeed, as Kardos [1934] pointed out, people tend not to include shading and shadows when describing a scene. In order to maintain the characteristics of an observed object, the human visual system is highly skilled at accounting for contextual information as well as surface characteristics. For example, a box that is black is seen as black regardless of how bright the room that contains the box is. This is called *lightness constancy*, and it is still a puzzle in visual science research [Gilchrist et al. 1999; Brainard 2003].

Specific methods to measure this phenomenon have been developed, such as asymmetric matching and apparent dissimilarity of the surfaces (see Logvinenko and Maloney [2006]). Lightness constancy is particularly vital in preserving object identity when part of the object is not illuminated, i.e. in shadow. Shadows and shading are essential in shape information acquisition [Mingolla 1983] and have long been utilised in vision and computer graphics to present or retrieve the three-dimensional shape of an object [Bruckstein 1988].

Another related topic to lightness constancy is *colour constancy*—the perceived colour of a surface stays more or less constant under variations of light intensity of the spectral component component [Foster 2011; Ebner 2007]. The perception of colour was not a focused factor in this thesis but we had to use colour stimuli in most experiments so our results would be applicable in the industry. We only used white light and assigned the same set of materials to each character across experiments to minimise the effect of colour and colour constancy.

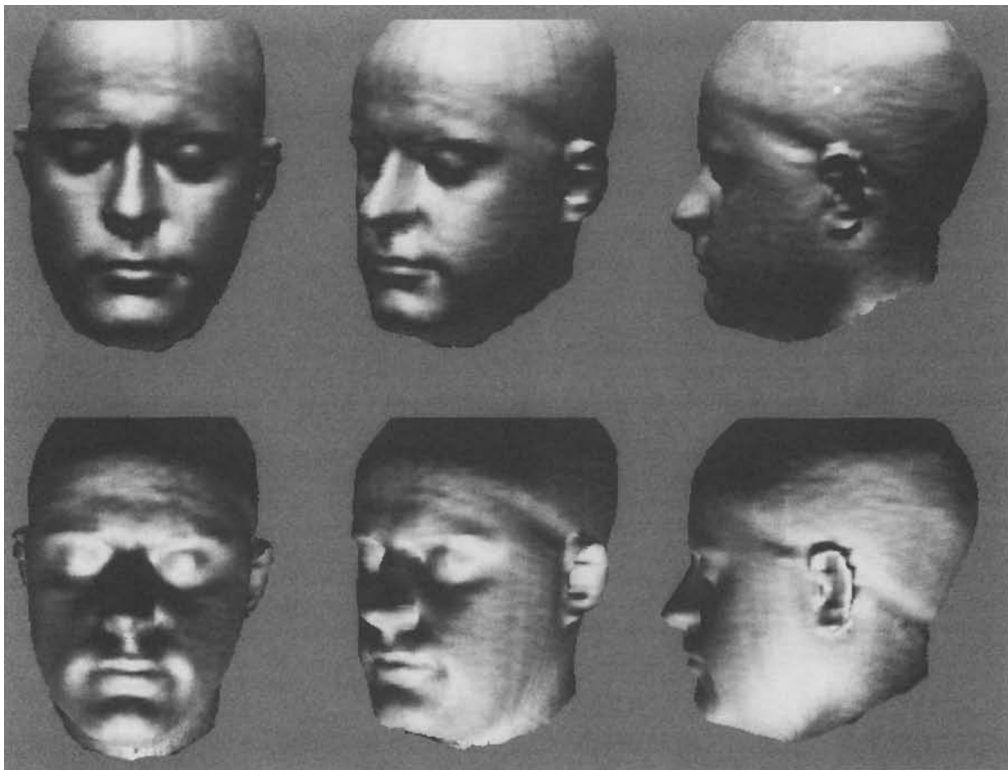
In summary, lighting perception is the intricate relationship between changes in physical lighting and specifics of the human visual system adapting to attempting to preserve the consistency of the object properties.

### 2.3.2 Shading, Shadows and Light Direction

Besides illuminating the scene, light or dark, another important role of lighting is shading the in-between that helps reveals the 3D shape from a 2D image. Recall the *chiaroscuro* (Section 2.1.1), the gradations of light and shadow created by the interaction between light and surfaces. Our brain can perceive underlying three-dimensional forms of objects in an image from the *chiaroscuro*, and the process is known as *shape from shading*. Nonetheless, how the brain recovers shape from shading and what the relevant visual cues remain challenging questions [Thompson et al. 2016]. Computationally, solving shape from shading is under-constrained, which means analysing the image data alone cannot uniquely determine the underlying geometry. The illumination information and the material properties are also required for a better estimate of the three-dimensional form. However, other ambiguities, such as the *bas-relief*, could also present in the image and make the precise shape recovery impossible.

In contrast to computing shape of shading, the human visual system is incredible at telling the 3D shape from just looking at an image based on the prior experience. Although light

intensity is mostly discounted in the perception of a surface's albedo, it is still processed and used for additional perceptual tasks. For example, changes in brightness across the surface of an object that is due to illumination and not the texture, i.e., shadows, are an essential factor in acquiring shape information [Mingolla 1983] and have long been used in computer graphics and vision to retrieve or present the 3D shape of an object [Bruckstein 1988]. There is no doubt that the perception of object properties such as shape and material varies as a function of light [Koenderink et al. 1996; Zhang et al. 2018; Zhang et al. 2019]. It is also worth noting that various lighting conditions affect the perception of shadow differently [Van Nes and Bouman 1967; Peli et al. 1991; Pamir and Boyaci 2016]. Shadows can be used to predict the shape and light direction. Alternatively, our brain can assume light coming from above and uses the information together with the shading pattern to recover the shape [Morgenstern et al. 2014; Morgenstern et al. 2011].



**Figure 2.4:** Facial shapes illuminated from above and below used in Hill and Bruce [1996]'s experiments

There were a number of early studies on lighting and face perception investigating the effect of direction and shadows on facial recognition. For example, Johnston et al. [1992] suggested that lighting a facial surface from below disrupted the recognition. Hill and Bruce [1996] later confirmed that participants performed recognition and matching tasks

more accurately when the facial shapes were illuminated from above than below (Figure 2.4). The link between, shading, shadows, light direction and (facial) shape motivated us to test the influence of contrast (soft vs. hard shadows) and light direction in our first experiment (Chapter 3).

### 2.3.3 Movement

Our eyes never stay still, and the world around us is always moving. Inevitably, our vision continually shifts. All this movement, even the retinal motion, provide us with useful visual information about how we move through the world, the geometry (e.g. shape, size and distance) of the objects we are passing by, also known as *structure of motion* [Gibson 1950; Ullman 1979].

The fact that the majority of perceptual studies of lighting have focused on still images is surprising since motion is an essential aspect of natural visual scenes. One study on colour constancy [Werner 2007] showed that the synergistic integration of colour and motion signals is an important mechanism for improving colour identification. Therefore, with added motion, colour constancy improves. More recently, perception of material properties have been studied on dynamic stimuli such as cloth [Bi and Xiao 2016], liquids [Assen and Fleming 2016], and optical flow characteristics [Doerschner et al. 2011].

### 2.3.4 Emotions

There exist several definitions of emotions in the field of psychology and neuroscience. Early studies of emotion believed our body responded a certain way to a specific stimulus, as stated by Frijda et al. “Input some event with its particular meaning; out comes an emotion of a particular kind” [Frijda 1988]. The assumption that emotional responses are universal or natural kinds have shaped the agenda of scientific studies of emotion that motivate contemporary researchers to look for observable common patterns of responses in face, voice and body such as Tomkins [1962]’s nine primary affects, Plutchick [1980]’s wheel of emotions, or Ekman [1992]’s six basic emotions.

In this thesis, we tested the emotions based on the Ekman’s classification that identifies the emotions that are universally recognised. Based on the visual similarity of physical facial expressions, he classified emotions into anger, sadness, surprise, disgust, fear and happiness. The classical view of emotions being “natural kinds” or hardwired in human



**Figure 2.5:** Characters from Pixar’s *Inside Out* representing basic emotions categorised by Ekman: disgust, fear, joy (happiness), sadness and anger. Note that surprise was not included in the movie.

DNA, such as the ones described by Ekman, has been heavily challenged by new evidences [Russell 1994; Barrett 2006]. However, we felt justified in using them since the observable static and dynamic facial features of Ekman’s system of emotion identification were simple to implement in the modelling and animation of our CG stimuli, plus they were widely accepted by the animation industry such as the ones seen in Pixar’s animated feature *Inside Out* (Figure 2.5.)

For stylised characters, the expression of emotion can be additionally manipulated by exaggerating the motion of the character [Lasseter 1987; Thompson et al. 2016]. This is a widely practised technique, designed to make the emotional expression of the character more salient [Lasseter 1987] and more appealing [Hyde et al. 2013]. Stylised appearance also plays an important role in emotion recognition and perceived intensity of the expression. In the study of Wallraven et al. [2007] the expressions of highly stylised characters (created with the brush-stroke method) were not only less recognisable than other stylisation techniques but also rated as less sincere and intense due to the noise of the outlines as the character was moving. According to this study, the results showed that while stylisation is often a preferred design choice which provides subjective certainty about the conveyed expression, high abstraction was found to hinder the recognition of facial expression and resulted in an unfavourable response to the character.

### 2.3.5 Appeal

In many face perception studies, the word *appeal* is used interchangeably with *attractiveness*, for instance, the attractiveness of a face could be defined by the appealing characteristics of facial features and the spatial relationships among them [Luo et al. 2011; Chin et al. 2006],

which focuses more on the physical appearance. However, in *Disney Animation: The Illusion of Life*, Thomas and Johnston [1995] suggested that *appeal*, one of the *twelve basic principles of animation* refers to “anything that a person likes to see,” rather than just being good looking. An appealing character does not need to be sympathetic; for example, villains can be appealing if they have the charm to attract our gaze although they appear ugly and repulsive. Thomas and Johnston’s definition of appeal involves a cognitive process, similar to the *artistic beauty* [Schulz and Hayn-Leichsenring 2017]. The perceptual process of appeal assessment is complex and quick. Evidently, people can judge the attractiveness of a person’s face at a glance, even when the face is not consciously perceived [Olson and Marshuetz 2005]. Due to this fact, appeal is crucial in engaging a viewer because the viewer only needs a fraction of time to judge a face and make a decision to pay attention or ignore. Once the appraisal has been done, it will rarely change with more time [Willis and Todorov 2006]. Therefore, more appealing virtual avatars means better chances to attract the attention of the audience and engage them.

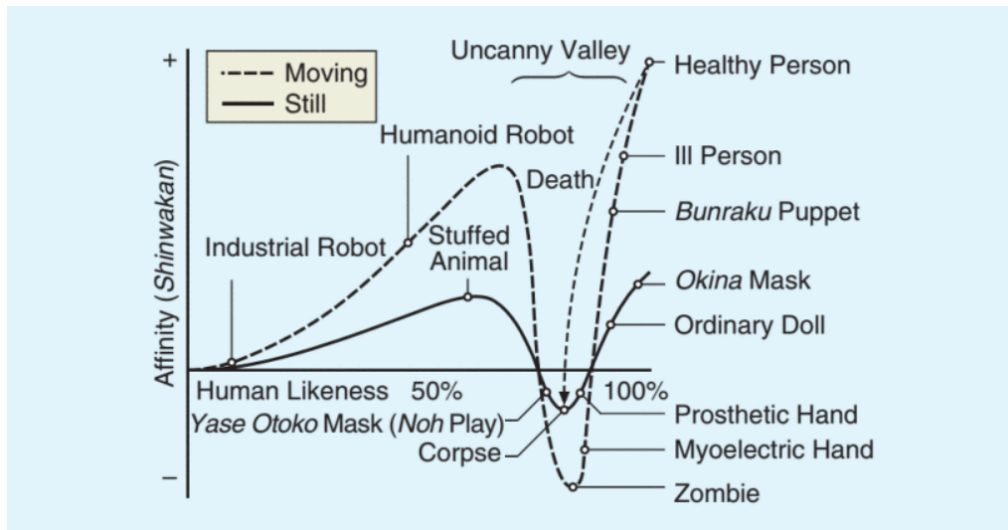
In computer graphics, there were attempts to quantify attractiveness of digital avatars. A recent work investigating into an appeal of a character’s face was done by Kokkinara and McDonnell [2015], in which they found more realistic facial motion capture data can increase the perceived appeal of the animated virtual face. In their work, “High appeal rating means that the virtual face is one that you would like to watch more of, and you would be captivated by a movie with that face on a character as the lead actor,” and suggested the judgement should be based on both visual and motion cues. This explanation of appeal rating is comparable to the one given by Thomas and Johnston and will be followed by this thesis.

Many of the recent character perception studies have linked stylisation to the overall appeal of virtual avatars, such as the render style investigation by McDonnell et al. [2012]. In their work, the creation of cartoon-like appearance was achieved by changing the character’s render technique. While certain styles of cartoon rendering have been found less appealing than others when assessed from still images, these differences were even more apparent when the character was moving. Stylisation and motion were found to be important factors of appeal. Zell et al. [2015] tested appeal and emotion perception of virtual characters from still images. The changes in realism were obtained by changing the geometry (exaggerated features, such as enlarged eyes, big nose, etc.) and materials (scanned texture from a real actor and artist created textures) of the character. Extreme



mismatches of material and shape had a negative effect on the perceived appeal, and cartoon shapes increased the reported intensity of specific emotions, especially for happy, sad and surprise. This study also examined different lighting conditions but found few effects. More closely related to our work, Lotman [2016] conducted a pilot study where they tested just two different lighting conditions (low and high contrast) on live-action footage of an actor's face, under three different emotional expressions. They found some evidence of lighting and emotional interactions, but not for the happy and angry expressions, only for disgust.

From the aforementioned previous work, we identified both the emotional state and appeal to be the most important aspects of character perception to investigate in this thesis, as these are central to storytelling, and artists use cinematic lighting extensively to enhance these aspects in order to captivate audiences [Calahan 2000]. There are some indications that emotional valence could be related to brightness in the field of Psychology. While some studies focused on how brightness can affect the viewer's emotions [Zhang et al. 2016], one study examined the perceptual effect of smiling vs. frowning faces and found smiles to be perceived brighter than frowns [Song et al. 2012].



**Figure 2.6:** A graph from [Mori et al. 2012] illustrating the Uncanny Valley effect of still and moving stimuli.

### 2.3.6 Uncanny Valley

While a positive response can be elicited by an attractive CG face, a negative response can also be evoked by the same face if it looks too realistic [Mori et al. 2012; Seyama and

Nagayama 2007]. According to evolutionary aesthetics, people favour facial features that are unique to their cultures [Cunningham et al. 1995] and hence could develop repulsiveness towards artificial faces because humans did not evolve with robots or 3D models [Rhodes et al. 2001]. The phenomena that virtual avatars may come across as creepy is known as the *Uncanny Valley Effect* [Mori 1970]. Although the eeriness of a humanlike figure is typically associated with photorealism—as its appearance approaches realism, its perceived affinity drops. There were studies investigating factors that could contribute to the perceived unfamiliarity of CG faces [MacDorman et al. 2009; Kätsyri et al. 2015]. We have already mentioned a few in Section 2.2.1 e.g. different render styles [McDonnell et al. 2012] and the mismatch in material and shape [Zell et al. 2015]. The question of eeriness was included in our study when we evaluated the effect of lighting on realistic characters in Chapter 5, Section 5.3. The original uncanny valley effect study by Mori et al. [2012] also pointed out the effect of movement raising the peaks and lowering valleys of perceived affinity of a human-like entity (Figure 2.6).

### 2.3.7 Language and Culture

Perceptual experiences are often associated with the abstract concept of affect. The fact that we have so many metaphors relating brightness to good and darkness to bad has been shown in psychology research to be results of developmental experiences which pair these factors, for example, social and physical rewards being more prevalent during the daytime while the darkness of night hides potential dangers [Landau et al. 2010]). A recent study investigating cross-cultural figurative language and emotions by Barchard et al. [2017] has confirmed that the dichotomous descriptor “bright” is associated with happiness and “dark” is associated with sadness, anger and fear (all rated above 87% by the participants). These studies imply that there is a semantic association between emotions and brightness, supported by another study using a Stroop-like test where positivity and brightness were found to be associated [Meier et al. 2004]. In addition, in [Xu and Labroo 2014], the authors found that bright light increases people’s perception of heat, which in turn activates their hot emotional system, leading to intensified affective reactions—positive and negative—to different kinds of stimuli.

From an anthropological standpoint, there is evidence, across cultures and times, that insight, health, optimism and virtue are commonly represented by brightness or light in local languages or folklore, whereas evil, danger and death are represented by darkness [Meier

and Robinson 2005]. The effect is also common in popular culture, either symbolically, e.g., in *Star Wars*, evil is paired with darkness and good with light, or illustratively, e.g., heroes dress in white and villains dress in black [Meier and Robinson 2005]. This is unsurprising as the *Jedi* is more or less a religious institute, good vs. evil and light vs. darkness are universal values in multiple distinct cultures and religions. In [Adams and Osgood 1973], participants from 20 countries with different religious beliefs were instructed to evaluate black, grey, red, yellow, blue, green and white colours. White was perceived as positive, and black was perceived as negative for all 20 countries.

## 2.4 Perceptual Experiment Design and Analysis

Early psychophysical and perceptual experiments require absolute control over as many variables as possible in order to infer the non-observable attributes of a perception process from the observable stimuli. Otherwise, the result can be difficult to model or explain. The experimenter needs to design the experiment carefully, what is shown, how it is shown, and to whom it is shown, so the outcome is uniquely interpretable. As a result, it is common to see perceptual studies have traditionally been conducted with abstract stimuli (e.g. simple lines or solid colour chips), because they can be 1) reproduced with precision, 2) systematically varied, and 3) described mathematically for both the variations and the overall relationship. It is crucial that participants only see the information we present to them. Hence, in traditional psychophysical experiments, the stimuli are shown in the completely dark room. The participant's head is required to maintain a fixed position (e.g. held in place using a chin rest) to ensure the consistent retinal image. Classic experiments were also conducted with thousands of repetitions for stable averaged results. Meeting all the above mentioned requirements was possible only via a custom-made machine until the recent advancement and improved reliability of computer and display technologies. Computer Graphics and computer-generated stimuli have become a new norm in perception experiment [Bartz et al. 2008].

In modern perception experiments, the head stabilisation and an excessive number of repetitions are no longer standard practices. Although near-absolute control over the experiment factors are still highly desired and many researchers continue to use simple stimuli, it is not conclusive if the results of such studies generalise to the more complex stimuli in the real world [Gibson 2014]. The perception studies of complex stimuli such as

the character lighting design will require to carefully adapt the classic methodology and systematically adjust the variables in order to address the high dimensionality of the stimuli. One of the chosen methods for this thesis work is multidimensional scaling analysis which was inspired by a low-level perception study of lightness [Logvinenko and Maloney 2006].

### 2.4.1 Multidimensional Scaling (MDS)

MDS is a class of algorithms that for a set of  $N$  data points, takes the number of dimensions ( $d$ ) and an  $N \times N$  dissimilarity matrix, of which element  $e_{i,j}$  represents the dissimilarity distance between point  $i$  and  $j$ , as input, and returns the coordinates of the  $N$  points in the specified  $d$ -dimensional space that best respect the dissimilarity matrix [Cunningham and Wallraven 2011; Cox and Cox 2001].

Most perceptual methods for determining the functional relationship between stimulus and perception require the experimenter to know the perceptual dimension or dimensions that are being studied. Richardson [1938] introduced a new family of methods, referred to as MDS, which is a data reduction technique, similar to *Principle Component Analysis (PCA)*. From a set of simple pairwise similarity ratings, MDS directly determines 1) the essential number of perceptual dimensions, typically lower than the actual number of the stimulus parameters, and 2) the proximity structure or the relative dissimilarity distance among the stimuli, scaled in the reduced perceptual space. MDS assumes Euclidean distances. Since we use Likert scales, it is not clear that our similarity structure is, in fact, Euclidean. Thus, we used a non-metric MDS which uses the ordinal relationships to obtain scaled proximities.

MDS has been utilised in many face [Sergent 1984; Papesh and Goldinger 2010] and lighting [Ramanarayanan et al. 2008; Tokunaga and Logvinenko 2010] studies. Further details of an impressive variety of the implementations and applications of MDS in other stimulus classes could be found in [Torgerson 1952] and [Schiffman et al. 1981].

## 2.5 Conclusion

This chapter provided the background to our work from the industry and visual perception standpoints. Many practices in art were based solely on observations, and have continuously evolved over hundreds of years. The process could be considered scientific due to the use of trial and error, that helped converge many lighting techniques into a handful of rules of thumb. On the other hand, utilising psychophysical experiment methods, scientists have

developed psychological theories explaining the perception of seeing and feeling light. The conclusions from the two approaches agree and do not agree on many aspects that are worth investigating. Learning from the previous work in both fields equipped us with the tools to create a comprehensive set of stimuli and experiments that explore the various facets of our research question. We hope that the new insights derived from our work will merge the two seemingly parallel approaches of illumination design, resulting in universal guidelines for impactful character lighting.



## Chapter 3

# Lighting Parameter Investigation

### 3.1 Introduction

The design of lighting in Computer Graphics is directly derived from cinematography, and many digital artists follow the conventional wisdom on how lighting is set up to convey drama, appeal, or emotion. In this chapter <sup>1</sup>, we conducted an initial broad study investigating the most used shading and lighting conditions in CGI.

Most CG artists closely follow the language of cinematography and have led to the development of rendering techniques that mimic lighting in the film. Meanwhile, there is also a significant fraction of digital artists that take inspiration from 2D storytelling arts such as comics and manga and focus on stylisation to achieve results closer to hand-drawn animation, intending to enhance emotional expression [McCloud1994]. Toon-shading is a common technique among artists in the latter group, and it is still popular, particularly in the Japanese entertainment industry.

Despite the difference in render styles, both groups have a strong belief in the connection between lighting design and an audiences' emotional response. Since the days of Leonardo da Vinci, known for sfumato or low-key lighting, and Caravaggio, known for tenebrism or high-key lighting, (see Section 2.1), certain assumptions of how lighting is perceived in art have been made through years of observations and trial & error. However, a recent study by Poland [Poland 2015] examined the influence of different contrast lighting designs (high-key vs. low-key) on the perception of a short film and surprisingly found her results contradicting with the conventional cinematic lighting design instructions.

---

<sup>1</sup>The content of this chapter was published in the ACM Symposium on Applied Perception 2016 (SAP'16)

Another lighting design technique commonly utilised in art to emphasise certain inner states of a character is *direction*. Light coming from below is considered unnatural, as opposed to sunlight or a lamppost illuminating from above, and often associated with negative meanings, for example, in stage lighting, a sinister character is often illuminated from the stage floor to emphasise evil intention or rage [Reid 2013; Wolf and Block 2014]. Although there were many studies on the perception of light direction, most of them concerned the shape from shading aspect or facial recognition [Hill and Bruce 1996].

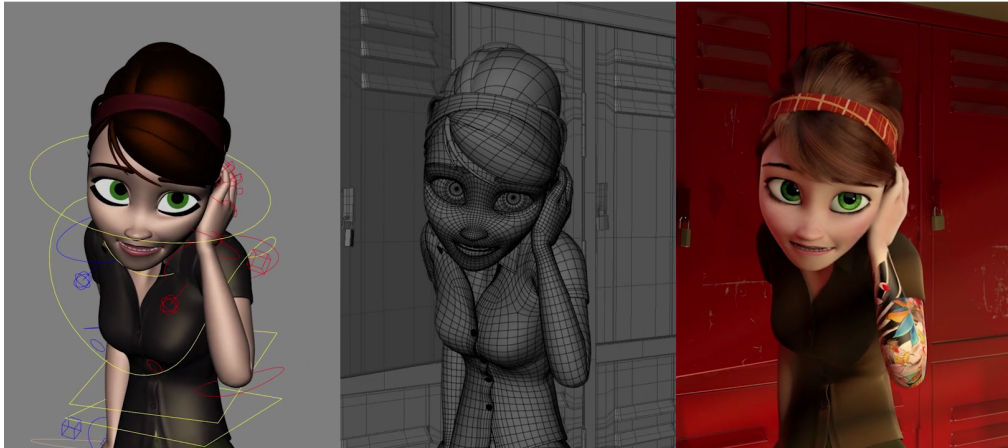
To the best of our knowledge, our study was the first to empirically determine the effects of shading and lighting conditions on our perception of animated virtual characters. We chose to examine the perceived emotion and appeal because they are fundamental to the human visual system from the evolutionary standpoint [Plutchik 2001] and crucial to the audience's engagement [Calahan 2000].

For this work, a professional animator was commissioned to create a sequence of dramatic emotional sentences for a typical CG cartoon character. The animated character was then rendered using a range of lighting directions, contrast levels, and shading techniques. Finally, the renders were shown to the participants who were asked to rate these animation clips on various aspects relating to the appeal and perception of emotion. This chapter represents a starting point for assessing the parameters of interest for the study of appeal and emotion of cartoon characters. We will use the results from this study to guide the choice of parameters for a more thorough investigation in later chapters.

## 3.2 Stimuli

In performing arts, properties of light are usually categorised into distribution, intensity, colour, and movement [Wolf and Block 2014]. Since computer animation inherits many lighting principles from theatre and film [Birn 2000; Calahan 2000], we will also adopt these lighting properties in our study. However, in order to minimise the length of the experiment (and fatigue of participants), we decided to reduce the number of variables and mainly focus on the distribution property which includes quality and direction (of the key-light). These lighting conditions will be applied to both *CG-shaded* (smooth shaded) and *toon-shaded* stimuli.





**Figure 3.1:** Frames showing the rig, mesh and a final render of the original Mery character

### 3.2.1 Character

As a character model, we chose Mery<sup>2</sup>, a female character with a typical cartoon appearance commonly seen in animated movies (Figure 3.1). This character has a highly controllable facial rig, with the ability to squash and stretch each part, allowing animators to create highly expressive cartoon animations.

### 3.2.2 Emotions

<b>Emotion</b>	<i>Affective sentence</i>
Anger	<i>I'm sick of you being late.</i>
Sadness	<i>My best friends is moving away.</i>
Disgust	<i>Aww, I'll never eat those noodles again.</i>
Fear	<i>Someone is following me.</i>
Happiness	<i>It's a beautiful day outside.</i>

**Table 3.1:** Validated affective sentences for anger, sadness, disgust, fear and happiness

We commissioned a set of animations from a professional animator which convey the six basic emotions: Anger, Happiness, Sadness, Fear, Surprise and Disgust, according to Ekman's classification [Ekman 1992]. For the dialogue, we provided the animator with audio replicating a previously validated list of affective sentences for spoken emotion identification [Martin et al. 2006; Ben-David et al. 2011] (Table 3.1). Each sentence lasted

<sup>2</sup><http://www.meryproject.com>

approximately 3 or 4 seconds. In the experiment, we chose not to play the audio, in order that participants could focus specifically on the appearance of the character when making their judgements and to ensure emotion recognition was not too easy due to the content of the audio track. The animation was created in Maya<sup>3</sup> 2016 digital content creation (DCC) software.

### 3.2.3 Lighting Parameters

In terms of quality, there are a diverse range of lighting styles across different disciplines [Alton 1995; Millerson 1991; Wolf and Block 2014] but we decided to broadly generalise them into two groups: low contrast and high contrast lighting. *Low contrast (Lo)* represents “soft light”—a large area light that produces subtle transitions from light to dark on the illuminated area and soft edge shadow [Lowell 1992; Millerson 1991], and also includes the middle ground between “gradated tonality” and “high-key” lighting [Malkiewicz and Mullen 2005]. *High contrast (Hi)* lighting represents “hard light”—a small but intense light source that produces hard-edge shadows, and also covers a subset of “low-key” lighting. To achieve the low and high contrast (first and second-to-last rows of Figures 3.3 and 3.4) in CG-shaded stimuli, we used an area light and a point light. However, to avoid the bias of overall brightness—considered one of the controlled variables, we could not go too extreme on either end—for example, we cannot produce true high-key or low-key lighting (Figure 3.2).



**Figure 3.2:** Frames from *Alfred Hitchcock's Rear Window (1954)* showing high-key (left) and low-key (right) lighting

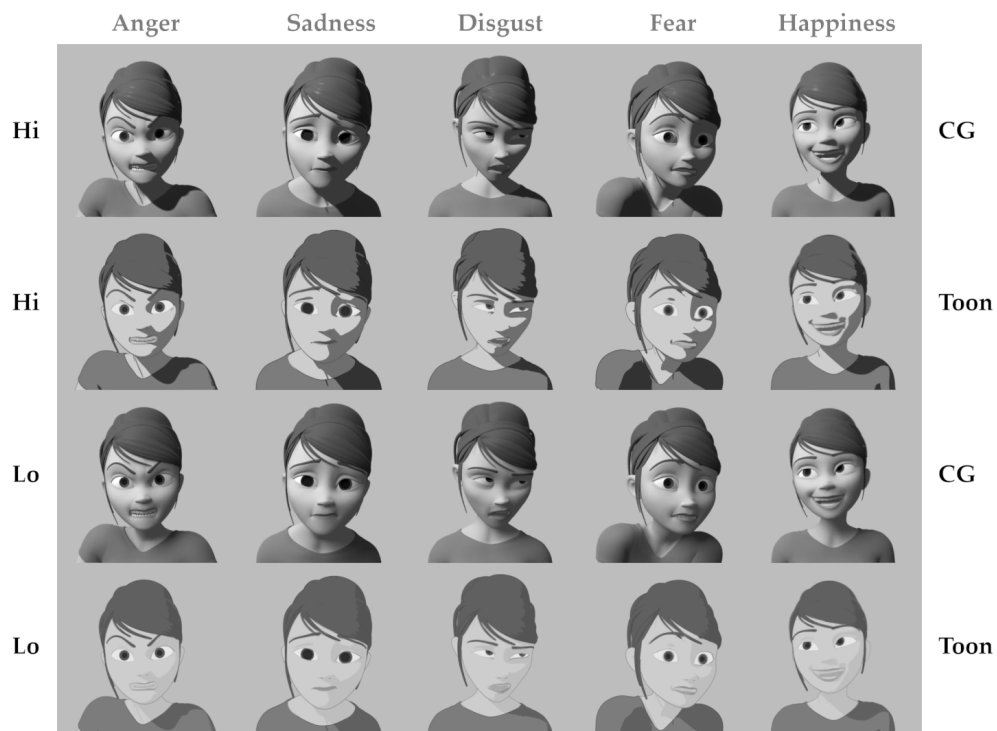
**Shading:** The animated character was exported to an Alembic<sup>4</sup> geometry cache file to be shaded, lit and rendered in Houdini<sup>5</sup> 15 DCC software. For the CG-shaded stimuli, we used

<sup>3</sup><https://www.autodesk.eu/products/maya/>

<sup>4</sup><http://www.alembic.io>

<sup>5</sup><https://www.sidefx.com/products/houdini/>

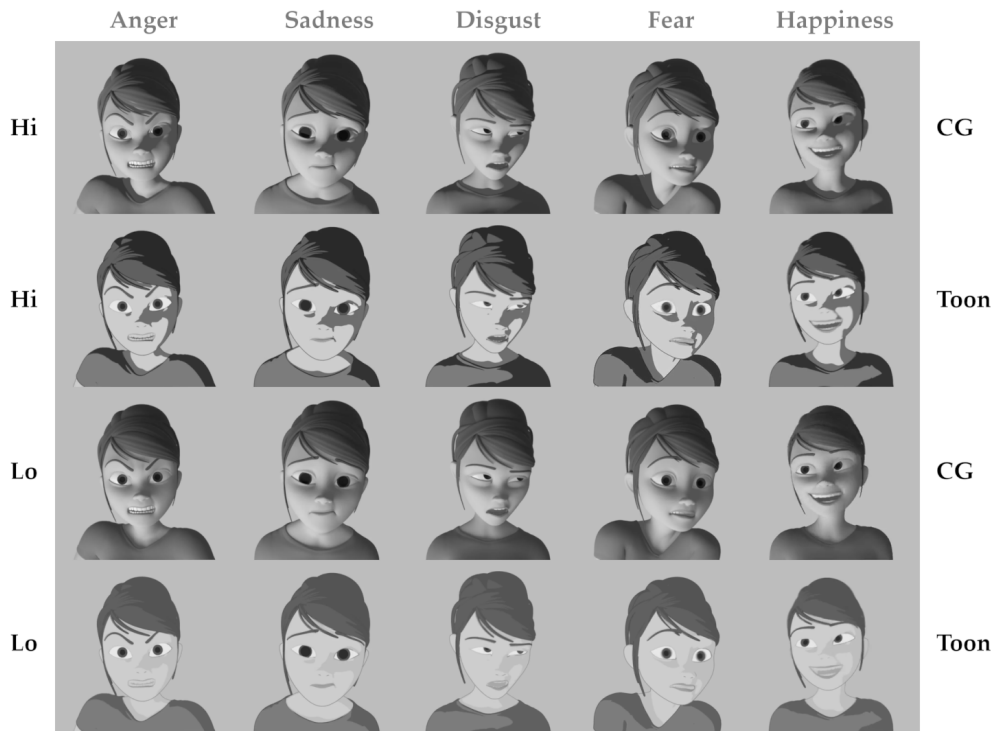
the Houdini principle shader, and for the toon-shaded stimuli, we wrote a custom shader based on the technique from Lake et al. [2000] in which the area light would have no effect (this was also a reason why we did not generalise the light quality into soft and hard light). Instead, a point light was used in both low and high contrast scenarios. The low and high contrast conditions were achieved by using different darkness levels for the shadowed area. We chose to use a greyscale rendering in this initial study to control for potential colour bias, as previous work [Seifi et al. 2012] showed that certain perceived emotions were enhanced when the colour matched the expression in painterly rendered images of human faces. greyscale was also easier to control in terms of matching the overall appearance of different stylisations to each other. In later chapters, we rendered Mery in colour for ecological validity.



**Figure 3.3:** Frames showing Mery emotions rendered in the combination of “light from above” (Ab), “high contrast” (Hi) / “low contrast” (Lo), and CG-shaded / Toon-shaded conditions

**Direction:** With regard to direction, we will investigate a light source coming from above (Ab) and below (Be) (Figures 3.3 and 3.4). Light coming from above or “motivated light” is a natural direction motivated by real world sources such as the sun or a ceiling lamp, hence the word “motivated” [Wolf and Block 2014]. Lighting coming from below or “unmotivated light” is commonly used to add dramatic effect, particularly in stage lighting, as the light

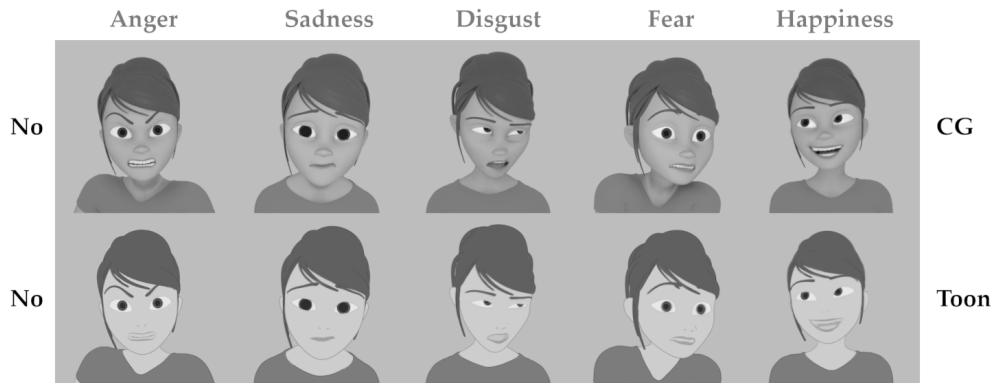
comes from an unnatural source such as fire light [Wolf and Block 2014; Gurney 2010]. To create the light direction, a key-light is placed in front of the character, 45 degree to the right, and 45 degree above (or below) the eye level [Calahan 2000; Gurney 2010].



**Figure 3.4:** Frames showing Mery emotions rendered in the combination of “light from below” (Be), “high contrast” (Hi) / “low contrast” (Lo), and CG-shaded / Toon-shaded conditions

Please note that there are countless possible light directions but we only use this setup, as it is a common key-light position, such as in the well-known three-point lighting or three-quarter lighting methods [Gurney 2010]. A key-light is often coupled with a fill-light to slightly brighten up the shadow and can be artistically placed anywhere. In the CG-shaded stimuli, we decided to use a low intensity dome light [Hery and Villemain 2013] instead of a fill-light to avoid the bias of fill direction in our key-light direction comparison. In the toon-shaded stimuli, the effect of the fill-light was added by simply controlling the darkness of the shadow and the dome light was not needed.

In this study, we also considered a special case of no visible directional light (**No light**), a popular light rig (a group of lights saved for sharing and reusing) for preschooler television cartoon series, as it is quick to setup for multiple shots without much fine-tuning [Birn 2000]. We created this with the combination of a dome light and ambient occlusion to evenly light



**Figure 3.5:** Frames showing Mery emotions rendered in the “no” directional light condition

the character in the CG-shaded stimuli, and completely removed shadows in the toon-shaded stimuli (Figure 3.5).

### 3.2.4 Movies

The character performing each of the emotional sentences was rendered in 2 shading styles (toon-shaded, and CG-shaded), under 5 different conditions: Lo/Ab, Lo/Be, Hi/Ab, Hi/Be, and No. Lo corresponded to the low contrast condition, and Hi to the high contrast. Ab and Be referred to the direction of the light coming from above or below the character’s head. Finally, the No condition showed no directional lighting information. The animations were rendered at a resolution of  $1280 \times 720$ .

## 3.3 Experiment

A within-subjects design was used for this experiment where all participants saw each condition. We used 5-point Likert rating scales in order to collect the subjective opinions of participants towards the different render styles. We were particularly interested in whether the lighting and shading changed their ability to recognise the emotion, and how intense that emotion came across. We were also interested in which styles they found most appealing.

There were 100 trials in total in the experiment: 5 emotions (angry, fear, happy, sad, disgust)  $\times$  2 shading styles (toon-shaded, CG-shaded)  $\times$  5 lighting conditions (Lo/Ab, Lo/Be, Hi/Ab, Hi/Be, No)  $\times$  2 repetitions. In order to avoid fatigue we omitted the surprise

emotion as the animation appeared happy to participants in a pilot study, and we included only 2 repetitions in order to account for participant variation. Participants viewed each video clip in a random order, and after each clip they were asked to answer three questions:

*Which emotion did the character portray?:* Participants were asked to indicate their choice by pressing corresponding keys on the keyboard, marked with the words Angry, Sad, Disgust, Fear and Happy.

*How expressive was the indicated emotion portrayed by the character?:* Participants were asked to rate expressiveness on a scale of 1-5 by mouse clicking a slider on the screen, with 1 representing a rating of “Not expressive at all” and 5 representing “Extremely expressive”. They were instructed to base their decision on how strong an impression of the indicated emotion they saw in the motions of the character.

*How appealing was the character overall?:* Participants were asked to rate appeal on a scale of 1-5 by mouse clicking a slider on the screen, with 1 representing a rating of “Not appealing at all” and 5 representing “Extremely appealing”. They were instructed to base their decision on how much they were captivated by the character appearance.

### 3.3.1 Participants

University ethical approval was granted for the experiment. 23 volunteers (11 male, 12 female) aged between 17-50 took part in this experiment. Participants were recruited mainly via university student and staff mailing lists with different disciplinary backgrounds. They had normal or corrected to normal vision and were naïve to the purpose of the experiment. As a reward for participation, they were given a 5 euro book voucher. The experiment lasted approximately 20 minutes.

## 3.4 Results

For the statistical analysis, we conducted three separate repeated-measures Analysis of Variances (ANOVAs), one each for the results on **recognition**, **intensity** and **appeal**. Each ANOVA had the within-participants factors *emotion* (5), *shading style* (2), and *lighting condition* (5). Posthoc tests were conducted using the Newman-Keuls comparison of means, for this and all subsequent tests.

### 3.4.1 Recognition Accuracy

Responses were converted to *correct* or *incorrect* and averaged over repetitions. A main effect of emotion was found ( $F(4, 88) = 5.4, p < 0.0007$ ). Results indicated that sad and disgust were the least recognised, with angry, fear and happy being equally recognised Figure 3.6). All were recognised above 80% accuracy, indicating the participants were very good at the task of recognising emotion. No other main effects or interactions were found, indicating that the shading style or the lighting condition did not affect accuracy. Note that the pattern of our result resembles the recognition scores of the *Western Literate* group discussed in [Russell 1994], except anger. The high recognition rate of anger in our experiment could be due to the exaggerated animation and we would examine the influence of motion in Chapter 5. We would also investigate darker lighting conditions to determine the point at which recognition starts to become affected.

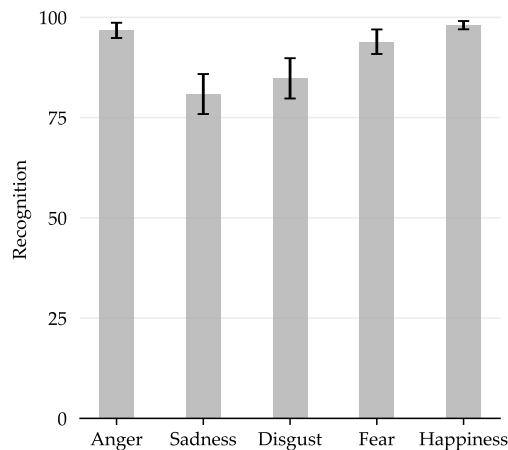


Figure 3.6: Main effect of Emotion on recognition accuracy

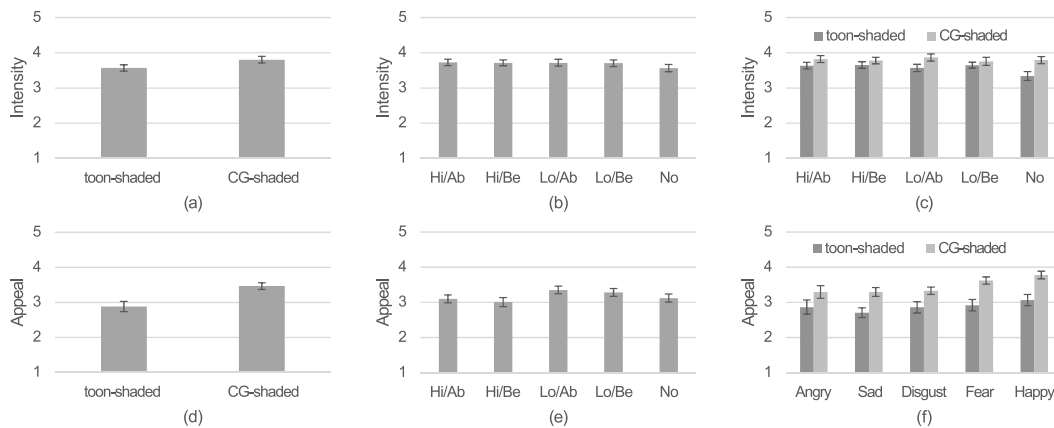
### 3.4.2 Intensity

A main effect of emotion was found ( $F(4, 88) = 25.07, p \approx 0$ ). Post hoc analysis indicated that the sad emotion was rated as the least intense ( $p < 0.002$  in all cases), fear and disgust next ( $p < 0.0004$  in all cases), and angry and happy were rated as the most intense ( $p < 0.0004$  in all cases). Intensity ratings were high in general, ranging from 3.11 for fear to 4.15 for angry. A main effect of shading style also occurred ( $F(1, 22) = 23.2, p \approx 0$ ), with posthoc showing

that animations rendered in CG-shading were rated as significantly more intense than those that were toon-shaded ( $p < 0.0003$ ) (Figure 3.7(a)).

A main effect of lighting condition was also found ( $F(4, 88) = 3.94, p < 0.006$ ), where posthoc tests showed that all conditions were rated equally intense except for No light which was rated as significantly less intense than all others ( $p < 0.02$  in all cases) (Figure 3.7(b)). An interaction between shading style and lighting condition ( $F(4, 88) = 3.94, p < 0.006$ ) gave us further insight, as posthoc tests showed that for CG-shading, there was no difference in ratings of intensity for any of the lighting conditions, whereas in toon-shading there was a difference, with the No light condition, rated significantly less intense than all others ( $p < 0.002$  in all cases) (Figure 3.7(c)).

No other interactions were found. In particular, we did not find an interaction between emotion and lighting condition, which was unexpected since we had hypothesised that different lighting conditions would have different effects across emotion (e.g., the angry emotion with a high contrast light below would be considered most intense). However, this was not the case and it seems that, for our examples, intensity of emotion perception is consistent across lighting conditions for different emotions. We will investigate this effect in more detail in later chapters.



**Figure 3.7:** Averaged ratings of some of the main effects and interactions for intensity (top) and appeal ratings (bottom). Hi: high contrast, Lo: low contrast, Ab: above light, Be: below light, No: no directional light.

### 3.4.3 Appeal

We found a main effect of emotion ( $F(4, 88) = 4.55, p < 0.003$ ), where posthoc analysis showed that the happy emotion was rated as more appealing than all others except for fear



( $p < 0.02$  in all cases). Ratings of appeal ranged from 3.0 for sad to 3.42 for the happy emotion. A main effect was also found for shading style ( $F(1, 22) = 16.8, p < 0.0005$ ), with CG-shaded being rated as significantly more appealing than toon-shaded (Figure 3.7(d)).

A main effect of lighting condition was also found ( $F(4, 88) = 5.69, p < 0.0004$ ), where posthoc analysis showed no difference between the light direction conditions (Hi/Ab vs. Hi/Be, or Lo/Ab vs. Lo/Be). However, both low contrast conditions (Lo/Ab and Lo/Be) were rated as significantly more appealing than both high contrast conditions (Hi/Ab and Hi/Be), with  $p < 0.02$  in all cases (Figure 3.7(e)). Finally, the No light condition was considered significantly less appealing than Lo/Ab.

One interaction was found between emotion and shading style ( $F(4, 88) = 2.68, p < 0.04$ ). Posthoc analysis showed that animations rendered in CG-shading were rated as significantly more appealing than toon-shaded renders for all emotions ( $p < 0.0002$  in all cases). As can be seen in Figure 3.7(f), in CG-shading style, all were rated as equally appealing except for fear and happy which were equally appealing, and more appealing than all others ( $p < 0.0007$  in all cases). In toon-shading style, the appeal ratings were more even across emotions, with only happy being rated as significantly more appealing than angry and sad ( $p < 0.04$  in all cases), and fear more appealing than sad ( $p < 0.05$ ).

### 3.5 Discussion

Our main finding of this experiment was that the shading style used (CG vs. toon) did not change the recognition of emotion, but did change the perception of intensity of that emotion, with toon-shaded renders being rated as displaying emotions as less intense than CG-shaded across the board. This implied that the smooth shading information in CG was important for the portrayal of intense emotions. Furthermore, we found that for CG-shading, there was no effect of shadow or lighting direction on emotion intensity. We believed this was due to the fact that shading information was present throughout all CG-shading conditions and was enough to convey high emotion intensity. For toon-shading, there was a big drop in intensity when there were no shadows or shading present (No light). These results indicated the importance of shading and lighting on the perception of emotion intensity. The analysis also showed that CG-shading was rated as more appealing than toon-shading and that low contrast lighting was preferred to high contrast lighting. Interestingly, we found no difference in appeal for the above and below lighting directions.

It was possible that our toon-shaded animations had lower intensity values than our CG-shaded animations due to the lack of visual information (Figure 3.5). In particular, the abstraction of the shaded surface could obscure facial shape and features crucial to emotion perception. This would be investigated further in Chapter 5 where more realistic characters (i.e. characters with more detailed shape information) were included.

Surprisingly, a main effect of shadow was not found in this study, contradicting the conventional practice of high-key lighting. We hypothesised that the threshold between the low and high contrast condition was quite large and we would like to investigate the influence of shadow intensity further. In the next chapter, we conducted a psychophysical experiment to better understand the perceptual space of the shadow and brightness. Modifying shadow inevitably affects the overall brightness of the scene as they are closely related. Although we tried to keep the brightness difference minimal across the stimuli in the same shading style, the averaged image intensity of CG-shaded and toon-shaded stimuli was distinguishable and it could have had an effect on the results. We hoped that a psychophysical model would provide us with well-defined brightness and shadow levels that were perceivable as different levels by the participants.

Chapter 4 focuses on how to create better stimuli with regular threshold of brightness and shadow in, and later in Chapter 5, we revisit and reexamine the perceived emotion and the appeal of virtual characters in greater detail. Hyde et al. [2013] found a large effect of auditory emotion level on perceived emotional intensity for their stimuli. In Chapter 4, we also aim to determine the effect of audio and other storytelling devices such as background and movement. Moreover, we aim to generalise the results over a larger set of characters and stylisation levels, all rendered in full-colour. However, we decided not to pursue the shading factor (toon vs. CG) further in this thesis as we moved to more practical lighting techniques, such as the use of the three-point setup and area lights, which are not compatible with toon shaders.

## Chapter 4

# Parameter Selection

### 4.1 Introduction

In the previous chapter, we took a broad view at the effect of lighting parameters on the perceived emotion and appeal of virtual characters, next we investigate brightness and shadow more closely. In this chapter <sup>1</sup>, we attempted to define levels of brightness and shadow intensity for usage in future experiments by mapping out different lighting conditions in a perceptual space. We did not investigate emotion or appeal here but sought to determine the range and threshold of the parameters that are perceptible to viewers.

We had already surveyed two key factors in lighting design, direction and contrast. Since the main effect of direction was not found in Chapter 3, we devoted this chapter to the psychophysics study of brightness and shadow in character lighting, by conducting several perceptual experiments, designed to investigate the ability of participants to discriminate lighting levels and the ratio of light intensity projected on the two sides of a cartoon character's face (key-to-fill ratio) in portrait lighting design. We applied a standard psychophysical method for measuring discrimination, typical in low-level perceptual studies but not normally considered for evaluating complex stimuli. We found that people can easily differentiate lighting intensities and distinguish between shadow strength and scene brightness under bright conditions but not under dark conditions. We provide a model of the results, and empirically validated the predictions of the model. We discuss the practical implications of our results and how they can be exploited to make the process of portrait lighting for CG cartoon characters more consistent, such as a tool for manipulating shadow while maintaining the level of perceived brightness.

---

<sup>1</sup>The content of this chapter was published in the ACM Symposium on Applied Perception 2019 (SAP'19)

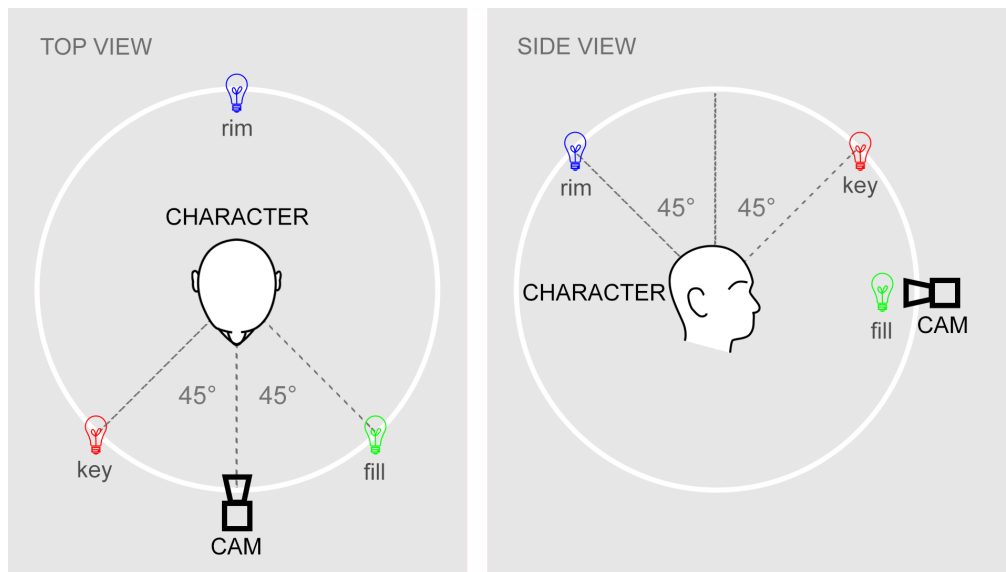


Figure 4.1: Three-point lighting setup with detailed light directions

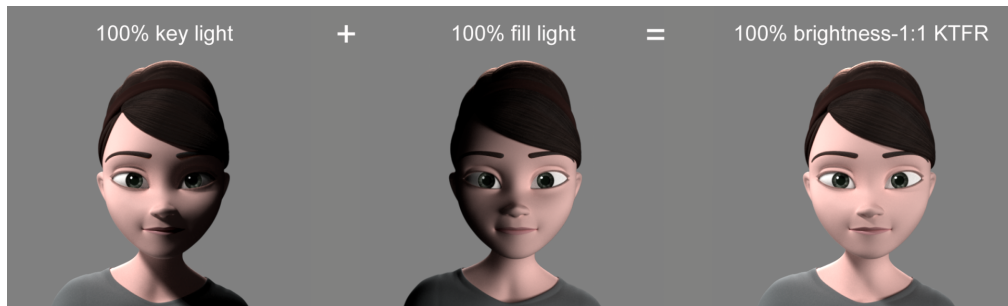
## 4.2 Stimuli

For the continuation from the previous study and going forward, we decided to use the same character model (Mery), however, in this study we excluded the effect of emotion by creating the stimuli with neutral expression and no audio. For more practical results than those of the previous experiment, we defined our range of brightness and shadow intensity based on a real-world design, and rendered our stimuli in colour.

### 4.2.1 Three-point Lighting

We set up our character, Mery, in a well-known, three-point lighting design [Millerson 1991] consisting of three light sources: a key light—the main source illuminating from one side of the camera, a fill light—from the opposite side—brightening up the shadow cast by the key light, and a rim light separating the background from the character. An additional key light (kicker) was also added to the key side of the character to lighten the edge of the character's face. The kicker's intensity is tied to the key light and considered part of the key intensity. We chose the three-point setup as it is the most basic and popular as it provided a wide range of controls that can be utilised to accentuate the emotion and appeal of a character [Birn 2000]. Figure 4.1 shows the relative angle of each light with respect to the camera and the character. We chose just one direction to test because we did not find a main effect of lighting direction on emotion intensity or appeal ratings in our previous experiment (Chapter 3), even for an

extreme key lighting direction from below.



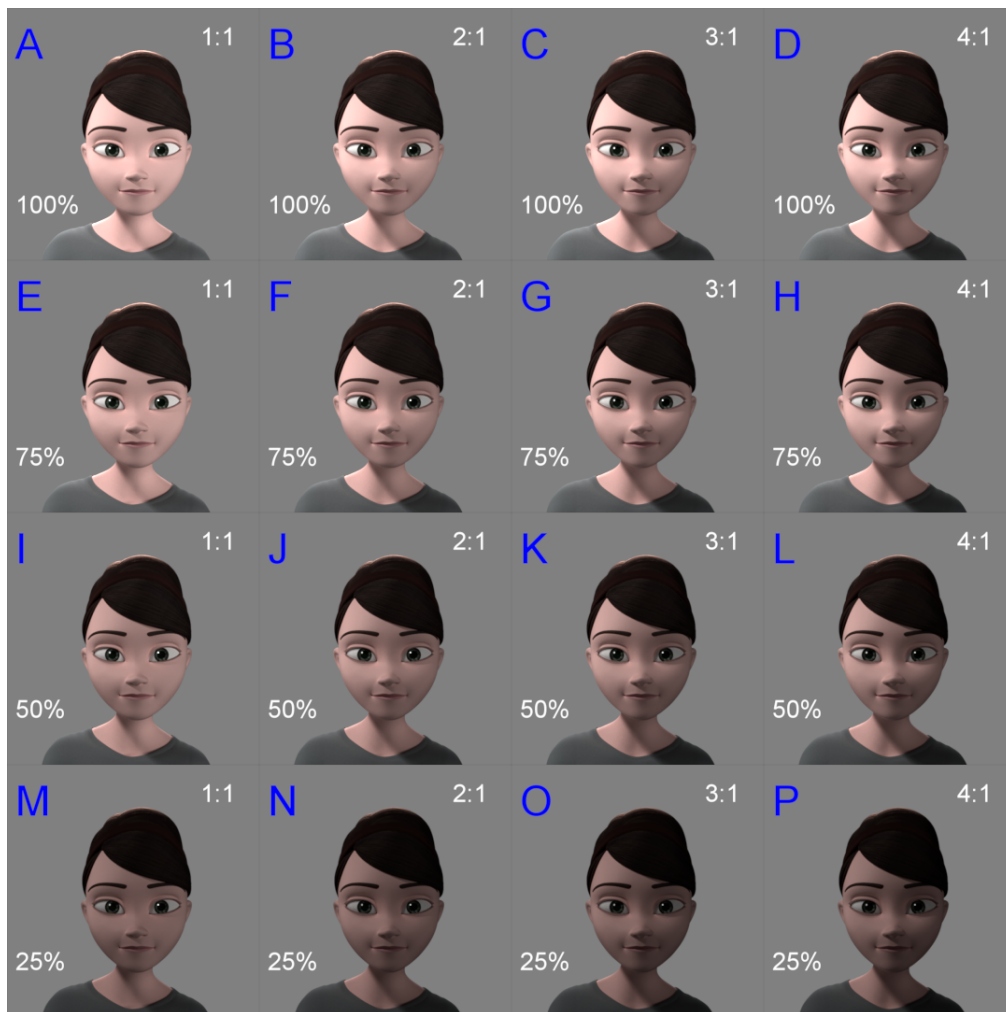
**Figure 4.2:** Contribution from key light and fill light were rendered separately and later combined and manipulated to create different levels of brightness and KTFR. The picture shows adding 100% key light brightness and 100% fill light brightness to make 100% brightness at 1:1 KTFR.

### 4.2.2 Light Intensities and Key-to-Fill Ratio

We followed the conventional 3D animation production pipeline and rendered the contribution from the key and fill light separately and later combined them for the desired *key-to-fill ratio (KTFR)*, the proportion of light intensity projected on the key side and fill side of the character face. (Note that it is valid to render images under separate light sources and then add them together since radiance can be summed together [Nicodemus et al. 1992].) KTFR is frequently used as a measure of contrast in cinematography and photography. We also adopted the conventional *expose-to-the-right* technique, illuminating the character with the maximum possible intensity before overexposing (no highlight clipping), to create the 100% *key intensity*—often referred to as key brightness or just **brightness** for the rest of the thesis—and 1:1 KTFR lighting condition (Figure 4.2). After that, we reduced the key and fill contributions in the gamma-corrected image space, with a gamma value of 2.2, to create 4 levels of key intensities: 100%, 75%, 50% and 25%, and 4 levels of KTFR: 1:1, 2:1, 3:1 and 4:1. We also ensured that there was no highlight clipping in any of the stimuli.

### 4.2.3 Movies

Mery was animated to the voiceover of neutral expression “there are magnets on the fridge” which lasted 50 frames or approximately 2 seconds when the animation was rendered at 24 frames per second. In studio portrait lighting, the background is typically in solid colour lit independently from the subject (no shadow interaction). To help reduce the influence of the background in our experiment, the character was rendered in front of a 18% gray



**Figure 4.3:** Still images taken from the 16 movies, rendered in different key light intensities (displayed in percentages) and key-to-fill ratios (displayed as ratios)

background, believed to be the middle gray according to the *Zone System* [Adams 1948] and perceived to be the midway between black and white in CIELAB colour space [Stone 2003]. Note that the use of a gray for the background is relatively common in psychophysics, but so is black or even speckled. Note that the surface with the highest intensity (and thus the white anchor [Gilchrist et al. 1999]) was the eyes of the character. Moreover, one eye was in the shadow and one was not, allowing a direct comparison. We rendered 16 animations (Figure 4.3) of the Mery character talking naturally with a neutral expression but without audible voiceover (key light intensity: 100%, 75%, 50%, 25% x 4 KTR: 1:1, 2:1, 3:1, 4:1).

#### 4.2.4 Environment

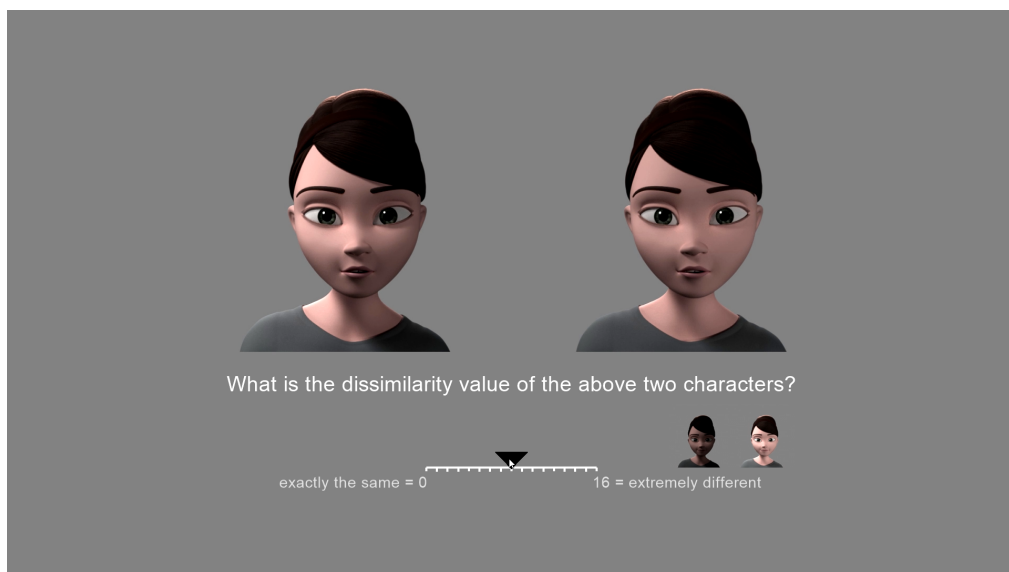
All lights used in our renders were white so the white point temperature takes on that of the monitor, Dell UP2713H monitor, which was calibrated to 100% sRGB colour gamut, 6500k white point, and  $80 \text{ cd/m}^2$  brightness. The experiment was conducted in a completely dark room to reduce the interference of outside light.

### 4.3 Experiment 1A - Dissimilarity

In this experiment, we aimed to identify combinations of brightness and shadow that were deemed perceptually different from each other using multidimensional scaling (MDS) analysis.

#### 4.3.1 Experiment

The task here was to rate the dissimilarity of two movie clips with different lighting conditions on a scale from 0 to 16. Note that the original study by Logvinenko and Maloney [2006] used 30-point scale but we reduced ours to just 17 as we had 16 different stimuli plus one zero-value for the exactly-the-same rating. While rating scales with more than 9 points probably do not add resolution, they also generally do not decrease accuracy either [Cunningham and Wallraven 2011].



**Figure 4.4:** A screenshot of a trial in the anchored MDS experiment (Experiment 1A)

Participants started with a training session during which they were shown example pairs of identical stimuli (A-A, P-P) and informed that this was a dissimilarity of 0, as well as example pairs of extreme difference (A-P, P-A) and told that this dissimilarity was a 16. Next, all 256 possible pairs were displayed in a random order. Still image thumbnails of an A-P pair were displayed beside the 16 end of the rating scale at all times, anchoring the maximum rating (Figure 4.4). The movies were looped continuously until the participant made a decision. After each rating, the participant was shown a blank screen with the middle-gray background and a white fixation point in the middle screen. Participants pressed the mouse button to start the next trial. The size of each movie in the pair was 10.3 x 10.3 degrees of visual angle. The pair was displayed 3.43 degrees of visual angle apart. The experiment lasted approximately 15-20 minutes.

### 4.3.2 Participants

Fifteen volunteers (7 females, 8 males, aged 23-50, avg. 35) took part in this experiment. The experiment lasted approximately 15-20 minutes. For this and the subsequent experiments, university ethical approval was granted. Participants were from different disciplinary backgrounds and were recruited primarily via university student and staff mailing lists. All had normal or corrected to normal vision and were naïve to the purpose of the experiment. A book voucher was given to participants as a reward for taking part.

### 4.3.3 Multidimensional Scaling (MDS)

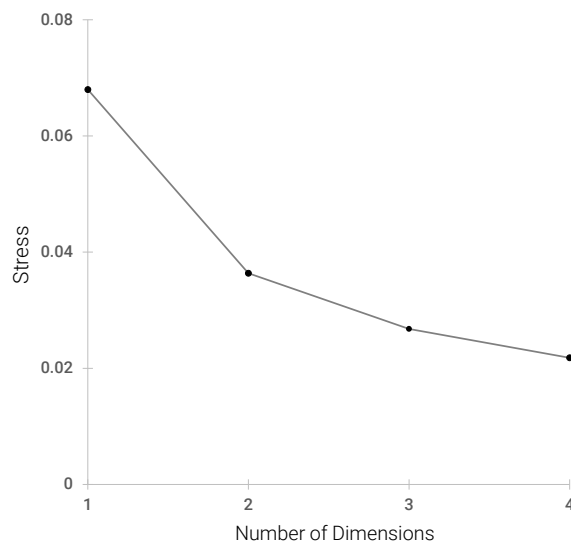
Most perceptual methods for determining the functional relationship between stimulus and perception require the experimenter to know the perceptual dimension or dimensions that are being studied. Richardson [1938] (see also [Torgerson 1952]) introduced a new class of method, referred to as MDS, to directly determine from a set of simple pairwise similarity ratings the number of perceptual dimensions involved, as well as their scale values. MDS takes an  $N \times N$  dissimilarity matrix of  $N$  stimuli and returns the coordinates of the  $N$  stimuli in a  $d$ -dimensional space that best fits the dissimilarity matrix (for more, see [Cox and Cox 2001; Cunningham and Wallraven 2011]). MDS assumes Euclidean distances. Since we use Likert scales, it is not clear that our similarity structure is, in fact, Euclidean. Thus, we used a non-metric MDS which uses the ordinal relationships to obtain scaled proximities.

All analyses were done in Statistica 7.1 and Matlab 2019a software packages and the computational details can be found in Borg [2012].



### 4.3.4 Results

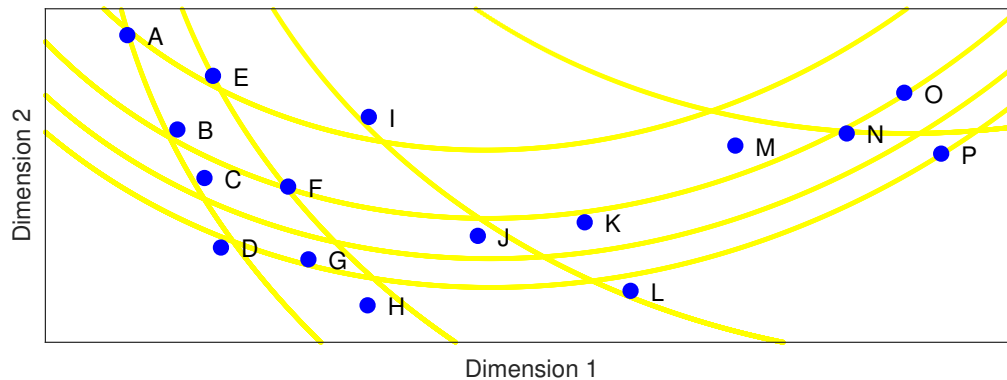
The ratings were averaged across participants to create the 16 x 16 dissimilarity matrix. The stress plot (Figure 4.5), Kaiser criteria, parallel analysis, and the optimal coordinates all suggest that two dimensions are sufficient to explain the dissimilarity in the matrix. The resulting MDS-derived perceptual similarity space (which has a stress value of 0.036, indicating a good fit) is presented in Figure 4.6, and Table 4.1.



**Figure 4.5:** The MDS stress plot suggests that two dimensions are sufficient to explain the dissimilarity as the stress is reduced more than half from one to two dimensions but does not decrease much from two to three dimensions.

As can be seen by analysing the positions of the blue data-points in the figure, the four key intensities for each KTFR (e.g., A, E, I and M) vary strongly and consistently along *Dimension 1*. Analyses show that the average pixel value (in any single frame of the animation) correlates nearly perfectly with *Dimension 1* ( $r^2 = 0.95$ ). It seems that *Dimension 1* represents scene brightness.

The effect of varying KTFR, known to control the contrast caused by shadows, for a given key intensity seems to be roughly orthogonal to the effect of key intensity. This, combined with a close examination of the stimuli, suggests that *Dimension 2* relates to the perception of shadow strength. One critical insight from the MDS space is that the effect of KTFR changes as a function of key intensity. When the key illumination is high (100%), participants appear



**Figure 4.6:** The blue data-points, representing the different levels of key intensity and KTFR are shown in the MDS perceptual space along the two dimensions. The yellow lines represent the proposed parametric log-polar model that best fits the data.

to perceive the “true” scene brightness (A, B, C, D). This is consistent with them (mostly) “discounting” the effect of shadows. Note that slight deviation from vertical at the highest key intensity might be due to the fact that as the KTFR decreases (at a given key level), the average pixel intensity will decrease slightly. As the key intensity decreases, participants seem to lose the ability to discount shadows. At the 25% key level, the effects of KTFR are perceived almost entirely as changes in scene brightness (M, N, O, P). This is consistent with Peli et al. [1991]’s finding that contrast perception changes as a function of illumination.

## 4.4 Experiment 1B - Un-anchored MDS

Note that the MDS Experiment 1A was performed using anchored scales. While this is standard in psychophysics, the exact choice of the anchors might cause biases and/or noise. As a control, we re-ran the experiment without anchors.

### 4.4.1 Experiment

We added a training session in which participants were shown all possible stimuli pairs in a random order without being asked to rate them. During this main part of the experiment, participants were told that 0 meant that the stimuli are identical and 16 meant that the stimuli are extremely different. To additionally test for accuracy, each pair was viewed 4 times in this experiment, with each pair being seen once before any pair was seen a second time (blockwise randomisation) resulting in a 50-60 minute experiment.

Stimulus	$x$	$y$
A	-1.2052	0.5052
B	-1.0251	0.1656
C	-0.9276	-0.0089
D	-0.8683	-0.2590
E	-0.8970	0.3587
F	-0.6266	-0.0398
G	-0.5536	-0.3015
H	-0.3408	-0.4670
I	-0.3362	0.2109
J	0.0560	-0.2170
K	0.4404	-0.1682
L	0.6054	-0.4148
M	0.9824	0.1077
N	1.3832	0.1514
O	1.5901	0.2979
P	1.7230	0.0787

**Table 4.1:** The coordinates of the blue data-points in Figure 4.6, representing the locations of the stimuli the MDS perceptual space.

#### 4.4.2 Participants

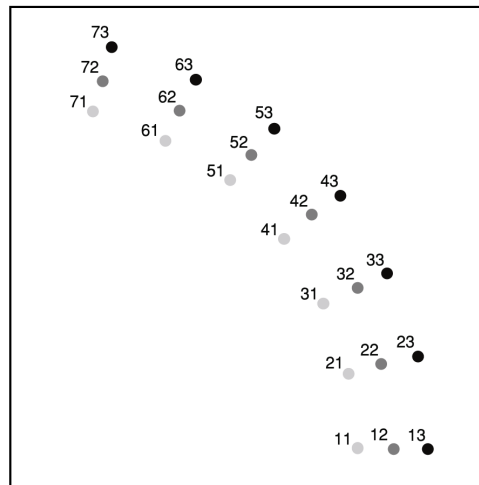
Sixteen new volunteers (6 females, 10 males, aged 18-37, avg. age 24) that had not taken part in the previous experiment, completed this un-anchored experiment. All other conditions remained the same.

#### 4.4.3 Results

The resulting MDS space was nearly identical to the anchored MDS space (Procrustes distance  $d = 0.0130$  [Dryden and Mardia 1998]). We therefore did not find any effect of anchoring or any significant improvement in the result with more repetitions per stimuli comparisons.

### 4.5 Parametric Model

Logvinenko and Maloney [2006] noticed that the MDS reconstruction of their data showed a “fan-like” structure (Figure 4.7). They thus parametrically modeled the perceived



**Figure 4.7:** The parametric model of Logvinenko and Maloney [2006] that best fits the observers' data. The points (light-surface pair  $ij$ ) lie along the radii of concentric ellipses.

dissimilarity ( $d$ ) of two achromatic Munsell chips, one with albedo  $A_i$  and illuminant  $L_k$ , and the other with albedo  $A_j$  and illuminant  $L_m$ , with the weight  $w_a$  and  $w_l$  as:

$$d = [w_a(\log A_i - \log A_j)^2 + w_l(\log L_k - \log L_m)^2]^{1.1} \quad (4.1)$$

which is an implicit function describing ellipses:

$$w_a \cdot X^2 + w_l \cdot Y^2 - D = 0 \quad (4.2)$$

where:

$$X = \log A_i - \log A_j$$

$$Y = \log L_k - \log L_m$$

and:

$$D = d^{0.91}$$

Since our data show a similar fan-like structure, we propose using a similar parametric model, describing a circular pattern. Specifically, we propose using the following parametric model:

$$[x, y] = [R \cdot \cos(T) + c_x, R \cdot \sin(T) + c_y] \quad (4.3)$$

where:

$$R = w_{br} \cdot \log_2\left(\frac{100}{B}\right) + w_{kr} \cdot \log_2(K) + c_r$$

$$T = w_{bt} \cdot \log_2\left(\frac{100}{B}\right) + w_{kt} \cdot \log_2(K) + c_t$$

and:

$B = \text{key intensity percentage (100, 75, 50, 25, etc.)}$

$K = \text{key-to-fill ratio (1/1, 2/1, 3/1, 4/1, etc.)}$

**Table 4.2:** Weight and constant values of the parametric model (Equation 4.3) that best fits the MDS perceptual space—the yellow lines in Figure 4.6.

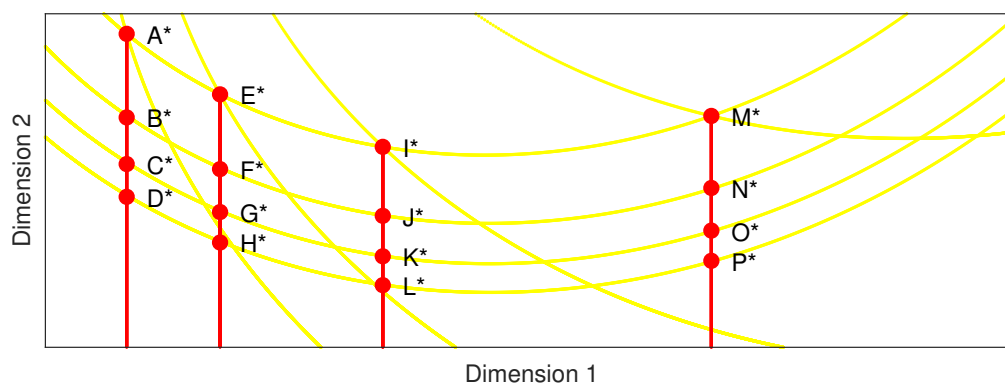
Variable	Value
$c_x$	-0.8844
$c_y$	2.3602
$w_{br}$	0.4569
$w_{kr}$	0.3687
$c_r$	1.8608
$w_{bt}$	5.6586
$w_{kt}$	24.9604
$c_t$	-100.0000

We used a least-squares curve fitting tool in Matlab (*lsqcurvefit()*) to fit the model to the MDS plot coordinates (Table 4.1) which yielded the 4 unknown weights ( $w_{br}$ ,  $w_{kr}$ ,  $w_{bt}$  and  $w_{kt}$ ) and 4 unknown constants ( $c_x$ ,  $c_y$ ,  $c_r$  and  $c_t$ ), with the residual norm of 0.08, indicating a good fit. The weight and constant values of the fit can be found in Table 4.2. The yellow lines in Figures 4.6 and 4.8 depict the proposed log-polar model approximating the MDS experiment result. Note that we also tried fit our data to a quadratic model but the result was not better than the proposed parametric model.

Since *Dimension 1* seems to be perceived scene brightness, we can use the model to create iso-brightness lines (red lines in Figure 4.8). That is, we can produce multiple KTRF levels that maintain the same level of perceived scene brightness.

## 4.6 Experiment 2 - Model Evaluation

To evaluate that this model can be used to produce stimuli with different KTRF levels while maintaining similar overall brightness, we ran an online “matching to sample” experiment.



**Figure 4.8:** The red lines show the location in the perceptual space of the stimuli with ‘adjusted’ brightness.

The conditions for this experiment were less controlled than the previous—participants were allowed to use any monitor setting in any room environment—so that we could ensure our results were valid in a non-laboratory setting.

#### 4.6.1 Stimuli

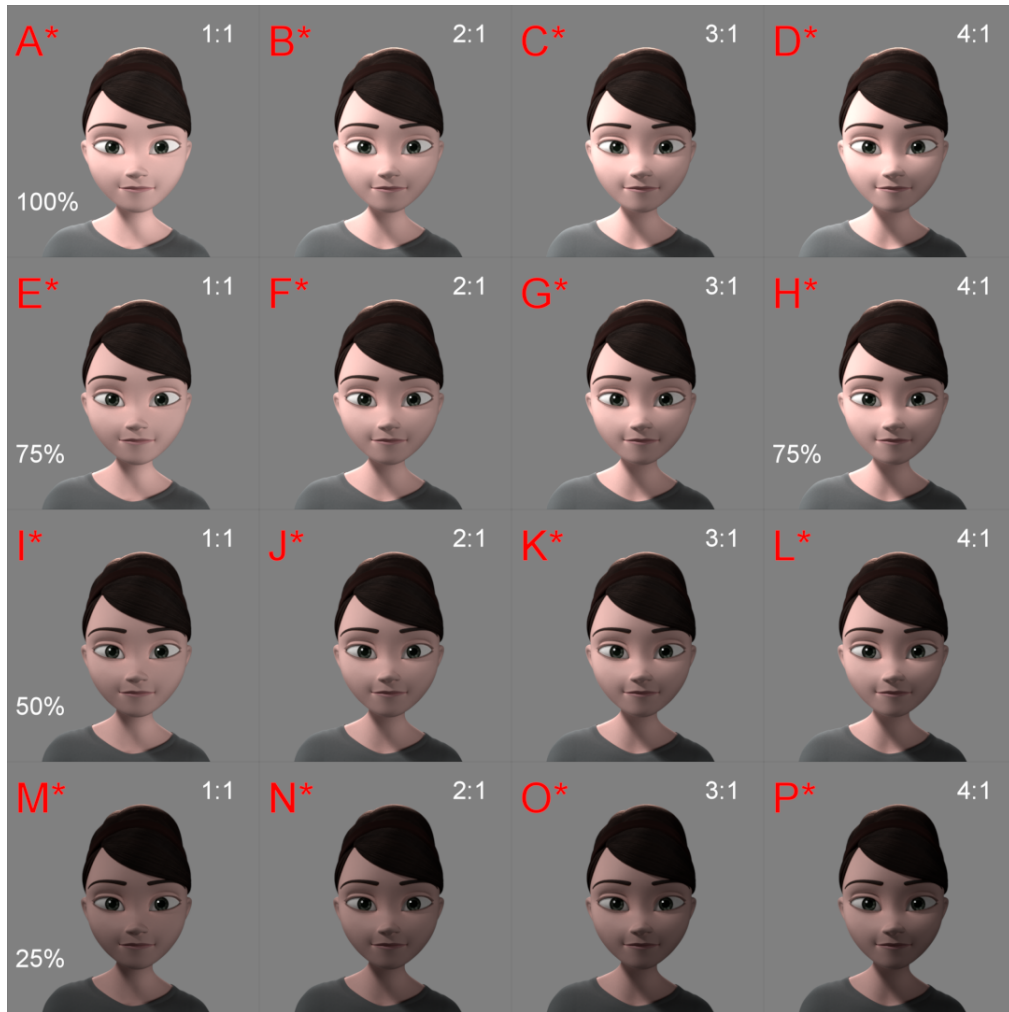
For each perceived iso-brightness level (red lines in Figure 4.8), we generated the adjusted stimuli (A\* - P\*) from the model (red points Figure 4.9 with the exact coordinates listed in Table 4.3).

#### 4.6.2 Experiment

In each trial (Figure 4.10), one key intensity (100%, 75%, 50%, or 25%) was randomly chosen. The 1:1 KTFR for the chosen key level was displayed at the top of the screen as the “sample” (i.e., A, E, I or M). One of the remaining KTFR values (2:1, 3:1, 4:1) was chosen for the two “match” stimuli (original vs. adjusted), which were displayed at the bottom. The participant had to choose which of the matches they believed to be more similar to the sample, in terms of overall brightness. Each combination of key intensity and KTFR was repeated 4 times in a random order and with a counterbalanced presentation of the left and right matches, yielding 48 trials in total.

#### 4.6.3 Participants

30 new volunteers (16 females, 4 males, aged 18-52, avg. age 28) that had not taken part in the previous two experiments, completed this online experiment. All were recruited similarly to



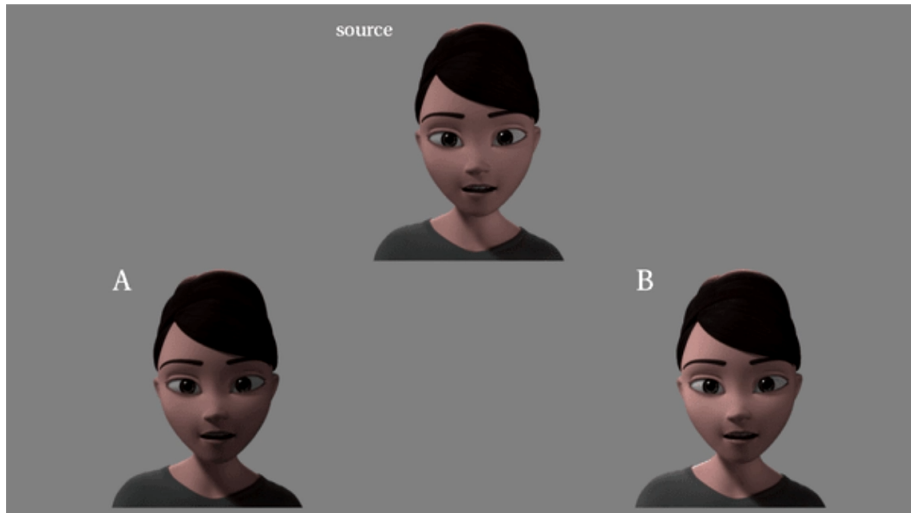
**Figure 4.9:** perceptually 'adjusted' stimuli set. Notice how the overall brightness is more consistent across the rows than in Figure 4.3.

the previous studies but they were not awarded with a book voucher.

#### 4.6.4 Results

On average, participants chose our adjusted stimuli 75.35% of the times, which is far above the 50% chance level of choosing the answers randomly. The percentage of time participants chose the adjusted stimulus were submitted to a two-way, repeated-measures ANOVA with brightness (4) and KTFR (3) as within-participants factors. No main effects or interactions were found, which implies that no systematic differences across brightness or KTFR were found in the percentage of times the adjusted stimuli were chosen. Note that some stimuli were altered more in Euclidean space than others (e.g., M\* - P\* were further from the original M - P than A\* - D\* were from A - D). Thus, our log-polar model can create different KTFR

44) Is the source movie more similar to A or B? \*  
(in terms of overall brightness)



A

B

**Figure 4.10:** A screenshot of the the Model Validation experiment (showing trial number 44)

levels with more consistent perceived brightness, regardless of the actual intensity.

## 4.7 Discussion

In this chapter, we investigated the perception of overall brightness levels and shadows strength (defined by the KTFR) in a carefully controlled, computer-generated scene. Our results extend previous knowledge from psychophysical experiments to a more complex three-point lighting setting with a cartoon-style virtual character.

Our key findings about human visual perception are: We have shown that *the ability to distinguish between shadows and overall illumination depends on the level of overall illumination*. When the overall illumination is high, people can reliably discount the effect of shadows to estimate the true light intensity. As the overall light intensity decreases, the effect of the shadows is increasingly mis-allocated to be a change in overall illumination. We have also shown that *the overall pattern of discrimination between brightness and shadow strength is very regular and can be captured with a log-polar function*. A set of well-spaced samples from the



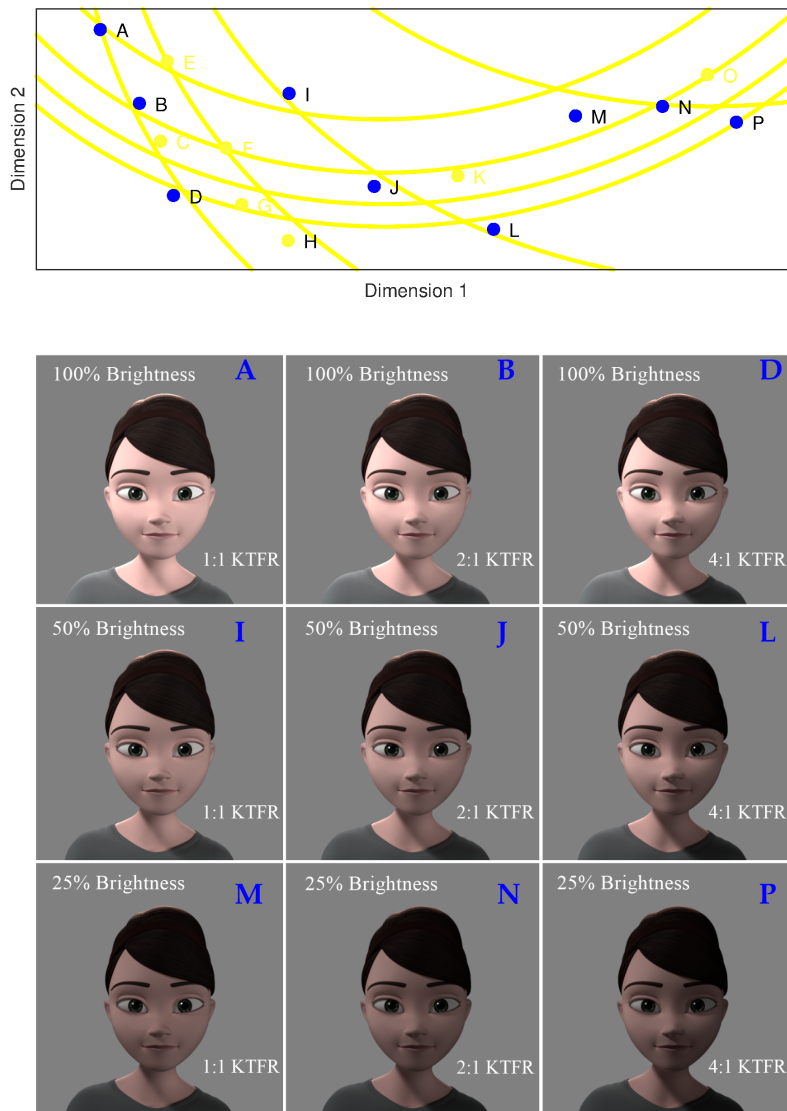
**Table 4.3:** The coordinates of the red data-points in Figure 4.8, showing the location in the perceptual space of the stimuli with “adjusted” brightness.

<b>Stimulus</b>	$x$	$y$
A*	-1.2075	0.5277
B*	-1.2075	0.2270
C*	-1.2075	0.0593
D*	-1.2075	-0.0586
E*	-0.8715	0.3099
F*	-0.8715	0.0411
G*	-0.8715	-0.1133
H*	-0.8715	-0.2230
I*	-0.2861	0.1212
J*	-0.2861	-0.1269
K*	-0.2861	-0.2725
L*	-0.2861	-0.3764
M*	0.8961	0.2324
N*	0.8961	-0.0267
O*	0.8961	-0.1800
P*	0.8961	-0.2888

lighting parameter space was crucial in continuing our study effectively. In other words, key intensity and KTFR could be systematically sampled for future experiments, such as in the next chapter, where we re-examined the effect of lighting design on the perception of higher level factors such as emotion and appeal that we briefly touched upon in Chapter 3.

In addition, our evaluation approach can be used for any CG character or scene after collecting a new dissimilarity matrix (which takes only 15 minutes per participant) and calculating the weights and constants. This model could also be applied to control illumination in studio photography, using the same approach. For a digital production, the model could be used to create a new type of lighting tool interface that we would demonstrate in Chapter 6.

Most importantly, we confirmed that a subset of the stimuli used in this chapter (Figure 4.11) were perceptibly different—being not too close to one another, and regularly distanced in the perceptual space of brightness and shadow intensity. This controlled set of stimuli, with smaller perceived thresholds than the one used in the previous chapter, will be used for further investigation into appeal and emotion in future experiments.



**Figure 4.11:** A subset of stimuli (top) representing the perceptual space of brightness and shadow intensity (bottom) that could be used in future experiments.

## Chapter 5

# Perception of Appeal and Emotion

### 5.1 Introduction

In Chapter 3, we examined the effect of high-contrast and low contrast lighting on the perceived emotion and appeal of virtual character, but our stimuli were broadly defined as high or low with a large parameter gap between them. In the last chapter, we learned the levels of brightness and shadow that were perceptually noticeable, and produced a set of stimuli evenly sampled from the perceptual space of character lighting. These new and verified stimuli put us in a better position to reassess our original study of the effect of lighting on emotion and appeal in greater detail. In this chapter <sup>1</sup>, we presented an extensive set of novel perceptual experiments designed to investigate the effects of brightness levels (key light intensity) and the proportion of light intensity illuminating the two sides of a character's face (key-to-fill ratio or KTFR). We dropped the lighting direction and shading from our study because a main effect of direction was not found in the earlier experiment and the toon shader is incompatible with three-point lighting used in this chapter.

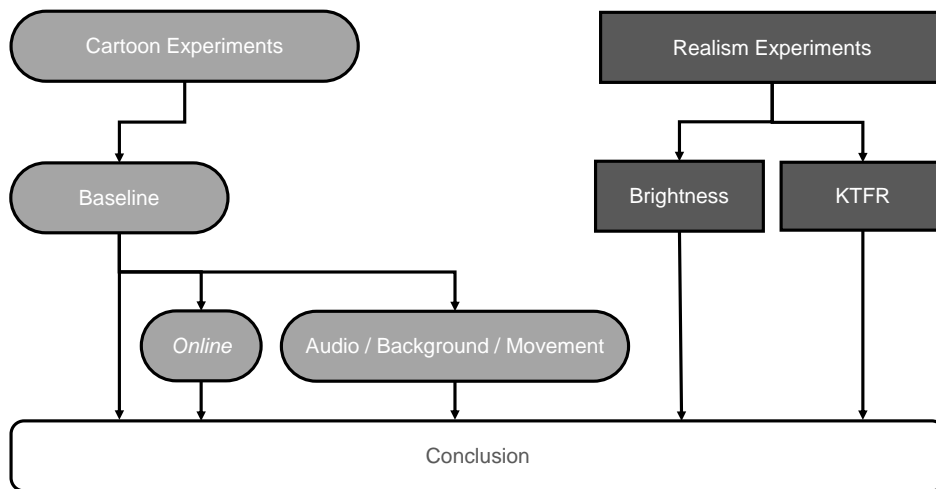
We divided our experiments into *cartoon experiments* where all conditions were tested on cartoon characters, and a set of *realism experiments* reviewing if our results hold for characters with more realistic proportion and details (Figure 5.1).

### 5.2 Cartoon Experiments

In this set of experiments, we were interested in how key light brightness and key-to-fill ratio affect the perception of the emotion and appeal of a cartoon animated, stylised CG

---

<sup>1</sup>The content of this chapter was published in the ACM Transaction on Graphics (TOG)



**Figure 5.1:** The experiments were divided into cartoon and realism experiments.

character. We hypothesised that lighting, based on the recent study by Barchard et al. [2017], will have certain associations with the perception of particular emotions portrayed by the characters. More specifically, we expected to see perceived happiness responding to the change of brightness, as well as perceived sadness, anger and fear responding to darkness.

### 5.2.1 Stimuli



**Figure 5.2:** (left to right) Mery, Jasmine, Franklin and Malcolm characters rendered in their original shaders

For the CG characters used in these experiments, we continue to work with Mery from the preceding experiments, and added Jasmine<sup>2</sup>, Franklin<sup>3</sup>, and Malcolm<sup>4</sup>. They all had typical appearance commonly seen in animated films (Figure 5.2), as well as detailed, controllable facial rigs enabling an animator to create high-quality animations. The appearance of each

<sup>2</sup><http://www.cgmeetup.net/home/jasmine-rose-rig-free-maya-character-rig-female-character-rig/>

<sup>3</sup><https://artella.leadpages.co/artella-character-giveaway/>

<sup>4</sup><https://www.animschool.com/>

character was normalised to avoid the effect of colour differences by using the same set of materials.



**Figure 5.3:** (left to right) Mery, Jasmine, Franklin and Malcolm characters rendered in our experiment shaders

## Recordings

<b>Emotion</b>	<i>Affective sentence</i>
Neutral	<i>There are magnets on the fridge.</i>
Anger	<i>I'm sick of you being late.</i>
Sadness	<i>My best friend is moving away.</i>
Fear	<i>Someone is following me.</i>
Happiness	<i>It's a beautiful day outside.</i>

**Table 5.1:** Validated affective sentences for neutral, anger, sadness, fear and happiness

Emotional dialogue was recorded in our studio from a male and female actor that were asked to convey a set of validated affective sentences for spoken emotion identification [Ben-David et al. 2011]. We added neutral emotion to test the effect of lighting on appeal without the modulation of expression, and in turn had to drop disgust to keep the length of the experiment manageable. We chose to omit disgust because it was excluded from [Ben-David et al. 2011] as well. The dialogue was recorded for a neutral expression and emotions of anger, sadness, fear and happiness (Table 5.1). Each sentence lasted approximately 2 to 4 seconds. We then commissioned a set of animations for each characters from 2 professional animators, who were provided with the audio recordings. The voice acting was performed by a female actor for Mery and Jasmine, and a male for Franklin and Malcolm as we wanted the recorded voice and acting to feel natural to the gender of the character. Although the dialogue for each emotion is the same (e.g., “I’m sick of you being late” for anger), the animator created different facial motions for each character to match the

voice, intonation, and character's appearance for each emotion (Figure 5.4). Therefore, in our experiments it is not possible to separate the effect of the character's appearance, gender and performance (we do not test for these separate factors).



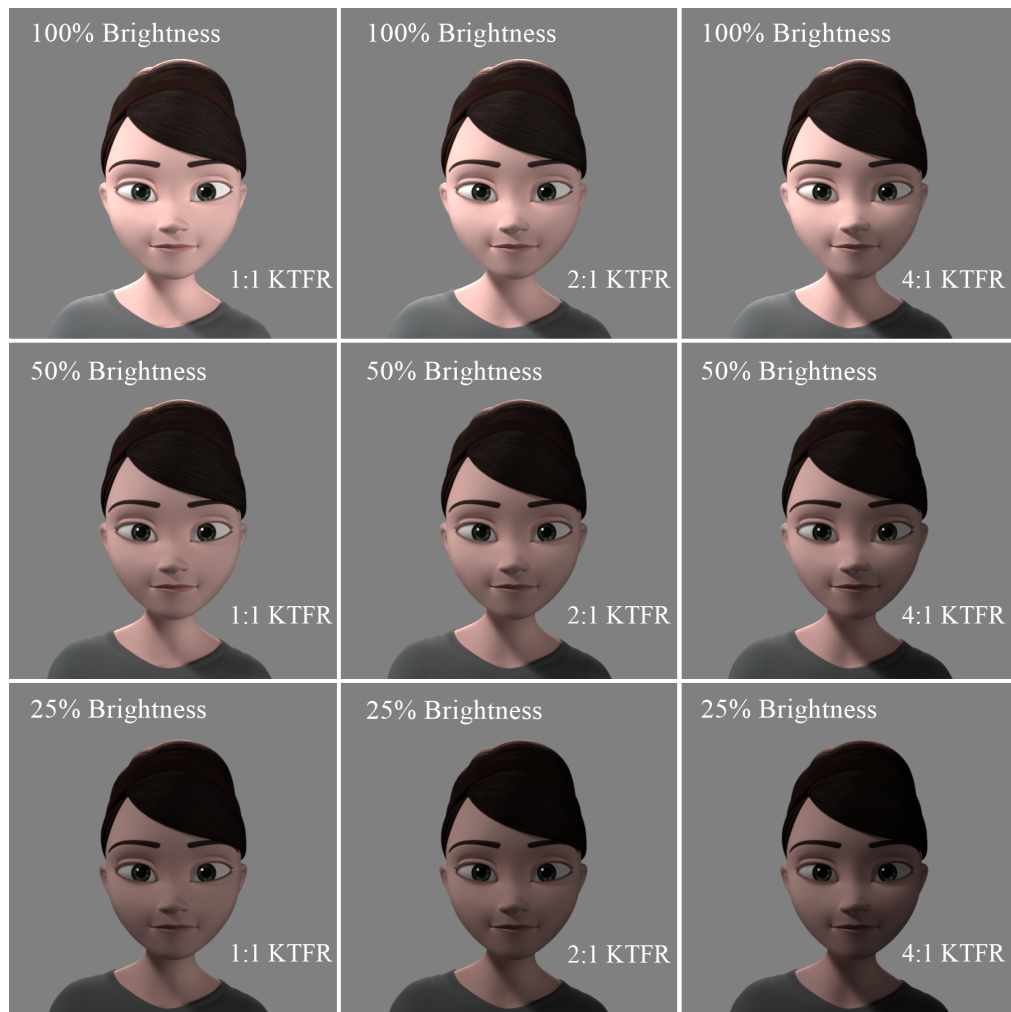
Figure 5.4: (left to right) Mery, Jasmine, Franklin and Malcolm portraying anger

### Lighting

Each character was again lit in the same three-point lighting setup used in the last chapter. It was essential to select lighting conditions that were uniformly sampled from the lighting perceptual space. In the previous psychophysical experiments on cartoon stimuli with ranging key light brightness and KTFR. Multidimensional scaling (MDS) analysis was used in order to model the connection between the lighting changes and the perceived scene brightness and shadow depth, which revealed the perceived (dis)similarity structure and dimensions of the interested entities. Based on this MDS plot, we selected 9 lighting conditions (3 levels of brightness: 100%, 50 %, 25% and 3 levels of KTFR: 1:1, 2:1, 4:1) that were a well-spaced, regular sampling of the perceptual space of KTFR and key light brightness (Figure 5.5). We also applied the same definition of the *100% brightness and 1:1 KTFR condition* and subsequently applied the lighting conditions used in the MDS experiments to all characters.

### Movies and presentation

Each character was rendered in front of a middle grey background to reduce the effects of relative lightness. The three brightness levels combined with three KTFR levels yielded 180 movie clips (5 emotions, 4 characters). These movies contained no audio so participants could focus specifically on the appearance of the character. A second subset (with just 2 characters) was generated with synchronised audio, to assess the effect of the character's voice. Audio levels were normalised to ensure they were at the same volume level for each emotion. A



**Figure 5.5:** Nine lighting conditions used in the Cartoon Experiments, which were rendered with different (key light) Brightness and key-to-fill ratios (KTFR).

third subset (with just 2 characters) was generated with complex backgrounds (Figure 5.6) instead of grey, to assess the effect of background complexity.

Since it is known that there is some information for emotion that is only available over time and not in any given frame [Cunningham and Wallraven 2009a] and that this dynamic information can compensate for degraded or noisy static information [Cunningham and Wallraven 2009b], we created another set of stimuli containing only a still image of an extreme pose chosen by hand from each movie of each emotion for Mery and Franklin characters. Comparing the results of static and dynamic stimuli allows us to begin to separate the contribution of shape from the contribution of motion.

Stimuli for all laboratory experiments were presented on a 27" Dell U2713Hb monitor (100% sRGB calibrated with the white point of 6500k and  $80 \text{ cd}/\text{m}^2$ ) in a completely dark



**Figure 5.6:** Franklin and Mery with neutral expression in front of two complex background scenes.

room to reduce the interference of outside light. The size of each movie was 600 x 600 pixels which corresponded to approximately 10.3 x 10.3 degrees of visual angle. The online experiments were presented on a range of devices from a tablet to a large desktop monitor and the movie size was reduced to 300 x 300 pixels for compatibility, and we did not control for the environment lighting conditions.

### 5.2.2 Laboratory Experiment (Baseline)

Our first experiment was our baseline, a controlled laboratory experiment where we were interested in whether lighting changed a participant's capability to recognise emotion, and how intense that emotion was conveyed. We were also interested in which lighting combinations were most appealing to the participant. For the baseline, movie clips of the Mery and Franklin characters with no audio as stimuli were chosen to determine the effect of lighting on appearance alone.

A within-participant design was used for this experiment: all participants saw each combination of character, emotion and lighting, and 7-point Likert rating scales were used to gather the subjective opinions of participants toward the different lighting conditions.

There were 180 trials in total in this experiment: 2 characters (Mery and Franklin, with different animations)  $\times$  5 emotions (neutral, anger, sadness, fear, happiness)  $\times$  3 levels of brightness (100%, 50%, 25%)  $\times$  3 levels of key-to-fill ratio (KTFR: 1:1, 2:1, 4:1)  $\times$  2 repetitions. In order to avoid fatigue, we included only 2 repetitions. Movie clips were viewed by the



participants in random order, and after each movie they were instructed to answer three questions:

*Which emotion did the character portray?:* Participants used a mouse to click on one of the words that were displayed on-screen: Neutral, Anger, Sadness, Fear, Happiness or Other.

*How intense was the indicated emotion portrayed by the character?:* Participants rated the intensity on a scale of 1-7 by moving a slider on the screen, using a mouse-click. 1 on the slider represented a rating of “Not at all” and 7 represented “Extremely”. They were informed to base their judgement on how strong their impression of the portrayed emotion was.

*How appealing was the character overall?:* Participants rated appeal on a scale of 1-7 by moving a slider on the screen, using a mouse. 1 on the slider represented a rating of “Not at all” and 7 represented “Extremely”. They were informed to base their judgement on how much they would be captivated by a movie with that character in the leading role, and would like to watch more of them.

### **Participants**

Fifteen volunteers (5F, 10M, aged 18-57, avg. 35) took part in this experiment. The experiment lasted approximately 30-40 minutes and the participants were allowed to take a 10-15 minute break mid-way through the experiment. In this and all subsequent laboratory experiments, university ethical approval was granted for the experiment. Participants were from different disciplinary backgrounds and were recruited primarily via university student and staff mailing lists. All had normal or corrected to normal vision and were naïve to the purpose of the experiment. A book voucher was given to participants as a reward for taking part.

### **Results**

For the statistical analysis, we conducted three separate repeated-measures Analysis of Variances (ANOVAs), one each for the results on **recognition**, **intensity** and **appeal**. Each ANOVA had the within-participants factors *character* (2), *emotion* (5), *brightness* (3), and *KTFR* (3). The ratings for each participant were averaged over the two repetitions. We ran Mauchly’s test for validating sphericity of the data, and whenever it is significant we report results with Greenhouse-Geisser correction applied and marked with an asterisk “\*”.

Posthoc tests were conducted using the Tukey Honestly Significant Difference (HSD) test for the comparison of means. We only report the main results of the experiment in this section. For the summary of all significant effects and interactions with post-hocs, see Table A.1 in the Appendix.

%	Neutral	Anger	Sadness	Fear	Happiness
Neutral	<b>77.96</b>	0.37	2.22	0.93	0.37
Anger	3.33	<b>98.52</b>	2.96	0.19	0.00
Sadness	0.74	0.56	<b>85.93</b>	3.70	0.19
Fear	0.34	0.56	4.07	<b>86.85</b>	0.19
Happiness	14.26	0.00	0.19	0.93	<b>98.33</b>
Others	3.33	0.00	4.63	7.41	0.93

**Table 5.2:** The confusion matrix of the recognition rating of Mery character

%	Neutral	Anger	Sadness	Fear	Happiness
Neutral	<b>90.19</b>	0.37	0.74	0.56	0.19
Anger	0.56	<b>97.96</b>	0.56	3.89	0.37
Sadness	6.85	0.37	<b>94.26</b>	14.26	0.00
Fear	0.37	0.56	1.11	<b>74.07</b>	0.19
Happiness	0.00	0.56	0.19	0.00	<b>94.44</b>
Others	2.04	0.19	3.15	7.22	4.81

**Table 5.3:** The confusion matrix of the recognition rating of Franklin character

**Recognition:** For the recognition of emotions, responses were converted to scores “1” (correct) or “0” (incorrect) and averaged over stimuli repetitions. Recognition rates for all emotions were very high, ranging from 81% to 98%, where all emotions were recognised equally well except for anger which was significantly more recognised than fear ( $p < 0.009$ ). The confusion matrices (Tables 5.2 and 5.3) showed that fear was often mistaken as sadness or others. The ambiguity of fear was well documented in previous facial recognition studies [Russell and Fernández-Dols 1997].

Some small differences in how brightness and KTFR combinations affected the recognition rates for individual emotions were found, but no trends.

**Intensity:** Intensity ratings were high in general, which is expected for exaggerated cartoon animations, where anger was rated as the most intense (average: 6.15), followed by



**Figure 5.7:** Interaction between Brightness and Emotion for ratings on Intensity in the baseline experiment. Star labelled lines point to significantly different means according to the post-hoc test. Error bars show standard error of the means.

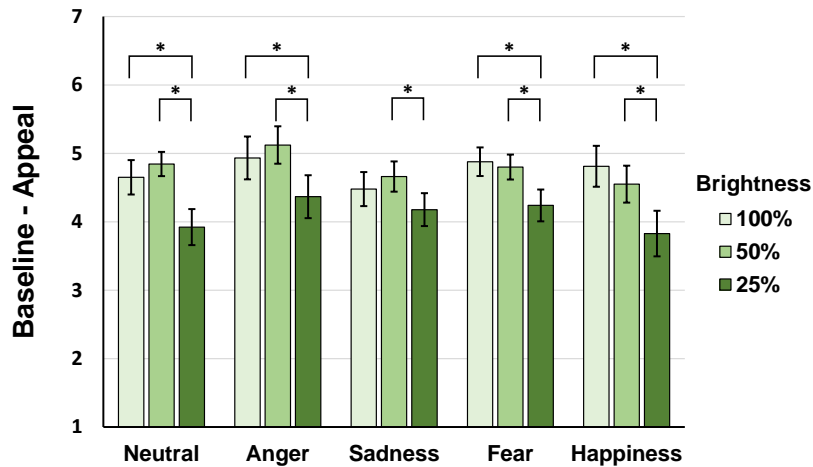
happiness (average: 5.63), then fear (average: 4.95) and sadness (average: 4.94) were rated as the least intense ( $p < 0.009$  in all cases).

Interestingly, an emotion specific effect of brightness was found, where for happiness 100% brightness was rated as significantly more intense than 25% ( $p < 0.004$ ). For anger, sad, and fear, there was no significant difference in intensity ratings for the different brightness levels (Figure 5.7). However, for sadness, there was a trend showing the darker condition was rated as more intense. We conclude that *brightness is important for the portrayal of intensity in a happy emotion*. There was no effect of or interaction with KTFR for intensity ratings.

**Appeal:** A main effect of brightness was modulated by an interaction between brightness and emotion (Figure 5.8). For all emotions except sadness, 25% brightness was less appealing than 50% and 100% ( $p < 0.002$ ). For sadness, 25% was only less appealing than 50% ( $p < 0.012$ ). In general, *brightness had a strong effect on appeal, with low brightness rated as less appealing than higher brightness for all emotions*. While an interaction between brightness and KTFR occurred, no significant differences were found in the post-hoc tests, implying that *KTFR had no major effect on appeal ratings*.

### 5.2.3 Online Experiments

Our online experiments were devised to confirm and generalise the effects found in the baseline experiment. In particular, we wished to confirm the impact brightness had on the appeal and intensity of the happiness and sadness emotions – as clear effects and tendencies



**Figure 5.8:** Interaction between Brightness and Emotion for ratings on Appeal in the baseline experiment. Star labelled lines point to significantly different means according to the post-hoc test. Error bars show standard error of the means.

were found in the baseline, and to generalise the results with additional characters and environmental conditions. All 4 cartoon characters (Jasmine, Malcolm, Mery, Franklin) were used with a reduced set of brightness levels: 100% and 25%, as we were only interested in the extreme changes which affected our results in the laboratory experiment. As KTFR did not affect our results previously, we used 1:1 KTFR throughout.

The well-known drawbacks of remote experiments such as the lack of experimental control and technical limitation [Reips 2000] forced us to conduct our experiments in a less-controlled environment than the Baseline (varying monitor size, brightness of the room, monitor settings, etc.). However, the disadvantages were compensated by the access of more diverse and larger pool of participants [Paolacci et al. 2010] (120 in total), which we acquired through US and European crowd-sourcing platforms MTurk<sup>5</sup> and Prolific<sup>6</sup>. In total, 80% of the participants were native English speakers. Additionally, some modifications had to be made in order to conduct the experiment online. Firstly, we only used one rating scale throughout the experiment to keep instructions minimal and clear. Therefore, we conducted two experiments in total. The first experiment investigated intensity alone, while the second investigated only appeal, and we altered the question from the Baseline to allow for the fact that participants were not asked to categorise the emotion, by specifically telling them that the emotion being portrayed was either happy or sad. For the happy emotion, the questions were:

<sup>5</sup><https://www.mturk.com>

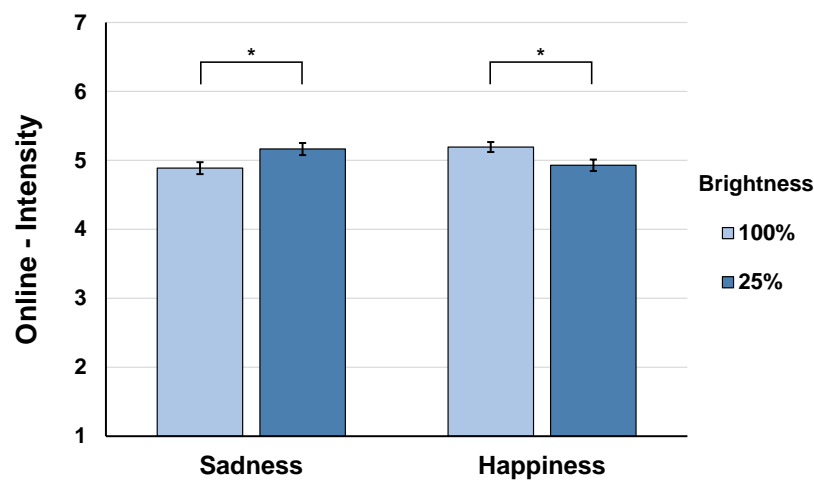
<sup>6</sup><https://prolific.ac>

- *How intense is the happy emotion portrayed by the character?:* Participants were informed to rate the intensity on a scale of 1 to 7 (1 = not intense at all and 7 = extremely intense).
- *The character is happy. How appealing is the character overall?:* Participants were informed to rate the appeal on a scale of 1-7 (1 = not appealing at all and 7 = extremely appealing).

Sixty participants (22F, 38M, aged 18-62, avg. 33) in total took part in the intensity experiment and an additional 60 participants (26F, 34M, aged 18-62, avg. 31) took part in the appeal experiment. Half of participants were sourced from MTurk and half from Prolific for each experiment, and ratings were averaged over the two repetitions, as before. In each experiment, 32 animation clips (4 characters, 2 emotions: happiness and sadness, 2 brightness levels: 100% and 25%, and 2 repetitions) were shown to the participants in random order. A within-participant design was used for each experiment: all participants saw each combination of character, emotion and lighting.

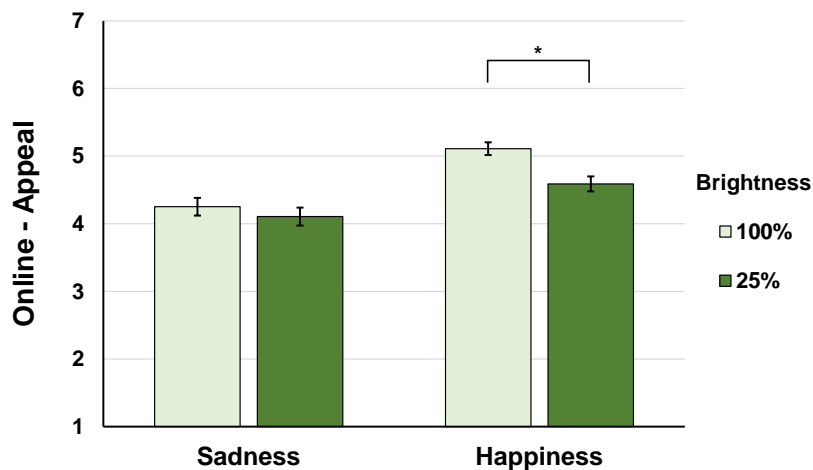
## Results

For the statistical analysis of each experiment, we conducted a repeated-measures ANOVA with within-groups factors: *character* (4), *emotion* (2), *brightness* (2) and *source* (2). We also ran the Mauchly's sphericity test and found no significance. Post-hoc tests were conducted using the Tukey Honestly Significant Difference (HSD) test for the comparison of means. Only relevant main effects and interactions are reported in the text. See Table A.2 in the Appendix for all significant results.



**Figure 5.9:** Interaction between Brightness and Emotion for ratings on Intensity in the online experiment. Star labelled lines point to significantly different means according to the post-hoc test. Error bars show standard error of the means.

**Intensity:** As in the Baseline, an emotion specific effect of brightness was found (Figure 5.9). For happiness, 100% brightness was perceived as more intense than 25% ( $p < 0.009$ ). For sadness, we previously found a trend, while here we find a significant effect where the 25% was perceived as more intense than 100% ( $p < 0.024$ ). Therefore *we replicated the effects of brightness on emotion intensity*. The source also affected the intensity ratings, with the Prolific participants rating sadness as more intense than the MTurk participants ( $p < 0.031$ ), while there was no difference for happiness. As expected, some characters were found to be more intense than others portraying the different emotions, but we found no interactions between character and brightness, implying that *the effect of brightness on the perception of intensity generalised across all characters tested*.



**Figure 5.10:** Interaction between Brightness and Emotion for ratings on Appeal in the online experiment. Star labelled lines point to significantly different means according to the post-hoc test. Error bars show standard error of the means.

**Appeal:** Happiness was found to be more appealing than sadness ( $p < 0.0001$ ) overall (Figure 5.10). Additionally, we found that 100% brightness is overall more appealing than 25%, which replicated our result from the Baseline. However, this time it was emotion-specific, where 100% brightness was rated as more appealing than 25% ( $p < 0.0002$ ), but for sadness, there was no significant difference, but a trend can be observed. With regard to characters, there was no interaction between brightness and character, implying that *brightness had the same effect on appeal ratings across all characters tested*.

Finally, it is worth to note the slight differences in the ratings between the lab and online experiments. This could be due to the higher quality of the lab monitor (brighter white and

darker black) and the extremely low-light experiment room contributing to the better perceptible levels of brightness, and hence wider gaps between 100% and 25% brightness ratings in the lab.

**Table 5.4:** Summary of conditions and participants in Cartoon Lab Experiments.

Experiment	Movement	Audio	Background	Participants
<b>Baseline</b>	Movie	No	Gray	5F, 10M aged 18-57 avg. 35
<b>Audio</b>	Movie	Yes	Gray	7F, 8M aged 18-37 avg. 31
<b>Background</b>	Movie	No	<i>Complex</i>	9F, 6M aged 18-42 avg. 26
<b>Movement</b>	<i>Still</i>	No	Gray	4F, 11M aged 18-47 avg. 30

#### 5.2.4 Audio, Background and Movement Experiments

In films and games, character lighting is generally accompanied by other storytelling devices such as audio and the background. Therefore, our final set of Cartoon Experiments aimed to investigate additional effects that were not included in the Baseline. A different group of 15 participants (see Table 5.4 for details) took part in each experiment and assignment to a block was random, based on order of appearance.

The first experiment investigated the addition of audio, which we hypothesised would change the intensity of the perceived emotions. Participants were shown the movie clips of Mery and Franklin, with synchronised audio, which they listened to on a set of closed-back headphones.

The second experiment investigated background complexity on user perception, which we hypothesised would lessen the perceived intensity of the emotion as there was more competing information to process. Participants viewed movie clips of Mery and Franklin in context with 2 different, more complex, background scenes (Figure 5.6).

The third experiment investigated the effect of movement, and we hypothesised that emotions would be perceived as more intense when still than when moving, since a still image would show the peak of the expression rather than the full sequence (with anticipation, etc.). Participants viewed only the still images of Mery and Franklin for 4 seconds each, the average length of the movies.

A within-participant design was used for each experiment and the details of the experiments were kept exactly the same as the Baseline (Section 5.2.2).

## Results

For the statistical analysis of each experiment, we conducted three separate repeated-measures Analysis of Variances (ANOVAs), one each for the results on **recognition, appeal and intensity**.

Since all elements of the new experiments were identical to the Baseline with the sole exception of the new manipulation, we treat each of these new conditions as a between-groups factors (*audio, background, or movement* respectively), and compare against the Baseline data (Section 5.2.2).

Therefore, for the analysis of each new condition, we conducted a between-groups ANOVA with 2 conditions - new condition and baseline, and within-participants factors *character* (2), *emotion* (5), *brightness* (3), and *KTFR* (3). We only discuss interesting new results. Please see Tables A.3, A.4 and A.5 in the Appendix for all significant effects.

### Audio

As before, recognition was high, but in this case *emotion recognition was higher with audio than without audio*. Recognition was over 98% for all emotions, compared with a range from 80% to 98% in the Baseline. This is reflected in the significant interaction between emotion and audio and interaction between emotion, character and audio which could be due to the tone and semantic content of the voice-over.

For appeal, an interaction occurred between emotion and audio, which showed that only the neutral expression was considered significantly less appealing in the audio condition than in the baseline ( $p < 0.005$ ), perhaps due to the addition of a monotone or boring voice portraying a neutral sentence.



### Background

There was no effect of background on recognition, intensity or appeal. Implying that *a complex background did not alter any ratings.*

### Movement

Movement affected emotion intensity, where the most intense emotions (anger and happiness) were rated even more intense in the movies than in the still images ( $p < 0.02$ ), whereas fear was less intense in the movie ( $p < 0.009$ ). This result could be influenced by the selection of the frame for the still image from the movie, as there were fewer 'peak' frames of emotional intensity in the fear sequence than in anger or happiness. Additionally, anticipation [Thomas and Johnston 1995] was used in some sequences more than others, which could have caused a heightening of intensity for certain movies over the still images. KTFR also interacted with movement, where KTFR 4:1 increased the intensity but only for still images ( $p < 0.02$ ).

### 5.2.5 Cartoon Experiments - Discussion

Emotion recognition for all of our cartoon characters and exaggerated emotions was very high, both for still images and movies, and generally unaffected by brightness and KTFR. Audio improved recognition in general, which was expected.

The main highlight from the results is the contrasting effect of brightness on the intensity of happy and sad emotions; where brightness intensifies happiness, while darkness intensifies sadness. This result aligns well with previous findings showing that smiling faces were perceived brighter than frowns [Song et al. 2012]. Perhaps there is an expectation for happiness to be bright in order to be intense, based on some association between happiness and brightness. These results on intensity are robust as they were present in our carefully controlled laboratory experiment, and repeated and generalised with more characters and environmental conditions in our online experiment.

The effect of KTFR on emotion intensity was not detected here or in our preliminary work (Chapter 3), perhaps due to the high level of expressiveness of the cartoon characters overpowering the effect of shadow, or the fact that our lowest KTFR was not dark enough. Therefore, we extend the investigation into KTFR to our Realism Experiments (Section 5.3).

Our main result regarding appeal was that brightness increased appeal in general for all emotions, while KTFR did not appear to have an effect. However, in Chapter 3 we found lower appeal for dark shadows (KTFR), so we investigate the effect further in our Realism Experiments (Section 5.3).

Another interesting finding was that fear and anger were entirely unaffected by lighting in our Baseline Experiment. This goes against our hypothesis based on both hypotheses from research in psychology (according to Barchard et al. [2017], we would expect darkness to affect both anger and fear, or based on Xu et al. [2014], we would have expected all emotions to have higher intensity ratings under brighter conditions). We are unsure why this occurred, so we investigate further in our Realism Experiments (Section 5.3), with a larger range of KTFR as well as more realistic depictions of anger and fear from expression scans of real humans.

The addition of voice had little effect on intensity, and only affected appeal in the case of the dull neutral sentence, being less appealing with the monotone voice added. This result is in contrast to previous work investigating the perception of virtual characters, where audio overpowered visual information (e.g., [McDonnell et al. 2012; Hodgins et al. 2010; Hyde et al. 2013]). Future work could investigate into further details examining the influence of tone and semantic meaning of the audio.

The addition of a more complex background also had little effect in general, implying that our results should generalise to more natural scenarios with characters portraying emotions in the context of a scene. However, further testing with dynamic scene lighting would be necessary.

### **5.3 Realism Experiments**

The results of our Cartoon Experiments surprisingly showed that lighting and shadow had less of an effect on emotion recognition, and intensity than we had hypothesised. The cuteness and exaggerated expressions of the selected cartoon characters might be one reason. We are interested in the effect of shape/realism on the perception of emotion and appeal since there is evidence in previous work of differences in emotional response to realistic and stylised virtual humans (e.g., [MacDorman et al. 2009; McDonnell et al. 2012; Volante et al. 2016]). We are also interested if our results on intensity and appeal of stylised characters will generalise to realistic faces.



**Figure 5.11:** Samples of stimuli used in the Realism Experiments. From left to right shows stylisation levels: Realistic, Middle and Toon. Row 1: sadness male-100% brightness-1:1 KTFR, Row 2: sadness male-100% brightness-16:1 KTFR, Row 3: happiness female-100% brightness-1:1 KTFR, Row 4: happiness female-100% brightness-16:1 KTFR.

Apart from the differences found in studies on emotional response, there is also a known effect where shading information is used to retrieve the shape information of a 3D object [Bruckstein 1988]. Studies have shown that shadows can actually alter an expression from happy to sad, as is the case of Noh masks [Kawai et al. 2013]. We would like to test if shadows would influence the perception of emotion on different facial geometries in our experiments, and hypothesise that the presence of shadows will affect emotion recognition

for realistic as opposed to stylised faces, since shading highlights shape information.

### 5.3.1 Stimuli

For the stimuli in these experiments, we carefully chose 3 facial shapes that corresponded to distinct levels of stylisation, ranging from highly cartoon stylised (Toon) to a middle level of stylisation (Middle) to a realistic photogrammetry scan (Realistic), and 2 genders, male and female, for generalisability (Figure 5.11). The models were shared from Zell et al. [2015]'s online character repository. Each model was posed to convey a static neutral, angry, sad and happy expression. The three-point lighting environment and render pipeline were set for these new models in a similar manner to the Cartoon Experiments, using the combination of 3ds Max, Maya and Nuke software to ensure consistency between lighting conditions across the different models.

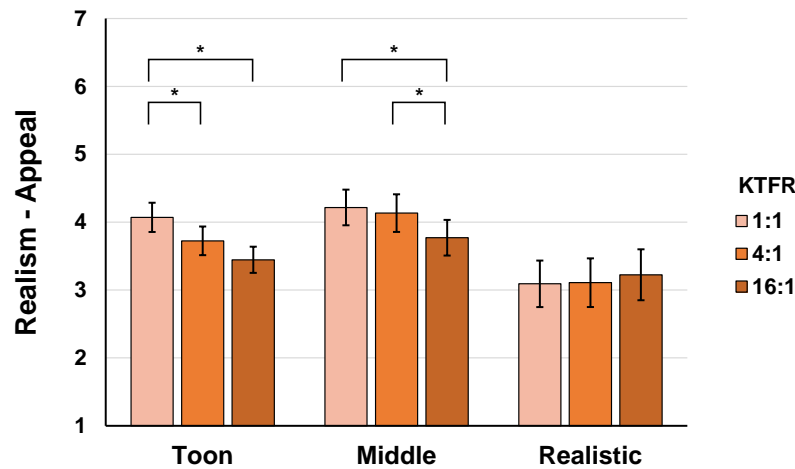
Since there were no major differences in perception of lighting between moving and still stimuli in the Cartoon Experiments, we focused on the perception of still images in these experiments, for simplicity. Additionally, since KTFR had little effect in the Cartoon Experiments, we extended the range to include a higher KTFR. Finally, there were no interactions between brightness and KTFR in the Cartoon Experiments, so we separate the conditions into two experiments (examining brightness level and KTFR separately) to lower the number of trials.

### 5.3.2 Experiment Design

Two groups of fifteen participants took part in these experiments (7F, 8M, aged 18-47, avg. 28, and 7F, 8M, aged 23-52, avg. 30). We conducted 2 within-groups experiments, one for key light brightness and one for KTFR. Each experiment lasted approximately 30 minutes. There were 144 trials in total in each experiment: 2 model genders x 3 shapes x 4 emotions (neutral, anger, sadness, happiness) x 3 lighting conditions (brightness in experiment 1 (100%, 50%, 25%) and KTFR in experiment 2 (1:1, 4:1, 16:1)) x 2 repetitions. An additional question on eeriness was included in these experiments, due to the use of realistic characters. All other experimental conditions were the same as before.

### 5.3.3 Results

Statistical analysis was conducted for **brightness** and **KTFR** separately, as separate repeated-measures Analysis of Variances (ANOVAs) for each of the three tasks: recognition, appeal,



**Figure 5.12:** Interaction between KTFR and Shape for ratings on Appeal. Star labelled lines point to significantly different means according to the post-hoc test. Error bars show standard error of the means.

and intensity. All other analysis details were as before. See Table A.6 and A.7 in the Appendix for the significant effects. Since Zell et al. [2015] investigated the effects of shape, emotion, and gender in their work, we only discuss these variables if our results differ to theirs, and focus on the main variables of interest to our study, namely brightness and KTFR.

**Brightness:** There was no main effect of brightness or interaction for emotion recognition or intensity. There was a main effect of brightness on appeal, where 25% was rated less appealing than 100% ( $p < 0.009$ ). There was also an interaction between brightness, gender, and shape for eeriness, where male Toon at 25% was rated more eerie than male Toon at 100%, male Middle at 100% and male Middle at 50% ( $p < 0.004$ ).

**KTFR:** KTFR affected emotion recognition, where 16:1 KTFR was less recognised than 4:1 ( $p < 0.022$ ). For intensity, there was an interaction between KTFR and Emotion, where darker shadows decreased the intensity of happiness ( $p < 0.0012$ ).

For appeal, there was a main effect of KTFR, where 16:1 was less appealing than 4:1 and 1:1 ( $p < 0.05$  in both cases). Further investigation showed the effect to be emotion specific, where anger at 16:1 was rated more unappealing than the other ratios ( $p < 0.0002$ ) and happiness at 16:1 and 4:1 were more unappealing than 1:1 ( $p < 0.0006$ ). There was no effect of KTFR on the appeal of neutral or sad. On further inspection, an interaction between gender, emotion and KTFR shows that the anger effect is coming from the female character and the happiness effect from the male. An interaction between shape and KTFR also occurred, which showed

no effect of KTFR on appeal ratings for Realistic (Figure 5.12). For Middle, 1:1 and 4:1 were more appealing than 16:1 ( $p < 0.0064$ ). For Toon, 1:1 was more appealing than 4:1 or 16:1 ( $p < 0.0011$ ).

A main effect of KTFR on eeriness showed 16:1 was more eerie than 4:1 or 1:1 ( $p < 0.0276$  in both cases), which was modulated by an interaction between gender, shape and KTFR, which showed this effect to be mainly coming from the male character.

### 5.3.4 Realism Experiments - Discussion

Emotion recognition was very high in general for all characters in these experiments (ranging 72% - 98%). We pushed the darkest KTFR to a ratio of 16:1 (compared to 4:1 being the darkest in the Cartoon Experiments) and found that this had an effect on the recognition, where very dark shadows significantly affected recognition rates. Therefore, care should be taken when using very dark shadows in practice to increase intensity, as the effect on intensity is less than we thought, while making the emotion less recognisable and the character less appealing.

Our hypothesis about dark shadows intensifying realistic shapes was not correct, as we found no effect of lighting on the angry or fear emotion in the Realism Experiments. We did, however, repeat our previous observations from the Cartoon Experiments about brighter conditions (in this case, KTFR) intensifying happiness.

In general, the darker the lighting condition (both for brightness and KTFR), the more unappealing. However, our main results of these experiments is that the appeal of shadow is also affected by stylisation level, as we observed a trend where the more realistic the character gets, the less effect the darkness of shadow has on appeal, as the Toon was considered only appealing for 1:1 KTFR, while middle was more appealing for 1:1 and 4:1, but for realistic, all three were rated equally. It should be noted though, that the realistic shape was rated less appealing in general than the others, regardless of the KTFR.

## 5.4 General Discussion

In this chapter, we have conducted a set of novel experiments investigating the role of portrait lighting on emotion on cartoon and realistic characters. We carefully designed a scene using a typical 3-point lighting setup that we could manipulate to produce a controlled set of stimuli, varying in key-light brightness and shadow intensity (KTFR).

We formed a number of hypotheses on how light would affect judgments of emotional expression, based on low and high level perceptual psychology research [Logvinenko and Maloney 2006; Xu and Labroo 2014; Barchard et al. 2017] and best-practice in the arts and cinematography literature [Thomas and Johnston 1995; Lasseter 1987]. Our results do not align with any one of these hypotheses completely, but have some similarities. For example, our results show that while brighter conditions intensify the expression of happiness, darker lighting conditions increase the intensity of a sad expression. This result is intuitive (e.g., based on evidence from the literature which associates brightness with emotional valence [Zhang et al. 2016; Song et al. 2012], and the fact that the dichotomous descriptor “bright” was found to be associated with happiness, and “dark” with sadness in the research by Barchard et al. [2017]), but has not been shown to directly affect perceived intensity of an expression before.

Surprisingly though, darker conditions did not intensify anger or fear for any of our characters emotions, across all experiments. This result goes against the hypothesis based on the dichotomous descriptors, as “dark” was found to be highly associated with anger and fear [Barchard et al. 2017], or well-known practice in cinematography for using low-key lighting to produce a gloomy mood [Pramaggiore and Wallis 2005]. Additionally, it does not align with the opposite hypothesis that brightness intensifies perceived emotions, due to activation of the hot emotional system [Xu and Labroo 2014].

In the case of anger, previous work has shown a strong relationship to the colour red [Fetterman et al. 2012]. Future work will investigate if the use of coloured lighting is more effective at heightening the intensity of anger.

For fear, previous work has shown a cross-cultural colour association with the colour black [Hupka et al. 1997] and so we are unsure why darkness had no effect on the perception of intensity of fear in our experiments. The authors suggested that the potential risk of nighttime darkness is the same everywhere, and thus black-fear is founded on our biology. Therefore, in practice, it might be the case that darkening the environment together with the character would be important for supporting fear in CG scenes, which could also be tested in future studies.

For the portrayal of fear and anger in film, our studies imply that using dark shadows to increase the perceived emotion of a character is not effective, however, it is possible that dark shadows could be used effectively in other ways to increase intensity throughout a scene,

which will be studied in future work.

For emotion recognition, we found it to be generally robust against lighting conditions, except for the extreme case of 16:1 KTFR which reduced recognition. Therefore, care should be taken when using extreme dark shadows to avoid negative effects on emotion recognition, particularly since dark shadows were not shown to have an impact on intensity ratings.

Interestingly, previous work [Meier et al. 2004] has shown evidence for an automatic association between brightness and affect for categorisation of words. We did not record reaction times to our emotion recognition task in this experiment, but this could be an interesting future direction to determine if this automatic association extends to more complex stimuli, such as those shown in our experiment.

Throughout our experiments, lighting had a much bigger effect on the appeal ratings than on emotional intensity or recognition. Across the board, increasing the brightness of the key light or lessening the key-to-fill ratio (lighter shadows), increased the appeal. We also found that brightness had little effect on eeriness, whereas lightening the shadows reduced eeriness ratings, which could be an important result for artists trying to reduce “uncanny valley” effects of their virtual characters. However, it should be noted that, for characters with realistic appearance, lightening the shadow did not improve appeal, so key-light brightness alone should be used to enhance appeal in those cases. These guidelines could be used to broaden the definition of the “appeal” principle of animation [Thomas and Johnston 1995] to include lighting.

Even though darkness increased intensity for sadness, it also reduced appeal, implying that there is not a direct correlation between how lighting is used for appeal and intensity. In the case of sadness, a trade-off between using darkness for intensity or lightness for appeal is necessary.

It is interesting to note that using bright scenes with no directional light is currently a popular setup in many applications with virtual characters (such as pre-schooler cartoon television series). Our results confirm that the addition of dark shadow does lower appeal for cartoon characters in particular. On the other hand, we found that a lower level of brightness (50%) can be equally appealing and intense for most emotions. This is a valuable guideline for artists that even negative emotions are more appealing under bright lights, which is something that was not shown previously.



---

Finally, our one-of-a-kind extensive set of experiments covered a wide range of brightness levels and shadow depth, as well as other important tools in storytelling (audio, background and movement) and an essential factor of stylisation in character design. Nonetheless, we only tested a fraction of the enormous character lighting parameter space. We believed one consideration that deterred other researchers from empirical studies of character lighting was the labour and time required to prepare the stimuli and conduct each experiment as we experienced it first-hand. In the next chapter, we explore a new experiment paradigm, utilising the method of adjustment and real-time graphics enabling us to explore the same lighting space in a shorter amount of time and hone in the results with exact parameter values usable by artists.



## Chapter 6

# Perceptual Lighting Tool

### 6.1 Introduction

At the end of the previous chapter, we attempted to derive character lighting guidelines through rigorous and labor-intensive rating-scale experiments exploring the parameters of light. This chapter <sup>1</sup> proposed an alternative approach based on the method of adjustment, similar to common lighting tools in 3D content creation software, but with the power of modern real-time graphics. The method of constant stimuli was used extensively early on in this thesis as it was easy to administrate, could the variation of the sensory sensitivity of the observers and covered the range of the parameters; however, the method of constant stimuli is not suitable for threshold estimate as it requires an enormous number of trials [Pelli and Farell 1995; Gescheider 2013]. The method of adjustment, while being more intuitive and able hone in smaller thresholds in less time, can produce unreliable results influenced by the participant's personal experience, especially when the task description is ambiguous [Pelli and Farell 1995; Gescheider 2013]. That being said, as we wanted to refine our previous findings in a shorter experiment, we decided to take the advantage of the method of adjustment with carefully defined experiment criterion.

This new framework allows users to interactively adjust lighting parameters and instantly assess the results on the animated characters, instead of having to wait for the render to complete. We show that using our system can help to speed-up experiment duration, allowing the experimenter to investigate more dependant variables. We focused specifically on brightness and shadow as a proof-of-concept, but more importantly, we wanted to validate the results of our previous experiments. We aimed to show that 1) the

---

<sup>1</sup>The content of this chapter was published in the ACM Symposium on Applied Perception 2020 (SAP'20)

prototypes of real-time lighting tools used in this experiment would promote the utilisation of visual perception in the content creator community, and 2) the two experiment paradigms would complement each other and accelerate the advancement of research in this field.

## 6.2 Tool Development

Using the results of our work throughout the thesis, we aimed to develop a new real-time lighting tool for content creators, based on perception of brightness and KTFR, that would allow them to achieve lighting conditions quickly and easily, maximising appeal and emotional intensity of their CGI character.

The main insight is that the tool has a power-of-two adjustment which is equivalent to the conventional unit of *stop*, commonly used in cinematography and photography (a “stop up” is doubling and a “stop down” is halving a light quantity.) Chapter 4 has confirmed the doubling and halving to be the natural perceived interval where the human visual system can discriminate brightness and shadows. The lighting tool was designed specifically to test if limiting the selections to a small subset of data points evenly sampled from the perceived lighting space can further speed-up the experiment process, by allowing users to only explore lighting conditions that give perceptually dissimilar results.

### 6.2.1 Real-time Lighting & Rendering

We adopted the three-point lighting design, with the same light directions and sizes that were used to illuminate the stimuli in Chapters 4 and 5. However, the previous setups deployed three area lights (key, fill and rim), but in the new real-time graphics environment of the Unreal Game Engine version 4.21 (UE4)<sup>2</sup>, an individual area light was replaced by a grid of point lights for efficiency while retaining comparable shadow quality. The brightness of the key and fill lights illuminating the character were altered directly by the user and the changes were rendered in real-time.

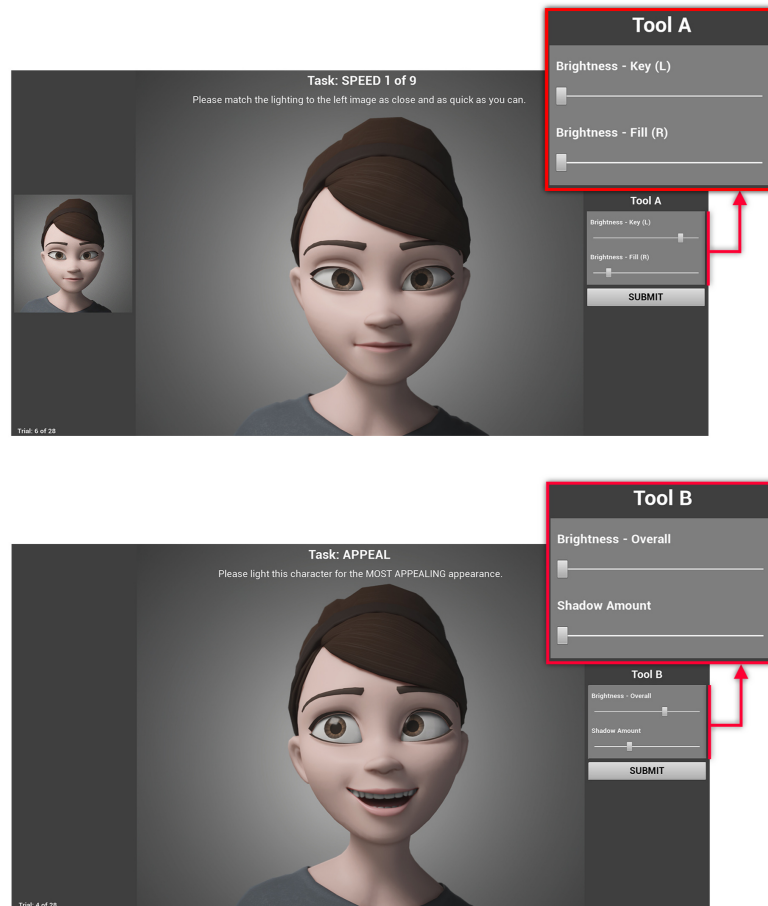
### 6.2.2 Lighting Control & Interface

We delivered two lighting tools, one was our control condition with the traditional continuous adjustment (tool A) and the other was our proposed perceptual tool with the power-of-two adjustment (tool B) implemented based on parametric equations that

---

<sup>2</sup><https://www.unrealengine.com>

captured the psychophysical model of the perceived brightness and shadow space of character lighting (please see Chapter 4, Section 4.5 for details). Users changed the lighting conditions of the scene via the one of two provided lighting control interfaces, *Tools A or B*, located on the right side of the screen and the result was immediately rendered in the middle of the screen (Figure 6.1).



**Figure 6.1:** Experiment setup in Unreal Engine 4. Top: digital lighting Tool A being used in the Speed Task. Bottom: perceptually-based Lighting Tool B being used in the Appeal Task.

### Tool A

Tool A had two independent control sliders (Figure 6.1 left), one for key light brightness (light coming from the left side of the screen) and the other for fill light brightness (light coming from the right side of the screen). The tick on each slider could be moved continuously in the range between 12% and 110%. The 100% is set to be the maximum possible intensity before highlight clipping occurs and the range 12% to 110% was set to match the range of Tool B.

The numeric values of the selected values were not shown on the sliders for either tool. We took up the definition of *100% brightness* from Chapter 4 which was the maximum possible intensity before highlight clipping occurred.

### Tool B

Tool B also had two control sliders (Figure 6.1) right, one for the overall brightness and the other for the shadow amount, defined by the key-to-fill ratio (KTFR). The KTFR is the proportion of the key light brightness to the fill light brightness that project on to the character face. High KTFR means there is less illumination, and hence, more shadow on the fill side of the face. The sliders of Tool B could only select a finite combinations of 4 levels of overall brightness (100%, 50%, 25% and 12.5% and 4 levels of KTFR (1:1, 2:1, 4:1 and 8:1). For each pair of brightness and KTFR, the intensity of key and fill light in the scene were adjusted automatically according to an approximation of the parametric model described in Chapter 4. Note that the model maintains consistent iso-brightness level for different KTFRs, and hence the individual light intensity can vary from 12% to 110%. Please see Chapter 4, Section 4.5 for computation details.

## 6.3 Experiment

We conducted an experiment to test the usability of the new lighting tool, as well as to validate the results of our experiments in Chapter 5 using a different experiment paradigm, where participants adjusted the lighting themselves to achieve appeal and intensity. Our experiments were run on the UE4 platform.

### 6.3.1 Character

In order to compare between the results of the new method-of-adjustment tools and those of the method-of-constant-stimuli experiments done in the last chapter, we picked the same animated CG character *Mery* for our test model, and we closely followed the steps described in Chapter 4 to recreate a similar CG scene used to render earlier stimuli. The animations of *Mery* displaying neutral expression, happiness, sadness stored in Alembic<sup>3</sup> geometry cache files were imported to the UE4 scene editor where the materials and the lighting were setup. We duplicated the types and values of shaders from the previous stimuli. Each animation clip lasted 2-3 seconds, and during the experiment, the animation was looped.

---

<sup>3</sup><http://www.alembic.io>

### 6.3.2 Presentation

The major difference from the earlier experiments was, instead of rendering the stimuli to movie clips with predefined set of lighting conditions, in this experiment, the participants were asked to alter the lighting parameters and observe the changes in the final render interactively.

All lights were white with the white-point temperature and peak luminance of experiment monitor, Dell UP2713H, calibrated to 100% sRGB color gamut, 6500k white point, and  $80 \text{ cd/m}^2$  brightness. The entire experiment was conducted in a completely dark room in which the character and the tool were displayed on a 100% sRGB calibrated monitor. The participant instructions were outlined on a piece of paper that was explained clearly to the participant who was also given a copy before the experiment started.

### 6.3.3 Design

The experiment was divided into two blocks, one for each of the lighting tools, A and B, and we counter-balanced the ordering of the blocks. In each block, the experiment started with a training session, in which the participant was explained how each slider modified the lighting in the scene. The participant took as much time as needed before continuing to the actual experiment.

After the training, participants always started with either the Intensity or Appeal task, counter-balanced, each consisting of two trials, happy and sad, also counter-balanced. Next, the participant completed the speed & accuracy task and a usability questionnaire. There were 26 trials in total (13 for each block), and participants were allowed to take a short break before the beginning of each trial.

For the first two tasks, we chose to test happiness and sadness because they were studied extensively previously. For the final task, we used the neutral animation.

### 6.3.4 Intensity Task

For the intensity task, the participant was instructed to “*please light the character for the happiest appearance*” for the happy animation, and to light the character for the saddest appearance for the sad depiction. The instruction was also displayed on the top of the screen. The participant was also advised to take as much time as needed. The final selected

slider values were recorded for each of the two trials (one for sad and one for happy presented in random order).

### 6.3.5 Appeal Task

The participant was instructed to *“please light the character for the most appealing appearance”*. The instruction was also displayed on the top of the screen. They were also advised to take as much time as needed and explained the definition of appeal as *“If a character is appealing then you would be captivated by a movie with that character in the leading role, and would like to watch more of them.”* [Kokkinara and McDonnell 2015]. As before, the final selected slider values were recorded for each of the two trials (one for sad and one for happy presented in random order).

### 6.3.6 Speed & Accuracy Task


The participant was instructed to *“please match your lighting to the target image as close and as quick as possible.”*. The instruction was also displayed on the top of the screen. There were 9 trials in total for this task. The 9 targets, 3 levels of brightness (100%, 50% and 25%) times 3 levels of KTFR (1:1, 2:1 and 4:1), were regular samples of the character lighting perceptual space detailed in Chapter 5. The slider values and the time taken to complete each trial were recorded. This task simulated the real-world use case of a lighting artist performing a set of lighting assignments under a time-constraint.

### 6.3.7 Usability Questionnaire


After completing the three tasks, the participant was asked to answer a short perceived usability questionnaire designed based on the Usability Metric for User Experience (UMUX) [Finstad 2010]. UMUX was designed to replace the common 10-item, five-point Likert scale System Usability Scale (SUS) with just four seven-point questions written in less ambiguous language that still conforms to the ISO 9241-11 (1998) definition of usability (overall usability, effectiveness, efficiency and satisfaction), and has been proven to be reliable and highly correlated to SUS [Berkman and Karahoca 2016]. Each question was rated on a seven-point Likert’s scale from strongly disagree to strongly agree (Figure 6.2).




This lighting tool meets my requirements to complete the task.




Using this lighting tool is a frustrating experience.



This lighting tool is easy to use.



I have to spend too much time correcting things with this lighting tool.



Continue

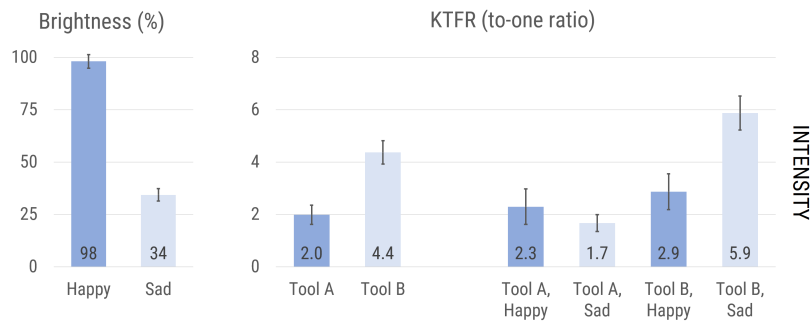
Figure 6.2: Usability Questionnaire

### 6.3.8 Participants

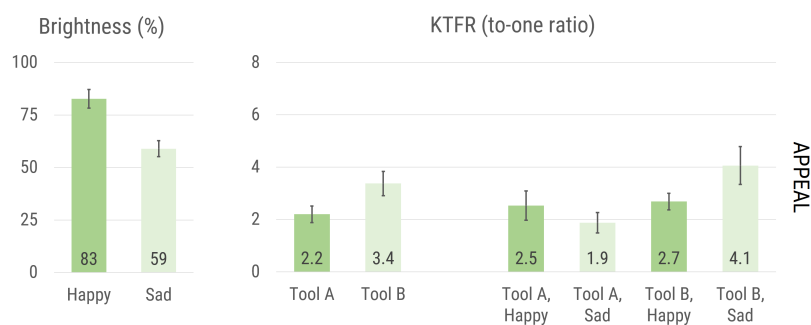
Sixteen participants took part in the experiments (7 females, 9 males, aged 18-37, avg. 28), all with normal or corrected to normal vision, and recruited primarily via university student and staff mailing lists from the same population as the earlier experiments. They had different degrees of experience in CG content creation and were naïve to the purpose of the experiment. A €5 voucher was rewarded to each volunteer for taking part.

## 6.4 Results

For each task, the chosen values of lighting parameters from all participants were analyzed together using a two-way repeated-measures ANOVA with the within-group factors of emotion (happiness, sadness) and tool (A, B). In order to be able to compare results across the tools, we generated KTFR values using the key and fill slider values for Tool A. We show the main effects in Figures 6.3 and 6.3, and discuss significant results in the text. We ran post-hoc analysis using Tukey's Honestly Significant Difference (HSD) tests throughout.



**Figure 6.3:** Main effects and interactions of intensity task.



**Figure 6.4:** Main effects and interactions of appeal task.

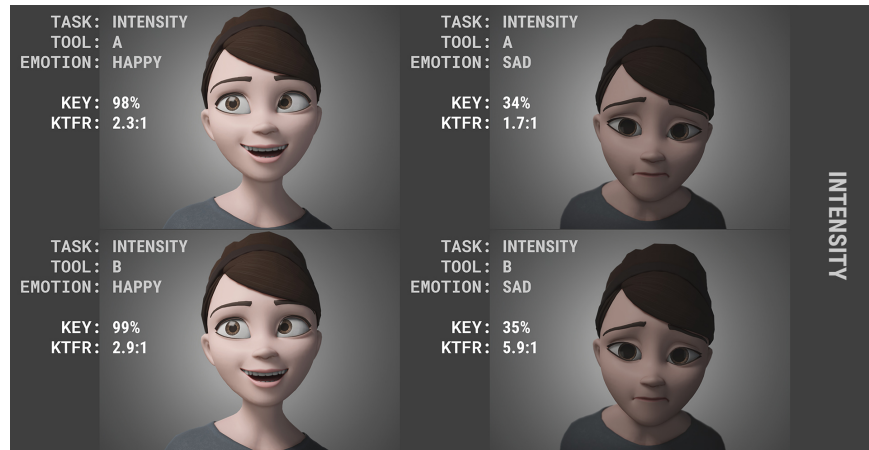
### 6.4.1 Intensity

Participants were asked to light the character for the most intense appearance and took on average 37.40s to complete each trial. For brightness, we found a main effect of emotion ( $F(1, 15) = 143.89, p < 0.0001$ ) where happiness (98%) was illuminated brighter than sadness (34%), as expected. A main effect of tool was not found, indicating that participants did not differ in their brightness selections using either tool.

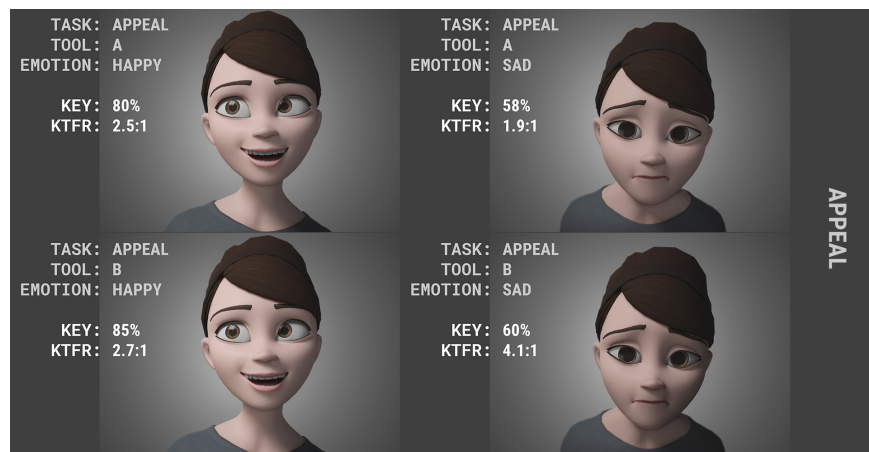
For KTFR, there was no main effect of emotion, indicating that the same KTFR levels were considered intense for happiness and sadness equally. A main effect of tool ( $F(1, 15) = 17.03, p < 0.001$ ) showed us that participants used higher KTFR with Tool B (4.4:1) than Tool A (2:1), perhaps due to the discrete allowed values of Tool B. However, the mean KTFR from Tool B was almost two levels higher than the mean from Tool A indicating that different levels of discretization are needed for a conclusive explanation.

An interaction between tool and emotion ( $F(1, 15) = 17.22, p < 0.0003$ ) for KTFR showed that sad Mery was lit at higher ratio with Tool B (5.9:1) than Tool A (1.7:1). Also, when using Tool B, participants chose higher KTFR for sad (5.9:1) than for happy (2.9:1) ( $p < 0.002$ ).

See Figure 6.5 for a visualisation of the average chosen values for Tool A and Tool B, and Figure 6.10 for the individual values selected by participants.



**Figure 6.5:** Screenshots from our real-time system showing the averaged key light and KTFR values chosen by the participants for the intensity task.



**Figure 6.6:** Screenshots from our real-time system showing the averaged key light and KTFR values chosen by the participants for the appeal task.

### 6.4.2 Appeal

Participants were asked to light the character for the most appealing appearance and took on average 35.65s to complete each trial. For brightness, the analysis showed a main effect of emotion ( $F(1, 15) = 17.27, p < 0.001$ ) where happiness (83%) was lit brighter than sadness (59%). For KTFR, a main effect of tool ( $F(1, 15) = 4.61, p < 0.049$ ) showed that participants chose higher KTFR with Tool B (3.4:1) than Tool A (2.2:1). When asked to light the sad character, participants selected higher KTFR with Tool B (3.4:1) than Tool A (2.2:1) (as indicated by the tool x emotion interaction ( $F(1, 15) = 7.02, p < 0.019$ )). There was no main

effect of emotion on the choice of KTFR, indicating that appealing KTFRs were the same for sad and happy. See Figure 6.6 for a visualisation of the average chosen values for Tool A and Tool B, and Figure 6.11 for the individual values selected by participants.

### 6.4.3 Speed

Participants were asked to match the lighting of a character to provided targets. The ANOVA showed a main effect of tool ( $F(1, 15) = 11.21, p < 0.005$ ) where, on average, that participants completed the task significantly faster when they used Tool B (15.47s) than they did with Tool A (18.73s), indicating that the perceptually-based tool indeed reduced the experiment time for participants as they were not exploring parameters that had no perceptible difference, like in Tool A.

### 6.4.4 Accuracy

We also tested the accuracy of both tools by mapping the results into the perceptual space using the parametric model from Chapter 4 and then measuring the perceived dissimilarity distances between the results and targets. The ANOVA analysis showed the accuracy of Tool A and B are not significantly different.

### 6.4.5 Usability

The ANOVA Analysis of the perceived usability questionnaire showed no significant differences of perceived effectiveness, satisfaction, efficiency and the over experience between Tool A and Tool B. In general, effectiveness was rated high, satisfaction was good, overall ease of use was high, and the efficiency was better than average (Figure 6.7).

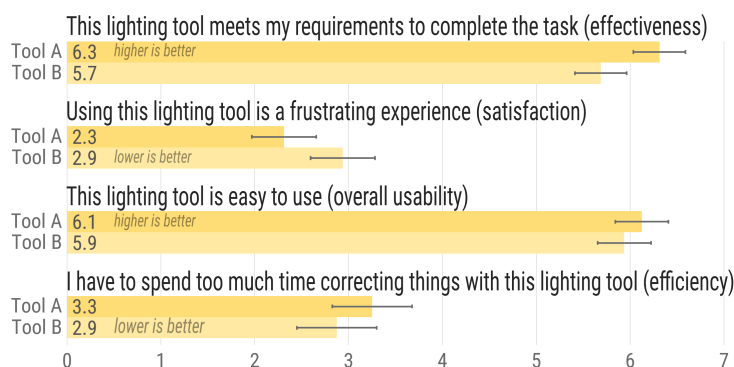


Figure 6.7: Usability questionnaire ratings

## 6.5 Validation

Previous work has discussed a need to navigate through the perceptual space of appeal quicker and more effectively [Zell et al. 2019]. This study served as a bridge between the recent findings from the constant stimuli to the new paradigm of interactive experiment design utilising the modern graphics hardware and game engines. Here, we compare our new results (Figures 6.10 and 6.11) to the results of the previous chapter that employed the method of constant stimuli to determine if we arrive at similar conclusions in a shorter experiment.

### 6.5.1 Brightness

In the previous study, for happiness, brightness level 100% was rated more intense and more appealing than 25%. Our new results agreed as participants selected high levels of brightness for both appeal (98-99%) and intensity (80-85%). For sadness, previous work found an effect of brightness on intensity, where the emotion was rated more intense at 100% than at 25%. Our result here agrees, and we additionally hone-in on the optimal brightness level for intensity of sadness at 34%. In the previous study, there was no difference between ratings of appeal for the 100% or 25% brightness levels. Our result here complements that finding and also provides the optimum brightness for appeal of sadness of 58-60%.

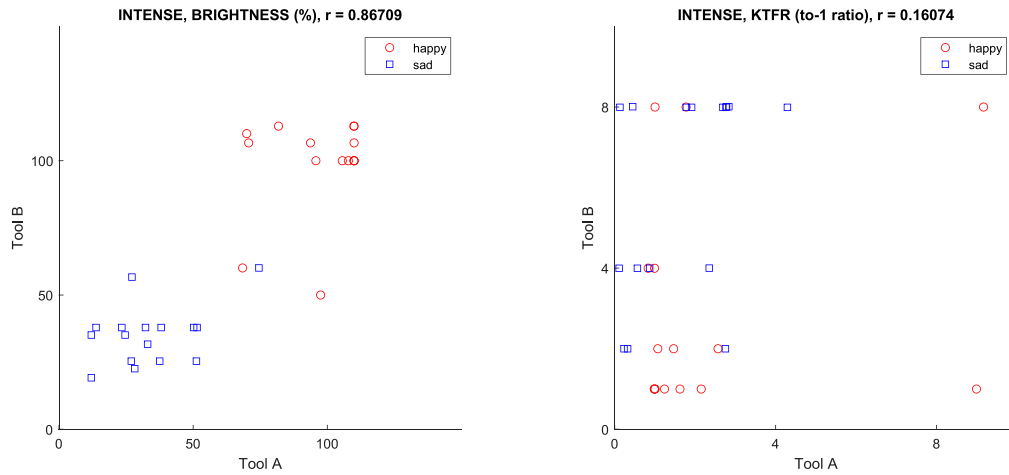
### 6.5.2 KTFR

In terms of shadow amount, the previous study did not find any conclusive evidence that indicated the effects of KTFR on the intensity and appeal of happiness and sadness which contradicts the artistic practice of using shadow to intensify the drama. However, our result gave us ranges of KTFR that the audience preferred for emotional and appealing characters. This could be influenced by their past experience and the debate on whether top-down or bottom-up processing is better for lighting perception is worth investigating in future work.

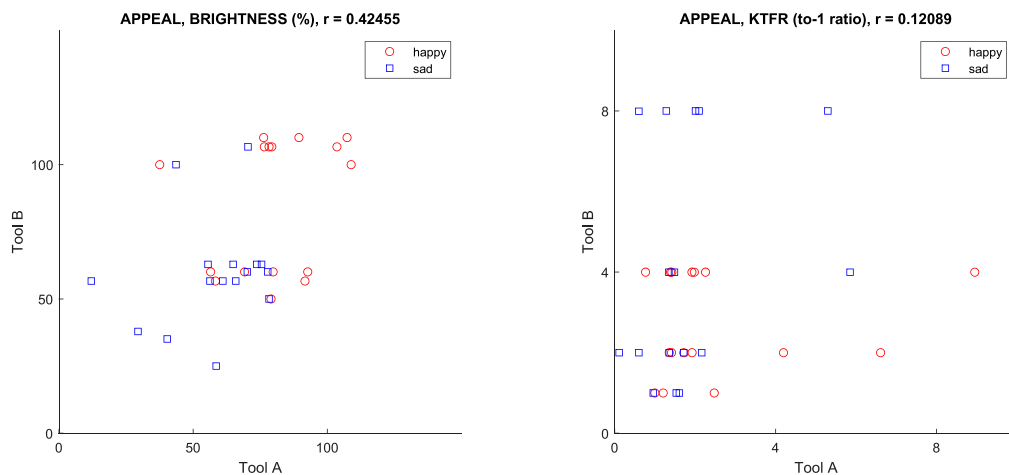
### 6.5.3 Time

Lastly, time taken to explore brightness and KTFR parameters was 36-37 seconds per trial for our method-of-adjustment experiment, and 35-40 seconds per trial for our previous method-of-constant-stimuli experiments in Chapter 5. Although the speed improvement seems negligible, note that we instructed our participants to take as much time as they

needed. Moreover, if we had matched the thresholds of the stimuli used in the experiments conducted in Chapter 5 to those that we used in Tool B, our previous experiment would have taken more than 60 minutes to complete, rendering the results less reliable due to fatigue. With the same amount of time, our experiment paradigm in this Chapter could explore a broader range of parameters and produced results with exact thresholds.



**Figure 6.8:** Correlation between Tool A and B of the individual parameter values of the appeal task selected by the participants for the intensity task.



**Figure 6.9:** Correlation between Tool A and B of the individual parameter values of the appeal task selected by the participants for the appeal.

## 6.6 Discussion

In this chapter, we presented real-time lighting tools that could be used in method-of-adjustment perceptual experiments investigating character lighting parameters. We evaluated the tools by allowing users to interactively alter the light brightness and shadow on animated cartoon characters. We included separate tasks for lighting for speed and accuracy, lighting for appeal, and lighting to improve emotional intensity. We validated the new results by comparing them to our old results in Chapter 5 that used a method-of-constant-stimuli with a Likert-scale response. It is worth to note that the previous experiment yielded recommendations of relative lighting values, for example, 100% brightness is more appealing than 25% brightness but our new experiment identified the absolute thresholds. We expect that our method of adjustment will speed-up investigation of experiments with large ranges of parameters. However, one limitation of method of adjustment is that the decision task gets more difficult as the number of dimensions increase, so we will investigate its usability with 3 or more sliders in future work. In terms of usability, participants found the tools easy to use and met their requirements to complete the tasks, which indicates that this method could prove useful for future analysis of a wider range of lighting parameters, allowing researchers to analyse the perception of virtual characters more rapidly.

Finally, we also found that the perceptually-based tool is comparable to the traditional lighting tool in terms of usability and accuracy. However, the new tool reduced the time of performing repetitive lighting tasks which would be worth developing further for practical use, particularly, in the animation production of episodic or daily content, where quality and speed are essential. Another interesting discovery was the results of Tool A and B were highly correlated ( $r = 0.87$ ) only in the case of brightness for the intensity task (Figures 6.8 and 6.9). This evidence supports the findings of the previous chapter that brighter is happier and darker is sadder, regardless of the tools given to create the scene.

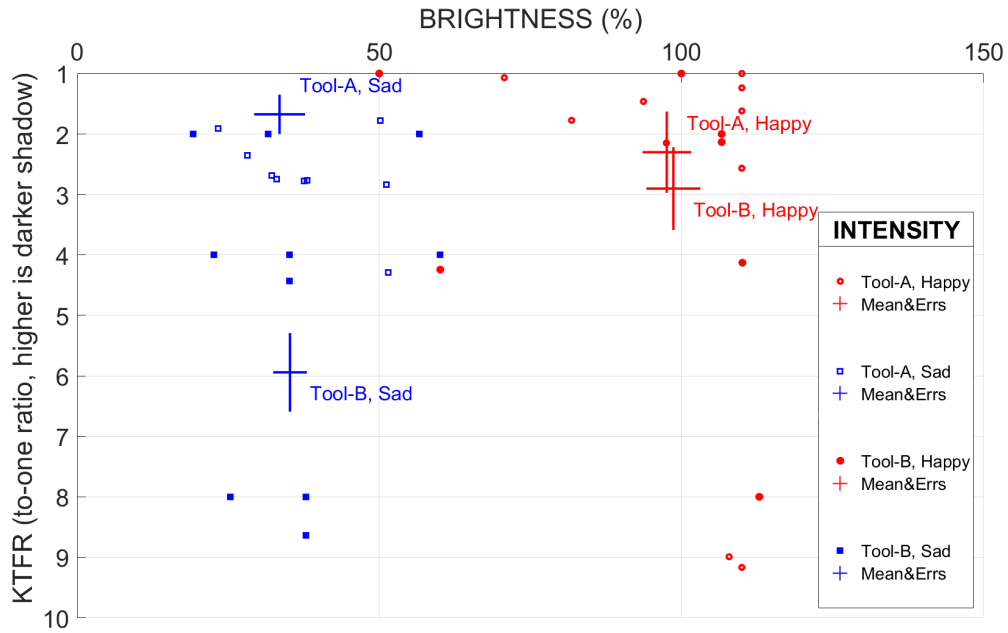


Figure 6.10: Mean and individual parameter values of the appeal task selected by the participants for the intensity task.

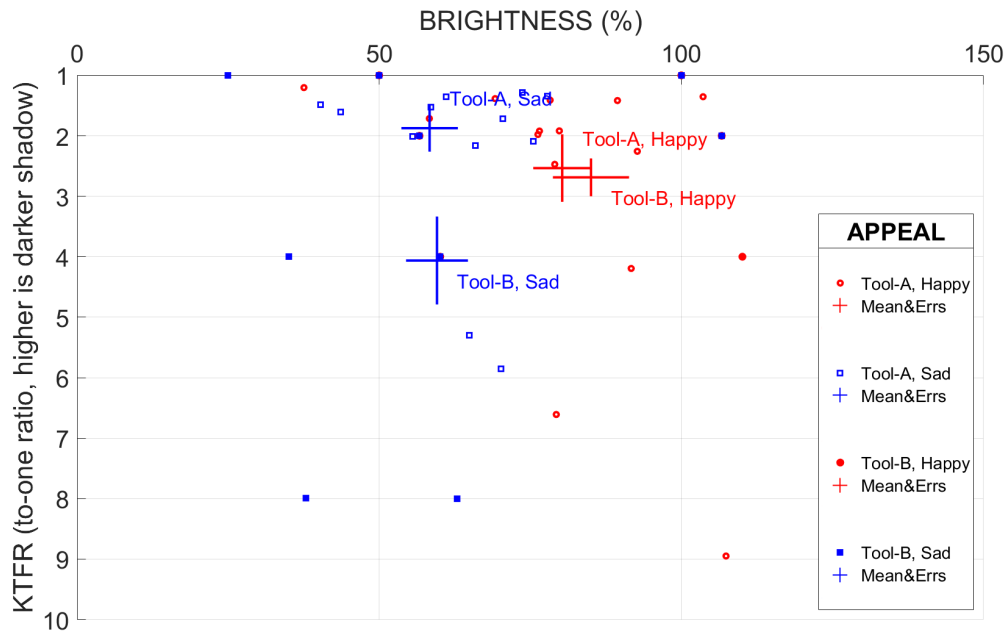


Figure 6.11: Mean and individual parameter values of the appeal task selected by the participants for the appeal task.



## Chapter 7

# Conclusion

The purpose of this thesis was to empirically evaluate the effect of conventional practices of CG character lighting on the perceived emotion and appeal of virtual characters (Chapters 3 and 5) and to standardise the lighting design guidelines for virtual characters grounded by collected psychophysical data and analysis (Chapters 4 and 5). Our study also breathed new life into traditional experiment techniques, the multidimensional scaling (MDS) analysis and the method of adjustment, that proved to be highly capable in solving our research questions (Chapters 4 and 6). The latter technique could also be deployed in conjunction with the lighting guidelines as a perceptually-based lighting tool (Chapter 6). This chapter summarises the key findings of this thesis and presents some exciting paths this work may lead to for future research.

### 7.1 Contributions

This work provided new experiment results quantifying the perceived emotion and appeal of CG characters illuminated by a range of lighting parameters, as well as a method to derive a parametric model describing the perceptual space. These contributions found applications in both industry practice and academic research.

#### 7.1.1 The Character Lighting Guideline

We obtained new insights into the influence of several important parameters of CG character lighting that were compiled into a set of general guidelines for lighting design for expressive and appealing characters. The guidelines could be used to improve the overall quality of virtual avatar applications created by professionals and non-professionals alike.

- **Brightness level** – one of the significant findings of this thesis was the importance of brightness for the appeal of a character. Higher brightness level made the character appear more appealing for all type of characters, from stylised cartoon to realistic proportion regardless of the emotion portrayed by the characters. This outcome agreed with the conventional practices in film and theories in psychology. We also learned that the power-of-two technique called *stop up* and *stop down* in cinematography which doubles or halves the intensity level of a light respectively was an adequate measure for perceptible levels of brightness, confirmed by our rigorous psychophysical experiments. Moreover, with the new real-time method-of-adjustment, we could pinpoint the specific levels of brightness that were suitable for happy and sad expressions, for both emotional and appealing appearances.
- **Shadow depth** – for practical purposes, we decided to use the cinematic definition *key-to-fill ratio* (KTFR) as the gauge for shadow intensity, and hence the contrast between the key and fill sides of the character's face. We found that less shadow was more appealing in general, but dark shadow surprisingly had little to no effect on either the appeal or the emotion intensity, contradicting the film conventions of adding shadow to raise the tension, such as in *Film Noir*.
- **Direction** – another puzzle encountered in this work was the light-from-below, believed to convey unease feeling, from both stage lighting and visual perception standpoints. However, the main effect of direction was not found in our work.
- **Shading** – shape from shading theory emphasised the importance of surface information in recognising faces and expressions, and it was evident in the toon-shaded characters that were perceived as much less intense than the CG-shaded counterparts in our study.

### 7.1.2 The Psychophysics of Character Lighting

Besides creating lighting guidelines which was the main interest of this thesis, we also advanced the frontier of the virtual avatar perception research.

- **Lightness discrimination of complex stimuli** – Using the MDS analysis, we demonstrated for the first time the capability of the visual system in judging character lighting conditions, which were much more sophisticated than the typical stimuli used

in low-level perception studies. When the scene is too dark, humans become poor at discounting the shadows.

- **The psychophysical model** – we presented a parametric model describing the perceptual space of brightness and shadow of character lighting. The perceived brightness and shadow were tightly coupled and cannot be altered independently in a character lighting scene. However, the proposed mathematical relations could be used to level the perceived brightness of characters illuminated to display different shadow depth which has a potential use-case in animated film productions.

### 7.1.3 Method of Adjustment with Real-time Graphics

The method of adjustment was not new, but it was typically employed to perform simple tasks such as image manipulations due to the technology limitation. The modern graphics hardware and rendering algorithms has brought new possibilities to the old technique.

- **Perceptual lighting tool** – Leveraging the cutting-edge graphics hardware and the advanced game engine Unreal Engine 4 (UE4), we developed a new lighting tool based on the parametric model. We proved that the perceptual lighting tool was comparable to the traditional tool in terms of usability and accuracy, but had the potential to speed up repetitive lighting tasks common in episodic animated TV series.
- **Lighting parameter exploration** – we used the method of constant stimuli in the early experiments because producing industry-relevant results required high-fidelity CG scenes rivalling the blockbuster standard. Fortunately, with the latest version of UE4 and the interactive lighting tool just developed, we were able to create high-quality renders of an animated character in real-time. We showed that the method of adjustment could be utilised to explore the same lighting parameters investigated early by the method of constant stimuli, but with better precision, shorter time and less fatigue.

## 7.2 Future Work

Although we met the goal of the thesis by conducting a set of character lighting guidelines for brightness and shadows, we hope to further evaluate some of the parameters in more detail in the future. We also wish to generalise our findings over a broader range of characters and

applications, as well as to better explain the disagreements with the conventional practice of lighting in the art community.

We showed that light direction did not significantly influence the perception of emotion. However, theatrical lighting design typically deploys light from below to illuminate a sinister character. Exploring further into other traits of a character, besides emotion and appeal, could result in new indications.

Regarding shading, there was a significant decline in intensity in one of the test conditions that minimised the amount of shadows or shading information. These findings suggest the importance of shadowing and lighting on the perception of emotion intensity. With that being said, a change in shadow, however, does not significantly alter the recognition of emotion, which is surprising. It would be interesting to revisit the shape from shading theories and design new experiments to scrutinise the discrepancy. McCloud [McCloud 1994] also raised an interesting observation in his book, concerning the simplification of a face and the increase in expressivity. Future studies on the issue will improve the understanding of stylised virtual character design.

We have devoted a considerable amount of our work towards the study of brightness and shadows. However, to derive a more comprehensive set of lighting guidelines, there are still other parameters that should be explored in future experiments such as colour concerning both the light temperature and skin tone parameters, and well as the quality of light (soft light vs. hard light) produced from different light types and sizes. We also would like to examine other configurations of light placement in relation to the character or the camera which also lead us to the question of shot composition, another substantial storytelling device. The exploding number of related parameters was proven to be a challenge in designing a concise, self-contained experiment that did not cause any fatigue to the participants. We expect that our proposed method-of-adjustment lighting tool will speed-up the investigation of experiments with broad ranges of parameters. However, one limitation of the method-of-adjustment is that the decision task gets more difficult as the number of dimensions increase. In future work, we would like to investigate its usability with the current implementation and experiment design or our real-time lighting setup when users have control of three or more parameters (e.g., shadow, brightness, direction, colour, etc.). We also would like to improve our methodologies of character lighting studies with the adaptation of human-in-the-loop and optimisation techniques found in the sequential line search approach [Koyama et al. 2017]. Finally, the methods of the adaptive

sampling and machine learning with crowd workers from [Lagunas et al. 2019] could potentially be use to generalise our results of both the lighting parameter and perception judgement.

Due to the enormous number of parameters of character lighting, we have experienced first-hand the complexity of the problem and the difficulty to design a careful experiment. Our work has broken new ground in assessing the art of lighting with empirical methods of psychophysics and proven theories in visual perception. We hope that future researchers will benefit from the foundation we have laid down in this thesis and follow us in this endeavour to truly understand the art of lighting from the perspective of human perception.



## **Appendix A**

# **Summary of Main Effects and Interactions**

**Table A.1:** Chapter 5 - Baseline Experiment:  
main effects and interactions with post-hoc analysis.

<b>Recognition</b>		
Effect	F-Test	post-hoc
Emotion	$F(5,56) = 4.07, p < 0.006, \eta^2 = 0.23$	Anger recognized more than fear ( $p < 0.009$ ).
Character*Emotion	$F^*(2.81, 39.41) = 3.37, p^* < 0.031, \epsilon = 0.70, \eta^2 = 0.19$	For Mery, anger and happiness more recognized than sadness ( $p < 0.04$ ).
Emotion*Brightness	$F(8,112) = 2.71, p < 0.01, \eta^2 = 0.16$	No interpretable differences in post-hoc tests.
Emotion*Brightness*KTFR	$F(16,224) = 2.29, p < 0.004, \eta^2 = 0.14$	Fear better recognized at 100% brightness, 1:1 KTFR than 25% brightness, 4:1 KTFR ( $p < 0.02$ ).
<b>Intensity</b>		
Effect	F-Test	post-hoc
Emotion	$F(3,42) = 28.03, p < 0.0001, \eta^2 = 0.67$	Anger the most intense ( $p < 0.0009$ ), followed by happiness ( $p < 0.009$ ).
Character*Emotion	$F(3, 42) = 7.09, p < 0.0006, \eta^2 = 0.34$	For sadness, Franklin was significantly more intense than Mery ( $p < 0.001$ ).
Emotion*Brightness	$F^*(2.92, 40.82) = 3.42, p^* < 0.028, \eta^2 = 0.20, \epsilon = 0.49$	For happiness, 100% brightness more intense than 25% ( $p < 0.004$ ).
<b>Appeal</b>		
Effect	F-Test	post-hoc
Brightness	$F(2,28) = 9.86, p < 0.0006, \eta^2 = 0.41$	100% and 50% brightness more appealing than 25% ( $p < 0.004$ ).
Character*Brightness	$F(2, 28) = 7.38, p < 0.0027, \eta^2 = 0.35$	Franklin at 25% less appealing than all except Mery at 25%.
Emotion*Brightness	$F(8,112) = 2.95, p < 0.005, \eta^2 = 0.17$	For all emotions except sadness, 25% brightness less appealing than 50% and 100% ( $p < 0.002$ ). For sadness, 25% only less appealing than 50% ( $p < 0.012$ ).
Brightness*KTFR	$F^*(1.73, 24.25) = 3.72, p^* < 0.045, \eta^2 = 0.21, \epsilon = 0.43$	No interesting significant differences in post-hoc tests

**Table A.2:** Chapter 5 - Online Experiment:  
main effects and interactions with post-hoc analysis.

<b>Intensity</b>		
Effect	F-Test	post-hoc
Characters	$F(3, 174) = 84.64, p < 0.0001, \eta^2 = 0.60$	Franklin most intense, followed by Malcolm, Mery and Jasmine ( $p < 0.002$ ).
Character*Emotion	$F(3,174) = 53.525, p < 0.0001, \eta^2 = 0.48$	For Mery, happiness more intense than sadness ( $p < 0.0001$ ), but for Jasmine and Malcolm, sadness more intense than happiness ( $p < 0.047$ ).
Emotion*Brightness	$F(1, 58) = 22.972, p < 0.0001, \eta^2 = 0.28$	For happiness, 100% brightness more intense than 25% ( $p < 0.009$ ). For sadness, 25% brightness more intense than 100% ( $p < 0.006$ ).
Emotion*Source	$F(1,58) = 8.12, p < 0.007, \eta^2 = 0.12$	Sadness perceived as more intense by Prolific (EU) participants than MTurk (US) participants ( $p < 0.031$ ).
<b>Appeal</b>		
Effect	F-Test	post-hoc
Character	$F^*(2.28, 132.19), p^* < 0.0001, \eta^2 = 0.17, \epsilon = 0.76$	Franklin less appealing than the rest ( $p < 0.002$ ).
Emotion	$F(1, 58) = 31.56, p < 0.0001, \eta^2 = 0.13$	Happiness more appealing than sadness.
Brightness	$F(1, 58) = 23.85, p < 0.0001, \eta^2 = 0.29$	100% brightness more appealing than 25%.
Emotion*Brightness	$F(1, 58) = 17.01, p < 0.0002, \eta^2 = 0.23$	For happiness, 100% brightness more appealing than 25% ( $p < 0.0002$ ).



**Table A.3:** Chapter 5 - Effect of Audio:  
main effects and interactions with post-hoc analysis.

<b>Recognition</b>		
Effect	F-Test	post-hoc
Audio	$F(1, 28) = 12.14, p < 0.002, \eta^2 = 0.30$	Emotion recognition better in the audio condition.
Emotion	$F(4, 122) = 3.73, p < 0.007, \eta^2 = 0.12$	Fear less recognized in comparison to anger ( $p < 0.03$ ).
Emotion*Audio	$F(4, 122) = 3.94, p < 0.005, \eta^2 = 0.12$	Fear less recognized in no audio condition ( $p < 0.002$ ).
Character*Emotion	$F^*(2.90, 81.07) = 2.75, p^* < 0.05, \eta^2 = 0.09, \epsilon = 0.72$	No differences between characters according to specific emotions.
Character*Emotion*Audio	$F^*(2.90, 81.07) = 3.51, p^* < 0.03, \eta^2 = 0.11, \epsilon = 0.72$	In the audio condition, characters did not have an effect on recognition, whereas in no audio condition Mery sadness and Franklin fear were less recognized ( $p < 0.04$ ).
Emotion*Brightness *KTFR*Audio	$F^*(8.21, 229.76) = 2.37, p^* < 0.018, \eta^2 = 0.08, \epsilon = 0.51$	No significant interactions with KTFR or brightness.
Character*Emotion *Brightness*KTFR	$F^*(8.16, 228.41) = 2.05, p^* < 0.04, \eta^2 = 0.07, \epsilon = 0.51$	No significant interactions with KTFR or brightness.
<b>Intensity</b>		
Effect	F-Test	post-hoc
Character	$F(1, 28) = 5.34, p < 0.03, \eta^2 = 0.20$	Franklin more intense overall.
Emotion	$F(3, 84) = 39.14, p < 0.0001, \eta^2 = 0.58$	Sadness and fear are similar in intensity, while anger is the most intense emotion, followed by happiness ( $p < 0.02$ ).
Emotion*Audio	$F(3, 58) = 5.86, p < 0.0012, \eta^2 = 0.17$	No significant differences in intensity between corresponding emotions according to audio/ no audio condition.
Character*Emotion	$F^*(2.23, 62.52) = 15.36, p^* < 0.0001, \eta^2 = 0.38, \epsilon = 0.76$	Only Mery sadness less intense than Franklin ( $p < 0.0002$ ).
Emotion*Brightness	$F^*(3.09, 86.40) = 6.07, p^* < 0.0008, \eta^2 = 0.18, \epsilon = 0.51$	Only happiness least intense at 25% brightness compared to both other brighter levels ( $p < 0.008$ ).
<b>Appeal</b>		
Effect	F-Test	post-hoc
Character	$F(1, 28) = 8.52, p < 0.007, \eta^2 = 0.23$	Mery more appealing than Franklin.
Emotion	$F^*(2.22, 62.07) = 6.61, p^* < 0.0018, \eta^2 = 0.19, \epsilon = 0.55$	Anger and fear more appealing than neutral expression ( $p < 0.003$ ).
Brightness	$F^*(1.49, 41.79) = 14.07, p^* < 0.0001, \eta^2 = 0.33, \epsilon = 0.75$	Brightness at 25% least appealing ( $p < 0.0003$ ).
Emotion*Audio	$F^*(2.22, 62.07) = 25.2, p^* < 0.033, \eta^2 = 0.11, \epsilon = 0.55$	Neutral was significantly less appealing than other emotions only in audio condition ( $p < 0.005$ ).
Character*Brightness	$F^*(2, 56) = 9.09, p^* < 0.0004, \eta^2 = 0.25, \epsilon = 0.96$	Mery more appealing than Franklin everywhere except at 25% brightness ( $p < 0.05$ ).
Emotion*Brightness	$F^*(4.92, 137.74) = 6.35, p^* < 0.0001, \eta^2 = 0.18, \epsilon = 0.61$	All expressions significantly less appealing at 25% brightness than both 100% and 50% ( $p < 0.0001$ ). Happiness 100% also more appealing than 50% brightness ( $p < 0.0007$ ).
Emotion*KTFR	$F^*(5.33, 149.33) = 2.63, p^* < 0.0234, \eta^2 = 0.09, \epsilon = 0.67$	No significant differences between emotions at different KTFR.
Brightness*KTFR	$F^*(3.00, 84.06) = 5.81, p^* < 0.0012, \eta^2 = 0.17, \epsilon = 0.75$	All brightness levels affect appeal in 4:1 ( $p < 0.02$ ), whereas only 25% lowers appeal in 2:1 and 1:1.
Character*Brightness*KTFR	$F(4, 112) = 3.65, p < 0.008, \eta^2 = 0.12$	Mery more appealing than Franklin at 100% brightness and 50% brightness and KTFR 1:1 ( $p < 0.005$ ).

**Table A.4:** Chapter 5 - Effect of Background:  
main effects and interactions with post-hoc analysis.

<b>Recognition</b>		
Effect	F-Test	post-hoc
Emotion	$F(2.29, 64.21) = 10.3, p^* < 0.0001, \eta^2 = 0.31, \epsilon = 0.57$	Fear less recognized than other emotions except sadness ( $p < 0.023$ ).
Character*Emotion	$F(2.74, 76.77) = 10.30, p^* < 0.0001, \eta^2 = 0.27, \epsilon = 0.69$	Franklin's fear less recognized than other Franklin's emotions and Mery's emotions ( $p < 0.042$ ).
Emotion*Brightness	$F(5.72, 160.11) = 2.62, p^* < 0.021, \eta^2 = 0.09, \epsilon = 0.71$	No interesting interactions.
Emotion*KTFR	$F(5.09, 142.57) = 2.41, p^* < 0.039, \eta^2 = 0.08, \epsilon = 0.64$	No interesting interactions.
<b>Intensity</b>		
Effect	F-Test	post-hoc
Character	$F(1, 28) = 4.97, p < 0.035, \eta^2 = 0.15$	Franklin more intense than Mery.
Emotion	$F(2.02, 56.48) = 56.08, p^* < 0.0001, \eta^2 = 0.67, \epsilon = 0.67$	Anger the most intense, followed by happiness ( $p < 0.0002$ ).
Character*Emotion	$F(3,84) = 13.69, p < 0.0001, \eta^2 = 0.33$	For only sadness, Franklin more intense than Mery ( $p < 0.0002$ ).
Emotion*Brightness	$F(3.99, 111.78) = 7.43, p^* < 0.0001, \eta^2 = 0.21, \epsilon = 0.67$	For only happiness, 100% brightness more intense than 50% and 25% ( $p < 0.005$ ).
Brightness*KTFR	$F(4, 112) = 2.65, p < 0.04, \eta^2 = 0.09$	No significant interactions.
<b>Appeal</b>		
Effect	F-Test	post-hoc
Character	$F(1,28) = 21.89, p < 0.0001, \eta^2 = 0.43$	Mery more appealing than Franklin.
Emotion	$F(2.57, 71.92) = 3.38, p^* < 0.029, \eta^2 = 0.11, \epsilon = 0.64$	Anger more appealing than all other emotions except fear ( $p < 0.044$ ).
Brightness	$F(1.62, 45.27) = 17.51, p^* < 0.0001, \eta^2 = 0.38, \epsilon = 0.83$	25% brightness less appealing than 50% and 100% ( $p < 0.0002$ ).
Character*Brightness	$F(2, 56) = 8.72, p < 0.0005, \eta^2 = 0.23$	For Mery, every higher brightness level is more appealing than its previous level ( $p < 0.0019$ ). For Franklin, 25% significant lower than the others ( $p < 0.0002$ ).
Emotion*Brightness	$F(5.24, 146.74) = 4.13, p^* < 0.002, \eta^2 = 0.13, \epsilon = 0.66$	All expressions significantly less appealing at 25% brightness than both 100% and 50% ( $p < 0.002$ ). Happiness 100% also more appealing than 50% brightness ( $p < 0.009$ ).

**Table A.5:** Chapter 5 - Effect of Movement: main effects and interactions with post-hoc analysis.

<b>Recognition</b>		
Effect	F-Test	post-hoc
Emotion	$F^*(3.02, 84.57) = 9.15, p^* < 0.0001, \eta^2 = 0.25, \epsilon = 0.76$	Anger and happiness higher than sadness, neutral and fear ( $p < 0.008$ ).
Emotion*Brightness	$F^*(4.70, 131.68) = 2.48, p^* < 0.039, \eta^2 = 0.08, \epsilon = 0.59$	No interesting interactions.
Character*Emotion	$F^*(2.59, 72.53) = 5.09, p^* < 0.005, \epsilon = 0.65, \eta^2 = 0.15$	Mery's neutral lower than anger or happiness ( $p < 0.0006$ ).
Character*Brightness	$F(2, 56) = 3.855, p < 0.03, \eta^2 = 0.12$	Franklin higher than Mery at 100% brightness ( $p < 0.04$ ).
Emotion*KTFR*movement	$F(8, 224) = 2.287, p < 0.03, \eta^2 = 0.08$	No interesting interactions.
<b>Intensity</b>		
Effect	F-Test	post-hoc
Emotion	$F^*(2.38, 66.58) = 38.127, p^* < 0.0001, \eta^2 = 0.58, \epsilon = 0.79$	All emotions are significantly different in intensity: anger most intense, followed by happiness, sadness and fear ( $p < 0.03$ ).
Brightness	$F^*(1.32, 36.81) = 3.77, p^* < 0.05, \eta^2 = 0.12, \epsilon = 0.66$	Expression at 25% brightness less intense than at 100% ( $p < 0.003$ ).
Emotion*Movement	$F^*(2.38, 66.58) = 7.75, p^* < 0.0005, \eta^2 = 0.22, \epsilon = 0.79$	Anger and happiness are more intense than other emotions in movie ( $p < 0.02$ ), fear is least intense than other emotions in still images ( $p < 0.009$ ).
KTFR*Movement	$F(2, 56) = 4.34, p < 0.02, \eta^2 = 0.13$	4:1 KTFR increases intensity only for still images ( $p < 0.02$ ).
Emotion*Brightness	$F^*(2.72, 76.26) = 10.17, p^* < 0.0001, \eta^2 = 0.27, \epsilon = 0.45$	Intensity for happiness significantly increases with every brightness level ( $p < 0.002$ ).
Character*Emotion	$F(3, 84) = 8.81, p < 0.0004, \eta^2 = 0.24$	Mery sadness less intense than Franklin sadness ( $p < 0.002$ ).
Character*Emotion*Movement	$F(3, 84) = 4.10, p < 0.01, \eta^2 = 0.13$	Mery sadness less intense than Franklin sadness only for movie, not for still image ( $p < 0.0003$ ).
Emotion*KTFR	$F(6, 168) = 4.63, p < 0.0003, \eta^2 = 0.14,$	Sadness at 4:1 ratio more intense than for other ratios ( $p < 0.04$ ).
Emotion*KTFR*Movement	$F(6, 168) = 2.20, p < 0.05, \eta^2 = 0.07$	No interesting interactions.
Character*Brightness	$F^*(1.61, 45.01) = 4.32, p^* < 0.03, \eta^2 = 0.13, \epsilon = 0.81$	Mery affected by brightness level extremes, Franklin not ( $p < 0.0002$ ).
<b>Appeal</b>		
Effect	F-Test	post-hoc
Character	$F(1, 28) = 28.87, p < 0.00002, \eta^2 = 0.51$	Mery more appealing than Franklin.
Brightness	$F^*(1.60, 44.92) = 32.76, p^* < 0.0001, \eta^2 = 0.54, \epsilon = 0.80$	25% brightness least appealing ( $p < 0.0001$ ).
KTFR	$F^*(1.65, 46.12) = 5.00, p^* < 0.016, \eta^2 = 0.15, \epsilon = 0.82$	KTFR 4:1 less appealing than both 2:1 and 1:1 levels ( $p < 0.03$ ).
Character*Brightness	$F^*(1.48, 41.64) = 16.14, p^* < 0.0001, \eta^2 = 0.37, \epsilon = 0.74$	Mery more appealing than Franklin at 100% and 50% brightness ( $p < 0.002$ ).
Emotion*Brightness	$F(8, 224) = 6.26, p < 0.0001, \eta^2 = 0.18$	All expressions significantly less appealing at 25% brightness than both 100% and 50% ( $p < 0.0001$ ). Happiness 100% also more appealing than 50% brightness ( $p < 0.0007$ ).
Brightness*KTFR	$F^*(2.51, 70.41) = 10.4, p^* < 0.0001, \eta^2 = 0.27, \epsilon = 0.63$	At 100% brightness, KTFR 2:1 and 4:1 more appealing than 1:1 ( $p < 0.02$ ), at 50% and 25% (4:1 more appealing than 1:1 ( $p < 0.03$ )).
Emotion*KTFR	$F(8, 224) = 3.75, p < 0.0004, \eta^2 = 0.12$	For happiness, 4:1 ratio least appealing ( $p < 0.03$ in all cases).
Character*KTFR*Movement	$F(2, 56) = 3.92, p < 0.03, \eta^2 = 0.12$	Mery more appealing than Franklin for all KTFR ratios for movies ( $p < 0.02$ ).
Character*Emotion*Brightness*KTFR	$F(16, 448) = 1.75, p < 0.04, \eta^2 = 0.06$	Increasing KTFR and brightness levels affect happiness for Mery, not Franklin ( $p < 0.009$ ).

**Table A.6:** Chapter 5 - Realism: Brightness:  
main effects and interactions with post-hoc analysis.

<b>Recognition</b>		
Effect	F-Test	post-hoc
Gender	$F(1, 14) = 7.26, p < 0.0174, \eta^2 = 0.34$	Male character more recognized than female character.
Emotion	$F(3, 42) = 14.84, p < 0.0000, \eta^2 = 0.51$	Neutral less recognized than others emotions ( $p < 0.0006$ ).
Shape	$F(2, 28) = 7.48, p < 0.0025, \eta^2 = 0.35$	Middle more recognized than other shapes ( $p < 0.01$ ).
Gender*Shape	$F(2, 28) = 15.26, p < 0.0003, \eta^2 = 0.52$	Female realistic and Male toon least recognized ( $p < 0.005$ ).
Emotion*Shape	$F^*(2.88, 40.29) = 5.10, p^* < 0.005, \eta^2 = 0.27, \epsilon = 0.48$	Neutral toon and neutral realistic least recognized ( $p < 0.021$ ).
Gender*Emotion*Shape	$F^*(3.29, 46.06) = 2.83, p^* < 0.044, \eta^2 = 0.17, \epsilon = 0.55$	For male, neutral toon least recognized ( $p < 0.0004$ ). For female, neutral and sadness realistic, and neutral toon least recognized ( $p < 0.0004$ ).
<b>Intensity</b>		
Effect	F-Test	post-hoc
Gender	$F(1, 14) = 22.48, p < 0.0004, \eta^2 = 0.60$	Male characters more intense than female characters.
Emotion	$F(2, 28) = 6.94, p < 0.0036, \eta^2 = 0.60$	Anger more intense than sadness and happiness ( $p < 0.025$ ).
Shape	$F(2, 28) = 10.06, p < 0.0006, \eta^2 = 0.40$	Realistic shape less intense than the other shapes ( $p < 0.002$ ).
Gender*Emotion	$F(2, 28) = 5.20, p < 0.013, \eta^2 = 0.35$	Male sadness and happiness more intense than the corresponding female emotions ( $p < 0.04$ ).
Gender*Shape	$F^*(1.29, 18.03) = 41.76, p^* < 0.001, \eta^2 = 0.66, \epsilon = 0.64$	Female realistic less intense than the rest ( $p < 0.0002$ ).
Emotion*Shape	$F(4, 56) = 7.05, p < 0.0002, \eta^2 = 0.23$	Realistic sad and realistic happy characters less intense than the rest ( $p < 0.029$ ).
Gender*Emotion*Shape	$F(4, 56) = 6.13, p < 0.0004, \eta^2 = 0.35$	Only for female, realistic sad and realistic happy emotions less intense than the rest ( $p < 0.0057$ ).
<b>Appeal</b>		
Effect	F-Test	post-hoc
Gender	$F(1, 14) = 5.78, p < 0.0306, \eta^2 = 0.29$	Male characters more appealing than female characters.
Emotion	$F^*(2.16, 30.24) = 4.73, p^* < 0.014, \eta^2 = 0.25, \epsilon = 0.72$	Anger and happiness less appealing than neutral ( $p < 0.041$ ).
Shape	$F^*(1.35, 18.87) = 15.21, p^* < 0.0005, \eta^2 = 0.52, \epsilon = 0.67$	Realistic characters less appealing than the rest ( $p < 0.0005$ ).
Brightness	$F^*(1.25, 17.43) = 5.20, p^* < 0.029, \eta^2 = 0.27, \epsilon = 0.62$	25% Brightness less appealing than 100% brightness ( $p < 0.009$ ).
Gender*Emotion	$F^*(2.02, 28.25) = 3.77, p^* < 0.035, \eta^2 = 0.21, \epsilon = 0.67$	For anger, male characters significantly more appealing than female ( $p < 0.0007$ ).
Emotion*Shape	$F(6, 84) = 5.99, p < 0.00003, \eta^2 = 0.30$	For each emotion, realistic shape is less appealing than other shapes ( $p < 0.004$ ).
Gender*Emotion*Shape	$F(6, 84) = 2.67, p < 0.02033, \eta^2 = 0.16$	For realistic happiness, the appeal of the male shape drops from the other shapes more significantly than the female shape ( $p < 0.0005$ ).
<b>Eeriness</b>		
Effect	F-Test	post-hoc
Emotion	$F^*(1.87, 26.13) = 8.92, p^* < 0.002, \eta^2 = 0.39, \epsilon = 0.62$	Anger and happiness more eerie than neutral and sadness ( $p < 0.026$ ).
Shape	$F^*(1.44, 20.12) = 21.80, p^* < 0.001, \eta^2 = 0.61, \epsilon = 0.72$	Realistic shapes more eerie than other shapes ( $p < 0.00015$ ).
Gender*Shape	$F(2, 28) = 5.96, p < 0.0070, \eta^2 = 0.30$	Female middle shape more eerie than male middle shape ( $p < 0.0014$ ).
Emotion*Shape	$F(6, 84) = 3.90, p < 0.0018, \eta^2 = 0.22$	For each individual emotion, realistic shapes more eerie than the other shapes in the same emotion ( $p < 0.031$ ).
Gender*Emotion*Shape	$F(6, 84) = 5.75, p < 0.0001, \eta^2 = 0.29$	For each individual gender/emotion pair, realistic shapes are more eerie than the other shapes in the same combination of gender/emotion except for female happiness, where middle shape becomes as eerie as the realistic one ( $p < 0.002$ ).
Gender*Shape*Brightness	$F(4, 56) = 3.39, p < 0.015, \eta^2 = 0.19$	Male toon at 25% more eerie than male toon at 100% ( $p < 0.004$ ).

**Table A.7:** Chapter 5 - Realism: KTFR: main effects and interactions with post-hoc analysis.

<b>Recognition</b>		
Effect	F-Test	post-hoc
Gender	$F(1, 14) = 13.66, p < 0.0024, \eta^2 = 0.49$	Emotions of male characters more recognized than female characters.
Emotion	$F^*(1.36, 19.02) = 9.36, p^* < 0.006, \eta^2 = 0.40, \epsilon = 0.45$	Neutral less recognized than other emotions ( $p < 0.0056$ ).
Shape	$F^*(1.46, 20.41) = 5.83, p^* < 0.017, \eta^2 = 0.29, \epsilon = 0.73$	Realistic shapes less recognized than middle shapes ( $p < 0.0067$ ).
KTFR	$F(2, 28) = 4.22, p < 0.025, \eta^2 = 0.23$	16:1 KTFR less recognized than 4:1 ( $p < 0.022$ ).
Gender*Shape	$F^*(1.39, 19.42) = 10.46, p^* < 0.003, \eta^2 = 0.43, \epsilon = 0.69$	Female realistic shapes less recognized than the rest ( $p < 0.049$ ), male toon less recognized than male middle ( $p < 0.038$ ).
Gender*Emotion*Shape	$F^*(2.8, 39.26) = 4.71, p^* < 0.008, \eta^2 = 0.25, \epsilon = 0.47$	For female, realistic sadness and neutral less recognized ( $p < 0.0002$ ), for male, toon and realistic neutral less recognized compared to middle neutral ( $p < 0.0003$ ).
Gender*Emotion*KTFR	$F(6, 84) = 2.83, p < 0.015, \eta^2 = 0.17$	For KTFR 1:1, sadness is less recognized for female than male character ( $p < 0.0003$ ).
<b>Intensity</b>		
Effect	F-Test	post-hoc
Gender	$F(1, 14) = 65.99, p < 0.0000, \eta^2 = 0.81$	Male characters more intense than female characters.
Emotion	$F(2, 28) = 12.41, p < 0.0002, \eta^2 = 0.62$	Anger more intense than sadness and happiness ( $p < 0.0027$ ).
Gender*Shape	$F^*(1.22, 17.09) = 69.51, p^* < 0.0001, \eta^2 = 0.82, \epsilon = 0.61$	Male realistic more intense than male middle and toon ( $p < 0.0002$ ). Female realistic less intense than all other combinations ( $p < 0.0002$ ).
Emotion*Shape	$F(4, 56) = 10.54, p < 0.0001, \eta^2 = 0.25$	Toon anger less intense than middle and realistic anger ( $p < 0.02$ ), toon sadness more intense than realistic sadness ( $p < 0.004$ ).
Emotion*KTFR	$F(4, 56) = 8.06, p < 0.0001, \eta^2 = 0.25$	In KTFR 1:1, happiness more intense than in KTFR 16:1 ( $p < 0.0012$ ).
Gender*Emotion*Shape	$F(4, 56) = 8.51, p < 0.0001, \eta^2 = 0.50$	Female realistic sadness and happiness less intense than all combinations ( $p < 0.0002$ ), male realistic anger more intense than all other combinations ( $p < 0.0002$ ).
<b>Appeal</b>		
Effect	F-Test	post-hoc
Gender	$F(1, 14) = 16.35, p < 0.0013, \eta^2 = 0.93$	Male characters more appealing than female characters.
Emotion	$F(3, 42) = 6.83, p < 0.0008, \eta^2 = 0.33$	Neutral and sadness more appealing than anger and happiness ( $p < 0.027$ ).
Shape	$F(2, 28) = 12.13, p < 0.0002, \eta^2 = 0.46$	Realistic shape less appealing than toon and middle shapes ( $p < 0.009$ ).
KTFR	$F(2, 28) = 9.92, p < 0.0006, \eta^2 = 0.41$	16:1 KTFR less appealing than 4:1 and 1:1 ( $p < 0.049$ ).
Gender*Emotion	$F(3, 42) = 4.39, p < 0.009, \eta^2 = 0.24$	Female emotions less appealing than male except for sadness ( $p < 0.007$ ).
Emotion*Shape	$F^*(3.72, 52.02) = 4.04, p^* < 0.008, \eta^2 = 0.22, \epsilon = 0.62$	for neutral, sadness and happiness, realistic shapes are the least appealing ( $p < 0.003$ ).
Emotion*KTFR	$F(6, 84) = 5.37, p < 0.001, \eta^2 = 0.28$	For anger, 1:1 KTFR more appealing than 16:1 ( $p < 0.0002$ ). For happiness, 1:1 KTFR more appealing than 4:1 and 16:1 KTFRs ( $p < 0.0006$ ).
Shape*KTFR	$F(4, 56) = 10.24, p < 0.0001, \eta^2 = 0.42$	For toon shape, 1:1 KTFR more appealing than 4:1 and 16:1 KTFRs ( $p < 0.011$ ). For middle shape, 16:1 KTFR less appealing than 4:1 and 1:1 KTFRs ( $p < 0.0064$ ).
Gender*Emotion*Shape	$F^*(3.21, 44.88) = 5.52, p^* < 0.003, \eta^2 = 0.28, \epsilon = 0.53$	Realistic shape least appealing for most emotions ( $p < 0.04$ ) except toon anger for both genders, and happiness for females, for which all shapes are equally less appealing.
Gender*Emotion*KTFR	$F(6, 84) = 3.13, p < 0.0082, \eta^2 = 0.18$	For male happiness, 1:1 KTFR more appealing than 4:1 and 16:1 KTFRs ( $p < 0.002$ ). For female anger, 1:1 KTFR more appealing than 16:1 KTFR ( $p < 0.0017$ ).

**Table A.8:** Chapter 5 - Realism: KTFR (continued):  
main effects and interactions with post-hoc analysis.

<b>Eeriness</b>		
Effect	F-Test	post-hoc
Gender	$F(1, 14) = 9.22, p < 0.0089, \eta^2 = 0.40$	Female characters more eerie than male characters.
Emotion	$F(3, 42) = 10.94, p < 0.0001, \eta^2 = 0.44$	Anger and happiness more eerie than neutral and sadness ( $p < 0.0016$ ).
Shape	$F(2, 28) = 11.99, p < 0.0002, \eta^2 = 0.46$	Realistic shape more eerie than toon and middle shapes ( $p < 0.0011$ ).
KTFR	$F^*(1.29, 18.03) = 10.28, p^* < 0.003, \eta^2 = 0.42, \epsilon = 0.64$	16:1 KTFR more eerie than 1:1 and 4:1 KTFRs ( $p < 0.0276$ ).
Gender*Emotion	$F(3, 42) = 4.38, p < 0.009, \eta^2 = 0.24$	Female anger and neutral more eerie than corresponding male emotions ( $p < 0.004$ ).
Gender*Emotion*Shape	$F(6, 84) = 2.47, p < 0.0302, \eta^2 = 0.15$	Male realistic happiness more eerie than most other combinations ( $p < 0.0123$ ).
Gender*Shape*KTFR	$F(4, 56) = 3.25, p < 0.0183, \eta^2 = 0.19$	For male toon and male middle, 16:1 more eerie than 1:1 but not 4:1 ( $p < 0.0089$ ). For male realistic, 16:1 more eerie than 4:1 but not 1:1 ( $p < 0.0073$ ).
Gender*Emotion*Shape*KTFR	$F(12, 168) = 1.91, p < 0.0367, \eta^2 = 0.12$	No interesting interactions.

# Bibliography

- Adams, Ansel (1948). "The Negative: Exposure and Development Basic Photo 2". In: *Morgan and Lester* 98.
- Akenine-Möller, Tomas, Eric Haines, and Naty Hoffman (2019). *Real-time rendering*. Crc Press.
- Alton, John (1995). *Painting with light*. University of California Press.
- Anderson, Barton L (2011). "Visual perception of materials and surfaces". In: *Current biology* 21.24, R978–R983.
- Assen, Jan Jaap R van and Roland W Fleming (2016). "Influence of optical material properties on the perception of liquids". In: *Journal of Vision* 16.15, p. 12.
- Baraff, David and Andrew Witkin (1997). "Physically-based Modeling, Principles and Practice". In: *Course Notes of ACM SIGGRAPH*.
- Barchard, Kimberly A., Kelly E. Grob, and Matthew J. Roe (2017). "Is sadness blue? The problem of using figurative language for emotions on psychological tests". In: *Behavior Research Methods* 49.2, pp. 443–456.
- Barrett, Lisa Feldman (2006). "Are emotions natural kinds?" In: *Perspectives on psychological science* 1.1, pp. 28–58.
- Bartz, D et al. (2008). "The Role of Perception for Computer Graphics". In: *29th Annual Conference Eurographics (EG 2008)*. The Eurographics Association, pp. 65–86.
- Belhumeur, Peter N, David J Kriegman, and Alan L Yuille (1999). "The bas-relief ambiguity". In: *International journal of computer vision* 35.1, pp. 33–44.
- Ben-David, Boaz M, Pascal H H M van Lieshout, and Talia Leszcz (2011). "A resource of validated affective and neutral sentences to assess identification of emotion in spoken language after a brain injury". In: *Brain injury* 25.2, pp. 206–220.
- Berkman, Mehmet Ilker and Dilek Karahoca (2016). "Re-assessing the usability metric for user experience (UMUX) scale". In: *Journal of Usability Studies* 11.3, pp. 89–109.
- Bi, Wenyan and Bei Xiao (2016). "Perceptual Constancy of Mechanical Properties of Cloth Under Variation of External Forces". In: *Proceedings of the ACM Symposium on Applied Perception*. SAP '16. New York, NY, USA: ACM, pp. 19–23.
- Birn, Jeremy (2000). *Digital Lighting and Rendering*. New Riders Publishing.
- Borg, Ingwer and James Lingo (2012). *Multidimensional similarity structure analysis*. Springer Science & Business Media.
- Brainard, David H (2003). "Color appearance and color difference specification". In: *The science of color 2*, pp. 191–216.
- Brown, Blain (2016). *Cinematography: theory and practice: image making for cinematographers and directors*. Taylor & Francis.
- Bruckstein, Alfred M (1988). "On shape from shading". In: *Computer Vision, Graphics, and Image Processing* 44.2, pp. 139–154.
- Calahan, Sharon (2000). *Storytelling through lighting: a computer graphics perspective*. Ed. by Anthony A Apodaca, Larry Gritz, and Ronen Barzel. Advanced RenderMan: Creating CGI for motion pictures. Morgan Kaufmann. Chap. 13, pp. 337–382.
- Catmull, Edwin (1975). "Computer Display of Curved Surfaces". In: *Proc. IEEE Conf. on Computer Graphics, Pattern Recognition, and Data Structure*. Vol. 11.

- Catmull, Edwin and James Clark (1978). "Recursively generated B-spline surfaces on arbitrary topological meshes". In: *Computer-aided design* 10.6, pp. 350–355.
- Chin, Stephanie F, T Joel Wade, and Cassandra French (2006). "Race and facial attractiveness: Individual differences in perceived adoptability of children". In: *Journal of Cultural and Evolutionary Psychology* 4.3-4, pp. 215–229.
- Cook, Robert L, Loren Carpenter, and Edwin Catmull (1987). "The Reyes image rendering architecture". In: *ACM SIGGRAPH Computer Graphics* 21.4, pp. 95–102.
- Cox, Trevor F and Michael A A Cox (2001). *Multidimensional scaling, 2nd Ed.* Boca Raton, FL: Chapman & Hall/CRC, 2001.
- Cunningham, Douglas W and Christian Wallraven (2009a). "Dynamic information for the recognition of conversational expressions". In: *Journal of Vision* 9.13, p. 7.
- (2009b). "The interaction between motion and form in expression recognition". In: *Proceedings of the 6th symposium on applied perception in graphics and visualization*. ACM, pp. 41–44.
- (2011). *Experimental design: From user studies to psychophysics*. AK Peters/CRC Press.
- Cunningham, Douglas W et al. (2007). "Perceptual reparameterization of material properties". In: *Proceedings of the Third Eurographics conference on Computational Aesthetics in Graphics, Visualization and Imaging*. Eurographics Association, pp. 89–96.
- Cunningham, Michael R et al. (1995). "'Their ideas of beauty are, on the whole, the same as ours': Consistency and variability in the cross-cultural perception of female physical attractiveness." In: *Journal of personality and social psychology* 68.2, p. 261.
- Dictionary, Oxford English (1989). "Oxford english dictionary". In: *Simpson, JA & Weiner, ESC.*
- Doerschner, Katja et al. (2011). "Visual motion and the perception of surface material". In: *Current Biology* 21.23, pp. 2010–2016.
- Dryden, Ian L and Kanti V Mardia (1998). *Statistical shape analysis*. Vol. 4. Wiley Chichester.
- Dutre, P, P. Bekaert, and K. Bala (2018). *Advanced Global Illumination*. CRC Press. ISBN: 9781315359878.
- Ebner, Marc (2007). *Color constancy*. Vol. 7. John Wiley & Sons.
- Ekman, Paul (1992). "An argument for basic emotions". In: *Cognition & Emotion* 6.3-4, pp. 169–200.
- Fechner, Gustav Theodor, Davis H Howes, and Edwin Garrigues Boring (1966). *Elements of psychophysics*.
- Fetterman, Adam K, Michael D Robinson, and Brian P Meier (2012). "Anger as "seeing red": Evidence for a perceptual association". In: *Cognition & emotion* 26.8, pp. 1445–1458.
- Filip, Jiří et al. (2018). "Evaluating physical and rendered material appearance". In: *The Visual Computer* 34.6-8, pp. 805–816.
- Finstad, Kraig (2010). "The usability metric for user experience". In: *Interacting with Computers* 22.5, pp. 323–327.
- Fleming, Roland W (2014). "Visual perception of materials and their properties". In: *Vision Research* 94, pp. 62–75.
- Foster, David H (2011). "Color constancy". In: *Vision Research* 51.7, pp. 674–700.
- Frijda, Nico H (1988). "The laws of emotion." In: *American Psychologist* 43.5, p. 349.
- Gescheider, George A (2013). *Psychophysics: the fundamentals*. Psychology Press.
- Gibson, James J (2014). *The ecological approach to visual perception: classic edition*. Psychology Press.
- Gilchrist, Alan et al. (1999). "An anchoring theory of lightness perception." In: *Psychological review* 106.4, p. 795.
- Gilchrist, Alan L (2013). *Lightness, brightness and transparency*. Psychology Press.
- Goldstein, E.B. (2009). *Sensation and Perception*. Cengage Learning. ISBN: 9781111783068. URL: <https://books.google.ie/books?id=EbgFAAAQBAJ>.



- Grodal, Torben (2005). "9 Film Lighting and Mood". In: *Moving Image Theory: Ecological Considerations*, p. 152.
- Gurney, James (2010). *Color and light: A guide for the realist painter*. Andrews McMeel Publishing.
- Hall, M B (1992). *Color and Meaning: Practice and Theory in Renaissance Painting*. Cambridge University Press.
- Hery, Christophe and Ryusuke Villemin (2013). *Physically Based Lighting at Pixar*.
- Hill, H. and Vicki Bruce (1996). "Effects of lighting on the perception of facial surfaces." In: *Journal of experimental psychology. Human perception and performance* 22.4, p. 986.
- Hodgins, Jessica et al. (2010). "The saliency of anomalies in animated human characters". In: *ACM Transactions on Applied Perception (TAP)* 7.4.
- Hupka, Ralph B et al. (1997). "The colors of anger, envy, fear, and jealousy: A cross-cultural study". In: *Journal of cross-cultural psychology* 28.2, pp. 156–171.
- Hyde, Jennifer et al. (2013). "Perceptual effects of damped and exaggerated facial motion in animated characters". In: *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on. IEEE*, pp. 1–6.
- Jensen, Henrik Wann (2001). "State of the Art in Monte Carlo Ray Tracing for Realistic Image Synthesis". In:
- Jensen, Henrik Wann et al. (2001). "A practical model for subsurface light transport". In: *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. ACM, pp. 511–518.
- Johnston, Alan, Harold Hill, and Nicole Carman (1992). "Recognising Faces: Effects of Lighting Direction, Inversion, and Brightness Reversal". In: *Perception* 42.11, pp. 1227–1237.
- Kajiya, James T (1986). "The Rendering Equation". In: *SIGGRAPH Comput. Graph.* 20.4, pp. 143–150.
- Kalat, James W. (2007). *Introduction to Psychology*. Thomson-Wadsworth.
- Kardos, Ludwig (1934). "Ding und Schatten. Eine experimentelle Untersuchung über die Grundlagen des Farbensehens." In: *Zeitschrift für Psychologie und Physiologie der Sinnesorgane. Abt. 1. Zeitschrift für Psychologie*.
- Kätsyri, Jari et al. (2015). "A review of empirical evidence on different Uncanny Valley hypotheses: Support for perceptual mismatch as one road to the valley of eeriness". In: *Frontiers in Psychology* 6, p. 390.
- Kawai, Nobuyuki et al. (2013). "Shadows alter facial expressions of Noh masks". In: *PloS one* 8.8, e71389.
- Khan, Erum Arif et al. (2006). *Image-based material editing*. Vol. 25. 3. ACM.
- Koenderink, Jan J et al. (1996). "Shape constancy in pictorial relief". In: *Perception* 25.2, pp. 155–164.
- Kokkinara, Elena and Rachel McDonnell (2015). "Animation realism affects perceived character appeal of a self-virtual face". In: *Proceedings of the 8th ACM SIGGRAPH Conference on Motion in Games*. ACM, pp. 221–226.
- Koyama, Yuki et al. (2017). "Sequential line search for efficient visual design optimization by crowds". In: *ACM Transactions on Graphics (TOG)* 36.4, pp. 1–11.
- Lagunas, Manuel et al. (2019). "A Similarity Measure for Material Appearance". In: *ACM Transactions on Graphics (SIGGRAPH 2019)* 38.4.
- Lake, Adam et al. (2000). "Stylized Rendering Techniques for Scalable Real-time 3D Animation". In: *Proceedings of the 1st International Symposium on Non-photorealistic Animation and Rendering*. NPAR '00. New York, NY, USA: ACM, pp. 13–20.
- Landau, David (2014). *Lighting for Cinematography: A Practical Guide to the Art and Craft of Lighting for the Moving Image*. The CineTech Guides to the Film Crafts. Bloomsbury Academic.

- Landau, Mark J, Brian P Meier, and Lucas A Keefer (2010). "A metaphor-enriched social cognition." In: *Psychological bulletin* 136.6, p. 1045.
- Lasseter, John (1987). "Principles of traditional animation applied to 3D computer animation". In: *ACM Siggraph Computer Graphics*. Vol. 21. 4. ACM, pp. 35–44.
- Logvinenko, Alexander D and Laurence T Maloney (2006). "The proximity structure of achromatic surface colors and the impossibility of asymmetric lightness matching". In: *Perception & Psychophysics* 68.1, pp. 76–83.
- Lotman, Elen (2016). "Exploring the Ways Cinematography Affects Viewers' Perceived Empathy towards Onscreen Characters". In: *Baltic Screen Media Review* 4.1.
- Lowell, Ross (1992). *Matters of light & depth: creating memorable images for video, film & stills through lighting*. Broad Street Books.
- Luo, Li Zhu, Hong Li, and Kang Lee (2011). "Are children's faces really more appealing than those of adults? Testing the baby schema hypothesis beyond infancy". In: *Journal of experimental child psychology* 110.1, pp. 115–124.
- MacDorman, Karl F et al. (2009). "Too real for comfort? uncanny responses to computer generated faces." In: *Computers in human behavior* 25.3, pp. 695–710.
- Magnenat-Thalmann, Nadia, Richard Laperrire, and Daniel Thalmann (1988). "Joint-dependent local deformations for hand animation and object grasping". In: *In Proceedings on Graphics interface'88*. Citeseer.
- Malkiewicz, Kris and M David Mullen (2005). *Cinematography*. Simon and Schuster.
- Maloney, Laurence T and David H Brainard (2010). "Color and material perception: Achievements and challenges". In: *Journal of vision* 10.9, pp. 19–19.
- Marschner, S. and P. Shirley (2016). *Fundamentals of Computer Graphics*. CRC Press. ISBN: 9781498785907.
- Martin, O et al. (2006). "The eNTERFACE'05 Audio-Visual Emotion Database". In: *Proceedings of the 22Nd International Conference on Data Engineering Workshops*. ICDEW '06. Washington, DC, USA: IEEE Computer Society, pp. 8–.
- Masson, T. (2007). *CG 101: A Computer Graphics Industry Reference*. Digital Fauxtography. ISBN: 9780977871001.
- McCloud, Scott (1993). *Understanding Comics*. William Morrow Paperbacks.
- (1994). *Understanding Comics: The Invisible Art*. HarperPerennial.
- McDonnell, Rachel, Martin Breidt, and Heinrich Buelthoff (2012). "Render me Real {I}nvestigating the Effect of Render Style on the Perception of Animated Virtual Humans". In: *ACM Transactions on Graphics* 31.4, 91:1–91:11.
- McIver, G. (2017). *Art History for Filmmakers: The Art of Visual Storytelling*. Bloomsbury Publishing. ISBN: 9781474246200.
- Meier, Brian P and Michael D Robinson (2005). "The metaphorical representation of affect". In: *Metaphor and symbol* 20.4, pp. 239–257.
- Meier, Brian P, Michael D Robinson, and Gerald L Clore (2004). "Why good guys wear white: Automatic inferences about stimulus valence based on brightness". In: *Psychological science* 15.2, pp. 82–87.
- Melo, Celso de and Ana Paiva (2007). "Affective Computing and Intelligent Interaction: Second International Conference, ACII 2007 Lisbon, Portugal, September 12-14, 2007 Proceedings". In: ed. by Ana C R Paiva, Rui Prada, and Rosalind W Picard. Berlin, Heidelberg: Springer Berlin Heidelberg. Chap. Expression, pp. 546–557.
- Millerson, Gerald (1991). *The Technique of Lighting for Television and Film*. Focus Press.
- Mingolla, Ennio (1983). "Perception of surface curvature and direction of illumination from patterns of shading". In: *Journal of Experimental Psychology* 9.4, pp. 583–595.
- Moffitt, J F (2004). *Caravaggio in Context: Learned Naturalism and Renaissance Humanism*. McFarland & Company.

- Morgenstern, Yaniv, Wilson S Geisler, and Richard F Murray (2014). "Human vision is attuned to the diffuseness of natural light". In: *Journal of Vision* 14.9, pp. 15–15.
- Morgenstern, Yaniv, Richard F Murray, and Laurence R Harris (2011). "The human visual system's assumption that light comes from above is weak". In: *Proceedings of the National Academy of Sciences* 108.30, pp. 12551–12553.
- Mori, Masahiro (1970). "Bukimi no tani [the uncanny valley]". In: *Energy* 7, pp. 33–35.
- Mori, Masahiro, Karl F MacDorman, and Norri Kageki (2012). "The uncanny valley [from the field]". In: *IEEE Robotics & Automation Magazine* 19.2, pp. 98–100.
- Nicodemus, Fred E et al. (1992). "Geometrical considerations and nomenclature for reflectance". In: *NBS monograph* 160, p. 4.
- O'Hailey, Tina (2018). *Rig it Right! Maya Animation Rigging Concepts, 2nd edition*. CRC Press. ISBN: 9780429866753.
- Olson, Ingrid R and Christy Marshuetz (2005). "Facial attractiveness is appraised in a glance." In: *Emotion* 5.4, p. 498.
- Orvalho, Verónica et al. (2012). "A Facial Rigging Survey." In: *Eurographics (STARs)*, pp. 183–204.
- Pamir, Zahide and Huseyin Boyaci (2016). "Context-dependent lightness affects perceived contrast". In: *Vision research* 124, pp. 24–33.
- Paolacci, Gabriele, Jesse Chandler, and Panagiotis G Ipeirotis (2010). "Running experiments on amazon mechanical turk". In: *Judgment and Decision making* 5.5, pp. 411–419.
- Papesh, Megan H and Stephen D Goldinger (2010). "A multidimensional scaling analysis of own-and cross-race face spaces". In: *Cognition* 116.2, pp. 283–288.
- Parke, Frederick I (1972). "Computer generated animation of faces". In: *Proceedings of the ACM annual conference-Volume 1*. ACM, pp. 451–457.
- Peli, Eli et al. (1991). "Effect of luminance on suprathreshold contrast perception". In: *JOSA A* 8.8, pp. 1352–1359.
- Pelli, Denis G and Bart Farell (1995). "Psychophysical methods". In: *Handbook of optics* 1, pp. 29–1.
- Plutchik, Robert (1980). "A general psychoevolutionary theory of emotion". In: *Theories of emotion*. Elsevier, pp. 3–33.
- (2001). "The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice". In: *American scientist* 89.4, pp. 344–350.
- Poland, Jennifer Lee (2015). "Light, Camera, Emotion! An Examination on Film Lighting and It's Impact on Audiences' Emotional Response". PhD thesis. Cleveland State University.
- Pramaggiore, Maria and Tom Wallis (2005). *Film: A Critical Introduction*. Laurence King.
- Ramanarayanan, Ganesh et al. (2008). "Dimensionality of visual complexity in computer graphics scenes". In: *Human Vision and Electronic Imaging XIII*. Vol. 6806. International Society for Optics and Photonics, 68060E.
- Reid, Francis (2013). *Stage lighting handbook*. Routledge.
- Reips, Ulf-Dietrich (2000). "The Web experiment method: Advantages, disadvantages, and solutions". In: *Psychological experiments on the Internet*. Elsevier, pp. 89–117.
- Research and Markets (2011). *Global Animation, VFX Video Games Industry: Strategies, Trends Opportunities, (2020-25)*. <https://www.researchandmarkets.com/reports/4900485/>.
- Rhodes, Gillian et al. (2001). "Attractiveness of facial averageness and symmetry in non-Western cultures: In search of biologically based standards of beauty". In: *Perception* 30.5, pp. 611–625.
- Richardson, Marion Webster (1938). "Multidimensional psychophysics". In: *Psychological Bulletin* 35, pp. 659–660.
- Ritchie, Kiaran, Jake Callery, and Karim. Biri (2005). *The Art of Rigging: A Definitive Guide to Character Technical Direction with Alias Maya*. Vol. 1. Alias.

- Ritschel, Tobias et al. (2012). "The state of the art in interactive global illumination". In: *Computer Graphics Forum*. Vol. 31. 1. Wiley Online Library, pp. 160–188.
- Rogers, D.F. (2001). *An Introduction to NURBS: With Historical Perspective*. Elsevier Science. ISBN: 9781558606692.
- Ruhland, Kerstin, Mukta Prasad, and Rachel McDonnell (2017). "Data-driven approach to synthesizing facial animation using motion capture". In: *IEEE computer graphics and applications* 37.4, pp. 30–41.
- Russell, James A (1994). "Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies." In: *Psychological bulletin* 115.1, p. 102.
- Russell, James A and José Miguel Fernández-Dols (1997). *The psychology of facial expression*. Vol. 10. Cambridge university press Cambridge.
- Schacter, D.L., D. T. Gilbert, and D. M. Wegner (2011). *Psychology (2nd Edition)*. New York: Worth.
- Schiffman, S.S. et al. (1981). *Introduction to Multidimensional Scaling: Theory, Methods and Applications*. Emerald Group Publishing Limited.
- Schulz, Katharina and Gregor U Hayn-Leichsenring (2017). "Face attractiveness versus artistic beauty in art portraits: a behavioral study". In: *Frontiers in psychology* 8, p. 2254.
- Seif El-Nasr, Magy et al. (2006). "Dynamic lighting for tension in games". In: *Game Studies* 7.1.
- Seifi, H, S DiPaola, and J T Enns (2012). "Exploring the Effect of Color Palette in Painterly Rendered Character Sequences". In: *Proceedings of the Eighth Annual Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging*. CAe '12. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, pp. 89–97.
- Sergent, Justine (1984). "An investigation into component and configural processes underlying face perception". In: *British journal of psychology* 75.2, pp. 221–242.
- Seyama, Jun'ichiro and Ruth S Nagayama (2007). "The uncanny valley: Effect of realism on the impression of artificial human faces". In: *Presence: Teleoperators and virtual environments* 16.4, pp. 337–351.
- Song, Hyunjin et al. (2012). "Brighten up: Smiles facilitate perceptual judgment of facial lightness". In: *Journal of Experimental Social Psychology* 48.1, pp. 450–452.
- Stevens, Stanley S (1957). "On the psychophysical law." In: *Psychological review* 64.3, p. 153.
- Stone, Maureen C (2003). "A field guide to digital color". In:
- Thomas, Frank and Ollie Johnston (1995). *The illusion of life: Disney animation*. Hyperion New York.
- Thompson, W. et al. (2016). *Visual Perception from a Computer Graphics Perspective*. CRC Press. ISBN: 9781466502765.
- Tokunaga, Rumi and Alexander D Logvinenko (2010). "Material and lighting dimensions of object colour". In: *Vision Research* 50.17, pp. 1740–1747.
- Tomkins, Silvan (1962). *Affect imagery consciousness: Volume I: The positive affects*. Springer publishing company.
- Torgerson, Warren S (1952). "Multidimensional scaling: I. Theory and method". In: *Psychometrika* 17.4, pp. 401–419.
- Van Nes, Floris L and Maarten A Bouman (1967). "Spatial modulation transfer in the human eye". In: *JOSA* 57.3, pp. 401–406.
- Vangorp, Peter, Jurgen Laurijssen, and Philip Dutré (2007). "The Influence of Shape on the Perception of Material Reflectance". In: *ACM Trans. Graph.* 26.3.
- Vangorp, Peter et al. (2015). "A Model of Local Adaptation". In: *ACM Transactions on Graphics* 34.6, 166:1–166:13.
- Volante, Matias et al. (2016). "Effects of Virtual Human Appearance Fidelity on Emotion Contagion in Affective Inter-Personal Simulations". In: *IEEE Transactions on Visualization & Computer Graphics* 22.4, pp. 1326–1335.

- Wallraven, Christian et al. (2007). "Evaluation of real-world and computer-generated stylized facial expressions". In: *ACM Transactions on Applied Perception* 4.3.
- Werner, Annette (2007). "Color constancy improves, when an object moves: High-level motion influences color perception". In: *Journal of Vision* 7.14, p. 19.
- Willis, Janine and Alexander Todorov (2006). "First impressions: Making up your mind after a 100-ms exposure to a face". In: *Psychological science* 17.7, pp. 592–598.
- Wolf, R and Dick Block (2014). *Scene design and stage lighting*. Cengage Learning.
- Xu, Alison Jing and Aparna A Labroo (2014). "Incandescent affect: Turning on the hot emotional system with bright light". In: *Journal of Consumer Psychology* 24.2, pp. 207–216.
- Yantis, Steven (2001). *Visual perception: Essential readings*. Psychology Press.
- Zell, Eduard, Katja Zibrek, and Rachel McDonnell (2019). "Perception of virtual characters". In: *ACM SIGGRAPH 2019 Courses*, pp. 1–17.
- Zell, Eduard et al. (2015). "To Stylize or Not to Stylize?: The Effect of Shape and Material Stylization on the Perception of Computer-generated Faces". In: *ACM Trans. Graph.* 34.6, 184:1–184:12.
- Zhang, Fan, Huib de Ridder, and Sylvia C Pont (2018). "Asymmetric perceptual confounds between canonical lightings and materials". In: *Journal of vision* 18.11, pp. 11–11.
- Zhang, Fan et al. (2019). "A systematic approach to testing and predicting light-material interactions". In: *Journal of vision* 19.4, pp. 11–11.
- Zhang, Xiaobin et al. (2016). "Feeling light or dark? Emotions affect perception of brightness". In: *Journal of Environmental Psychology* 47, pp. 107–111.