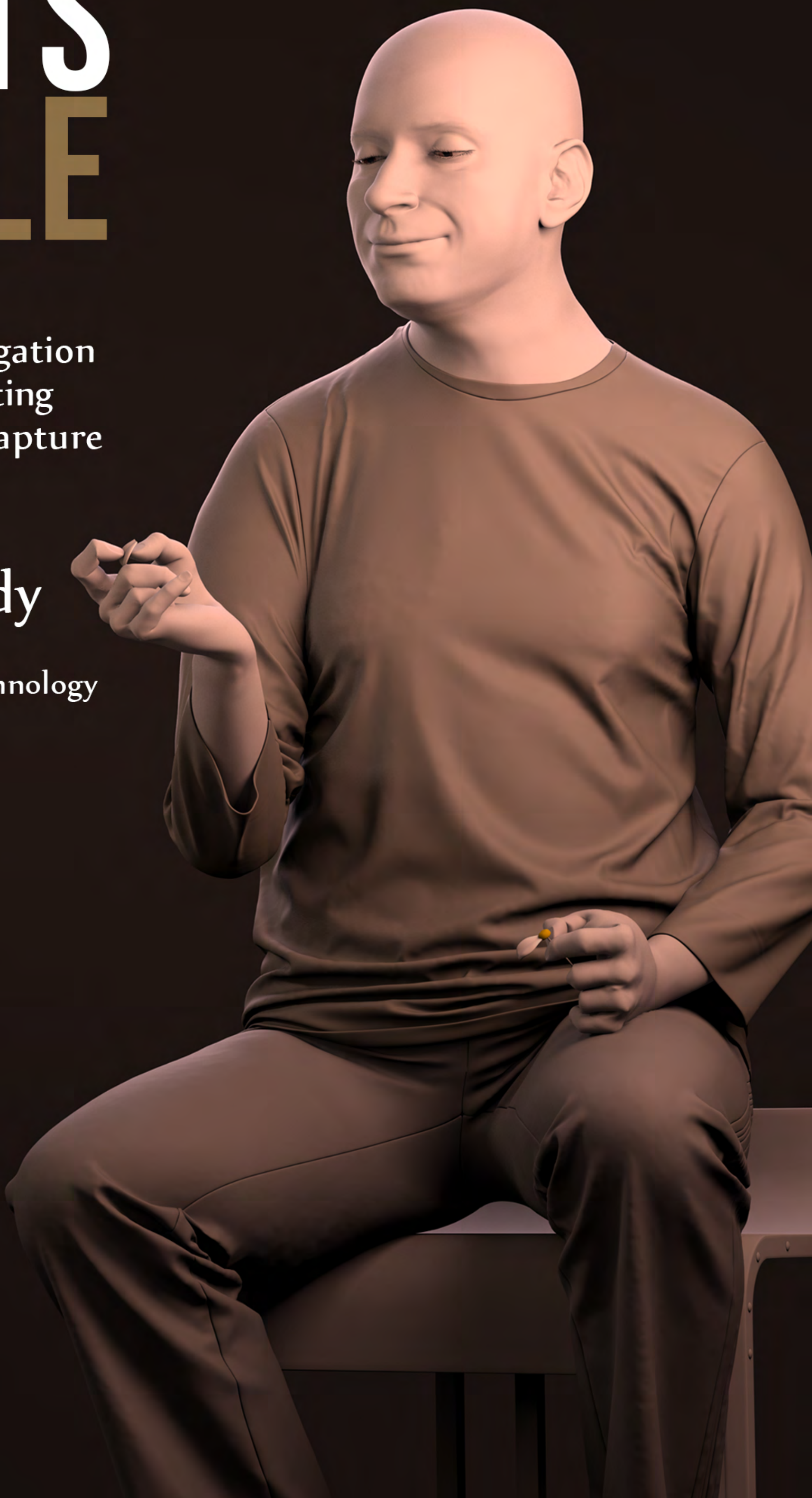
ACTING AND ITS DOUBLE

A Practice-Led Investigation of the Nature of Acting Within Performance Capture

PhD Thesis by
Jason Kennedy

Auckland University of Technology



Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Signature: ______________________________ Date: 26 January 2021

Acting and Its Double:
a Practice-Led Investigation of the Nature of Acting Within Performance Capture

Jason Kennedy

A thesis submitted to Auckland University of Technology
in fulfillment of the requirements of the degree of
Doctor of Philosophy

January 2021

Preface

Preface I. Abstract

This research deepens our understanding, as animators, actors, audiences, and academics, of how we see the practice of acting in performance capture (PeCap). While exploring the intersections between acting and animation, a central question emerges: what does acting become when the product of acting starts as data and finishes as computer-generated images that preserve the source-actor's "original" performance to varying degrees? This primary question is interrogated through a practice-led inquiry in the form of 3D animation experiments that seek to clarify the following sub-questions:

- What is the nature of acting within the contexts of animation and performance capture?
- What is the potential for a knowledge of acting to have on the practice of animating, and for a knowledge of animation to have on the practice of acting?
- What is the role of the animator in interpreting an actor's performance data and how does this affect our understanding of the authorship of a given performance?

This thesis is interdisciplinary and sits at the intersection between theories of acting, animation, film, and psychology. Additionally, this thesis engages with phenomenology and auto-ethnography to explore acting in performance capture from the perspective of a single individual as the actor, PeCap artist, and animator. This type of first-person experience-based insight is often missing from purely theoretical discussions about acting in performance capture and animation, and helps to provide a clearer understanding of the contributions of each creative role to the final PeCap result. This research provides a strong basis for the necessity of a paradigm revision for how acting is produced within a PeCap context.

Preface II. Keywords

Acting, animation, performance, motion capture, performance capture, MoCap, PeCap, digital performance, practice-led research, authorship, vactor, synthespian, digital counterpart, digital double, film production, actor training

Preface III. Acknowledgements

This project is wholeheartedly dedicated to my mother, Mary Patricia Yanacheak (née Kennedy). I frequently tell the story of how, when I was two-years-old, my mother read to me a booklet about dinosaurs published by the Sinclair Gas Company for the New York World's Fair 1964-65. I do not know where she found this booklet, but I was so fascinated by the dinosaurs in it that I asked her to read it to me every night. By the time I was three-years-old, I had memorised it – my mother's incorrect pronunciations of dinosaur names and all. My mother, a bartender, would occasionally take me to work. As a precocious child, I would sometimes ask a patron if he (they were all male) would like me to tell him a story. Upon his indulgence, I proceeded to recite the entire book verbatim, to which the patron would invariably lament to my mother: "Mary, this kid knows more words than I do."

At age three, I told my mother I would get my doctorate in palaeontology. I am sure I did not know what a doctorate was at that age, but she told me that the people who studied dinosaurs all had doctorates, and therefore I wanted to have one. I specifically went to university to study palaeontology, but as often happens when young, reality settled in and I realised I liked the idea of palaeontology more than actually doing it. A fortuitous viewing of "Monsters, Inc" in 2001 was the encouragement I needed to shift focus. I had always been fascinated by how the dinosaurs in "Jurassic Park" were created, and Sully's complex hair system in "Monsters, Inc" was so provocative that I realised I would rather animate dinosaurs than dig them up. At that moment, my path toward becoming a 3D animator was born. Yet, it really began seventeen years earlier with my mother reading to me. This is for you, mom; I fulfilled my promise.

Unfortunately, my mother died in 2015 a few months after I began this PhD. This research has outlived both my parents (my father, David Yanacheak, died in 2017) and at least five relationships. In some ways, that's a brutal figure, but it is a testament to dedication, perseverance, self-worth, and passion. In conceiving, executing, and writing a project of this magnitude, there are many people who contribute to its success without realising the role they play. I would like to acknowledge some of those people here. Khayreyah Oberon provided immeasurable support, humour, intelligence, and care as I toiled away at this thesis. Many of my colleagues have been great pillars of support and feedback, and have provided me with significant time to complete this work: thank you to Gregory Bennett, Dr Jan Kruse, Hossein Najafi, Dr Miriam Harris, John Piper, Dylen Velasquez, and Emily Ramsay. I would also like to acknowledge the many people who have provided support and feedback at conferences over the years, including Dr Lisa Bode, Dr Hannes Rall, Dr Mihaela Mihailova, Dr Birgitta Hosea, and Gray Hodgkinson. To my acting teachers Kacie Stetson and Dawn Glover, who have provided advice in so many ways over the years without even realising the roles they were contributing. To students who assisted me in various ways throughout this research, including Lily Martin-Babin, Marco Staines, and Drew English. To Farmehr Farhour and Arash Tayebi for their input and assistance with the Nansense motion capture system, as well as Lee Jackson for his tireless work in the AUT Motion Capture Lab. To friends who provided uncompromising support and love, including Jenilee Reddy, Shilpa Amin, Serenity Wise, Sonia Pinto, and Débora Borges. To my undergraduate professors at Albion College who demonstrated remarkable support and intellect, and who steered me along an academic path, including Doug Goering, Lynne Chytilo, Frank Machek, Robert Starko, Dr Robert AuFrance, Dr Ralph Davis, Dr William Bartels, Dr Ron Mourad, and Dr H. Eugene Cline. To the industry professionals who kindly gave me their time and attention, including Ed Hooks, Andy Serkis, and Dan Lemmon. And to anyone who has ever taught me a lesson, no matter how small. Teachers inspire greatness.

Finally, I wish to acknowledge the tireless, challenging, and insightful support of my PhD advisors, Dr Stefan Marks and Dr Frances Joseph, as well as Dr Sharon Mazer, who helped guide this project through its early conception.

Preface IV. Table of Contents

Preface V. List of Figures

Preface VI. Abbreviations

AU	action unit
BOE	Basic Object Exercise
CG	computer-generated
FACS	Facial Action Coding System
FK	forward kinematics
FPS	frames-per-second
IK	inverse kinematics
MoCap	motion capture
PeCap	performance capture
ROM	range of motion
RSO	retrospective self-observation
SWB	skin-weight-based
TA	think aloud
VFX	visual effects

1
OVERVIEW

1.1 Aims & Objectives

This research explores how we understand the practice of acting in performance capture for films. In this context, acting is explored from the position of the source-actor(s), the artists/animators responsible for adapting the performance to a virtual character, and the final on-screen digital counterpart. The realities of acting in performance capture draw into question many long-standing assumptions about how acting is produced and by whom. This thesis examines ambiguities in the performance terminology that is common to the experience of film actors and suggests updated definitions that help us better understand the nature of acting within a performance capture context. Through a practice-led inquiry involving two performance-captured animation experiments, this thesis provides both a theoretical and practical basis for understanding how performance capture leads to a highly mediated form of acting that often involves multiple authors. These production realities provide a strong case for the necessity of a paradigm revision for how acting is understood within performance capture.

This research focuses on a narrow subset of both animation (realistic character animation within mainstream narrative) and performance (as delineated by Method acting and techniques that descend from this tradition). This narrow focus is a conscious choice due to two main reasons:

1) the types of film characters examined in this thesis are well-represented by these categories; and

2) the scope and manageability of a thesis like this, where a full survey of all

possibilities of animation and performance would require significantly greater length.

By focusing on the cross-over between realistic character animation and an emotionalist approach to realistic acting, I am able to provide critical insights into the most common style of acting among motion-captured performances within live-action films.

The style of emotionalism I refer to here derives from a Strasbergian (Method) approach to acting. Strasberg provides a means for transfering Stanislavski's approach to acting for theatre to the process of acting for film, and therefore provides a bridge between stage and film acting (Section 3.2). This is important in that the process of acting within a motion capture context is at times similar to both stage acting and film acting.

Beyond just a focus on realism in terms of performance, this thesis is concerned with realism as a visual style understood within the context of visual effects production for film and animation. Between the visual poles of abstraction and hyperrealism within animation, there are limitless other styles and performance possibilities, and for this reason it is important to contain the scope of this discussion. This thesis is concerned with the performance of character animation within a realist style. That is, the animated characters that form the focal point of my analysis have real-world referents and behave "true to life". However, providing a meaningful definition of "character animation" is difficult and every definition I have found is recursive: generally, character animation is defined by animating a "character", which is a term entangled with multiple meanings in itself. I propose my own definition of *character animation* as any form of animation distinguished by its use of moving images to convey a degree of acting.[1] By this definition, graphical elements can range in form from the primitive – such as the eponymous 2D characters in Chuck Jones' "The Dot and the Line: A Romance in Lower Mathematics" (1963) – to the complex,

---

[1] This is as opposed to creature animation, which relies on figurative elements (generally non-human) to convey non-acting performances. Creature animation often does not necessarily result in acting but rather in various types of performance as outlined in subsection 4.4.2.

such as the articulated bodies of Bugs Bunny and the Na'vi aliens from "Avatar" (Cameron, 2009). The type and range of acting possibilities within character animation are determined by the style of the graphical elements used, and can likewise range from limited (graphic) to complex (cartoonal, realistic, or hyperrealistic) (Hosea 2012, p. 21).

1.2 Positioning the Researcher

This research sits at the nexus between two of my creative passions: acting and animation. I am a Senior Lecturer in Digital Design at Auckland University of Technology in Auckland, New Zealand. I have animated since 2003 and have taught 3D animation at the tertiary level since 2006. I am also an actor, and I have studied acting and performed in theatre and film productions in both leading and supporting roles since 1999. My undergraduate education at Albion College was liberal arts-focused, and my initial intention to become a palaeontologist was later supplanted by the more creative ambition to animate dinosaurs instead of digging them up. I completed a BA in studio art (drawing) with minors in geology and mathematics. I was also deemed to have an "unofficial major" in theatre by the Theatre Department due to auditing many of their papers and being an actor in many of their plays. I also completed four years of playwriting, which culminated in a musical that I wrote for my honours thesis. Unfortunately, animation was not a subject at Albion, so I completed an MFA in Electronic Art (Animation) at the University of Cincinnati in order to ground myself in the field. My master's thesis involved me acting in three short films that incorporated animation, visual effects, and live-action. At the University of Cincinnati, I was first introduced to motion capture by way of the MotionStar electromagnetic MoCap tracker by Ascension Technology Corporation. Most of my animation knowledge has been self-taught from a variety of books, tutorials, masterclasses, and Animation Mentor, a popular online-based animation training programme. Once I began working at Auckland University of Technology, I continued learning from and experimenting with optical motion capture systems available at the university.

I see myself foremost as an actor who uses animation and my physical body as my acting instruments. I combine acting and animation theories with practice to produce a multifaceted crossover between acting/animation and research/practice. The tacit knowledge I gain through producing my own animation provides a level of insight into the practical labour of 3D animation that may be missing from a purely theoretical understanding.

My experience as an actor and an animator positions me to participate in both of these roles in this research. Additionally, I believe that my experience across a variety of disciplines enables me to function as a liaison between specific technical languages and methodologies in order to produce an interdisciplinary practice. While currently there are few people who are both actors and animators, it is reasonable to assume that more actors will see MoCap/PeCap as an extension of their repertoire, and more animators will see how a knowledge of acting afftects their understanding of character performance. By being both the actor and animator in my experiments, I bring a unifying awareness to the performance, and in so doing provide a level of insight into all aspects of a virtual actor's performance that would otherwise be difficult to disentangle or elucidate. Hooks (2017) relates the story of Leonardo da Vinci being asked the secret of his talent, to which he responded, "Saper vedere" – "knowing how to see" (p. 13). I consider *saper vedere* an apt name for the technique I employ as actor-animator, whereby each side of that dichotomy provides insight into how to see the other side. This is an important contribution that provides balance to an anti-MoCap sentiment among some traditional animators (Furniss, 1999; Beiman, 2010) and animation theorists (Sobchack, 2009). By linking both theory and practice, this research provides a more positive (yet still critical) view of the potential for the role of motion capture within animation.

I am conscious of my own subjective preferences in the creation of this research, especially in terms of how I view and respond to the acting and animation I produce. It is important to attend to any unintended agenda or bias that could compromise my objective appraisal in terms of how the performances "read" and what they convey to an audience. This research is unique in that a single researcher is responsible for the entire process, from the creation of acting through the completion of the digital character on screen, including the evaluation of the performances at both ends. This type of research is generally done from either a highly specialised technical perspective or by an outsider who evaluates the production from a critical perspective. Subjectivity is a valuable part of my process in the creation of this work, but in order to maintain a degree of personal distance, I sought the feedback of an acting director with whom I have an existing professional relationship.

1.3 Structure of Thesis and Key Texts

This research deepens our understanding, as animators, actors, audiences, and academics, of how we see the practice of acting in performance capture (PeCap). It explores the intersection between acting and animation, and asks:

> What does acting become when the product of acting starts as data and finishes as computer-generated images that preserve, to varying degrees, the source-actor's "original" performance?

This primary question is interrogated through three sub-questions:

- What is the nature of acting within the contexts of animation and performance capture?
- What is the potential for a knowledge of acting to have on the practice of animating, and a knowledge of animation to have on the practice of acting?
- What is the role of the animator in interpreting an actor's performance data and how does this affect our understanding of the authorship of a given performance?

This thesis is interdisciplinary and sits at the intersection between theories of acting, animation, film, and psychology. Additionally, this thesis explores acting in motion capture from the perspective of a single individual as the actor, MoCap artist, and MoCap animator. This type of first-person experience-based insight is often missing from purely theoretical discussions about acting in motion capture and animation.

A glossary of key terms used throughout this thesis is provided on page 344. This glossary provides page numbers for when key terms are meaningfully introduced in the text.

1.3.1 Chapter Overview

The structure of this thesis provides for an extensive literature review, followed by a research methodology chapter, a chapter that discusses the practical implementation of this research, and a conclusion. This thesis also includes an extensive series of appendices that go into greater detail regarding technical processes related to the production of 3D animation and motion capture relevant to the practical experiments in this research.

The literature review is divided into three parts (Chapters 2-4). These chapters identify the core issues related to our understanding of acting within a performance capture context. They are structured so that each chapter provides a necessary foundation for the chapter that follows. Chapter 2 examines how the question of authorship of performance-captured roles has led to numerous controversies within the acting and visual effects industries in the past twenty years (Section 2.1). To better understand the technologies involved, this chapter presents an overview of definitions of animation and motion capture for the purposes of this thesis (Section 2.2). A breakdown of different motion capture technologies is presented to provide the reader with an understanding of the current "state-of-play" of the industry, and the advantages/limitations of the technologies used in the experiments in this thesis (Section 2.5). The chapter then turns to a discussion about the controversies surrounding the nature of ontology and indexicality between MoCap source-actors and their performance doubles (Section 2.7). This chapter provides a large-scale overview of current issues related to acting and authorship within performance capture, and opens these concepts to be interrogated in detail in the following chapters.

Chapter 3 introduces concepts of performance and acting that link to the discussions about

PeCap in the previous chapter (Section 3.1). This entails an analysis of key transformations to the concepts of performance and acting over the past 250 years, and specifically how we arrive at the Western realist style of acting within the Method tradition (Section 3.2). With this foundation laid, Section 3.3 discusses a model for identifying and understanding acting within performance-captured roles based on the degree of a character's technological mediation. This chapter establishes how non-living characters can be seen as performers, and leads to a discussion of how acting is produced for live-action and motion capture in the following chapter.

Chapter 4 provides meaningful definitions for terms common to an actor's experience that are often poorly understood or, at times, confused with one another. These terms are examined for how they are applicable within a performance capture context (Section 4.1). This proceeds to a discussion about the relationship between expressions and emotions, how expressions are understood, and a basis for how we empathise with animated characters (Section 4.2). Section 4.3 discusses how we define and understand screen presence among digital performers, while Section 4.4 provides a key contribution of this thesis: acting-centred definitions of various types of performance specific to digital characters produced via motion capture. The chapter concludes with a discussion about key differences between how live-action actors and animators participate in acting, and whether animators can be considered actors – and at what level (Section 4.5).

Chapter 5 provides an overview of the methodologies used throughout this research. Due to the scope and interdisciplinary nature of the research, several methodological approaches are employed and linkages between them are discussed here, including practice-led research (Section 5.1) a post-Merleau-Ponty phenomenology (Section 5.2), and autoethnography (Section 5.4). I detail ethical considerations for using myself as the sole actor in this research due to the vulnerability that the acting technique I use exposes me to (Section 5.5), and I provide a basis for not requiring ethics approval for two expert interviews (Section 5.6).

Chapter 6 presents the practice-led component of this research through two acting sessions produced via performance capture. Section 6.1 provides an overview of what the source-actor encounters when presented with the motion capture space, how to adapt to this, and what the source-actors' responsibilities are in terms of preparation and production. Section 6.2 details this study's first PeCap acting session based on an acting exercise designed to investigate subtlety in performance. Section 6.3 examines the role of the technical artists who are responsible for managing and retargeting the PeCap data onto the actor's digital counterpart, and Section 6.4 details the role of the animators responsible for preserving, interpreting, and adjusting the performance as needed. Section 6.5 details this study's second PeCap acting session, which involves creating a dramatic performance in the form of a scene from a stage play. Section 6.6 confronts the complicated issue of authorship within motion capture. This section contains interviews I separately conducted with actor Andy Serkis and Weta Digital Visual Effects Supervisor Dan Lemmon, both of whom respond to questions raised during the course of this research.

Finally, Chapter 7 presents my conclusions and a discussion of future work.

1.3.2 Key Texts

This research draws from a diversity of texts and it is useful to identify the "lay of the land" of relevant established research in the field, as well as the gaps in knowledge this thesis endeavours to address.

While there exist many examples of academic literature that discuss 3D animation and performance capture from a theoretical perspective, there are limited examples that examine these topics from a practical perspective. In the context of performance capture specifically, the two key practically oriented texts are Matthew Delbridge's PhD dissertation "The Cooling Steam of the Polar Express: Historical Origins, Properties and Implications of Performance Capture" (2014), and Shridhar Ravikumar's PhD thesis "Performance Driven Facial Animation with Blendshapes" (2018). To my knowledge, Delbridge's thesis is currently the only major scholarly work that specifically addresses concerns of acting and performance capture from practical perspectives. His research is supported by his experience in actor training and performance technologies at the tertiary level (Find an Expert, 2017). The central question of Delbridge's PhD asks: "[w]hat constitutes the field of performance capture in the early 21st Century?" (2014, p. 20). This interrogation pursues sub-questions that gauge the effectiveness of standard processes for capturing performance, how performance in physical and virtual space is to be understood, and how our understanding of performance capture can be expanded (2014, pp. 20-21).

The urgency of Delbridge's thesis is due to the limited number of of formal academic studies of PeCap (p. 16), especially in regard to the nature of performance within PeCap as it is captured in the moment, as opposed to how it is read by an audience afterward (p. 20). While texts for MoCap exist – such as "Mocap for Artists" (Kitagawa and Windsor, 2008), "The Animator's Motion Capture Guide" (Liverman, 2004), "Understanding Motion Capture for Computer Animation" (Menache, 2010), "The Mocap Book" (Tobon, 2010), and "Acting and Performance for Animation" (Hayes and Webster, 2013) – these are primarily "how-to" manuals that "overtly privilege the experience of the animator in the process of motion capture and discretely (yet significantly) minimise the experience of the performer, director and operator." (Delbridge, 2014, p. 42). By amplifying the voices of the artists responsible for producing performance, Delbridge's thesis provides a meaningful addition to the understanding of how performance is created in a PeCap context. However, while the technical language of performance capture is well articulated in Delbridge's thesis, what is missing is how performance language familiar to actors can be adapted to a PeCap context. He also considers performance capture from a more theatrical rather than filmic perspective, which means that thesis does not necessarily address performance concerns specific to PeCap within live-action filmmaking. Finally, because Delbridge is primarily concerned with using performance capture to produce live virtual performances, he does not discuss the role that animators play in developing the source-acting through to finished screen performances. These are all areas in critical need of academic study and which form critical components of this thesis.

By contrast to Delbridge's research, Ravikumar's thesis is based in computer vision and is technically focused, rather than performance-focused. Currently a software development engineer with the Amazon Video team in London, Ravikumar has previously worked as a software engineer encompassing research into computer graphics, vision, natural language, animation, and speech recognition and synthesis (Ravikumar, n.d.). His thesis is concerned with improving the fidelity of performance-driven facial capture results while making the modeling and capture stages more affordable, robust, and approachable (2018, p. 13). Ravikumar does this by providing a pipeline for automating actor-specific deformations for a 3D model of a given actor derived from 3D scanning or photogrammetry. He applies this workflow to a marker-based performance capture pipeline that utilises makeup patterns to train the software's representation of facial deformation, and which is capable of achieving better deformation optimisation with improved results in facial regions that are generally difficult to accurately capture with markers alone (2018, p. 13).

Key elements in Ravikumar's thesis are his explanations of the technical apparatuses that underpin the performance capture process. An inherent limitation of his thesis is that the scope does not extend beyond computer science. This is understandable and practical, and while his results intend to make motion capture technologies more user friendly for performers, the thesis does not discuss concerns from a performer's point of view *per sé*. Ravikumar conveys his explanations in the language of computer science, which I draw from and expand upon with language common to acting, animation, and performance in this thesis.

As a practice-oriented yet more theoretically grounded work, Birgitta Hosea's PhD dissertation "Substitutive Bodies and Constructed Actors: A Practice-Based Investigation of Animation as Performance" (2012) provides a performance-based counterpoint to the more technically oriented theses of Delbridge and Ravikumar. Hosea describes herself as a media artist, an inexact description used out of necessity to encompass her experience as an animator, video artist, interaction artist, motion graphics artist, theatre designer, and actress. Hosea indicates that her dissertation "represents [her] desire to synthesise these different areas that [she has] worked in, to explore theoretical concepts through interdisciplinary practice and to identify an area of practice that spans 'animation' and 'performance'." (2012, p. 2). Her research questions focus on ontology and performance within animation, drawing into question where a character exists, who is responsible for an animated character's performance, and the role of the animator in that performance. Additional questions are asked in regard to the relationship of performance between an animated character and the viewer, the most *apropos* of which being: "[i]s there a fundamental, ontological difference between animation and live performance?" (2012, p. 3). Similar to my own position, Hosea discounts the purely theoretical approach toward understanding animation that is predominant in the field. Instead, she engages with a practice-based methodology to experiment "*with* doing and *through* doing", whereby theory and practice inform each other as part of a "holistic search for knowledge." (2012, p. 5, *original emphasis*).

The animated work that forms the practical aspect of Hosea's PhD is a response to the notion that animation is a form of pre-recorded performance (2012, p. 159). Her work includes interactive public performances of animation featuring "live" feedback. Her research culminates in a one-on-one conversation between a human and an animated character named Miss Smith. This project examines whether the character of Miss Smith is determined by the animator as she is created, or if the character "is produced by the participant's act of engagement with her." (2012, p. 146). She concludes that "animation is a form of performance by proxy that extends the body of the performer" through a substitutive performance (2012, p. 163). As Hosea is not a 3D animator or motion capture artist, she does not fully extend the theoretical underpinnings of her dissertation to these specific categories of animation. Her discussion briefly touches on motion capture from a theoretical perspective, but the range of motion capture discussed is limited in terms of the types of technology and films explored. However, her findings and conclusions are useful and I endeavour to more fully apply them to 3D animation and motion capture in this thesis.

2

AND THE AWARD GOES TO…

2.1 Awarding Motion Capture Performances

During critical discussions about motion capture (MoCap), the conversation often veers toward Andy Serkis, who rose to Hollywood fame for his portrayal of Gollum, the feral hobbit in "The Lord of the Rings" trilogy (Jackson, 2001; 2002; 2003). Serkis has been called "the Godfather of Motion Capture" (Stern, 2014), "The King of Post-Human Acting" (Hiatt, 2014), and the "great character creator of our generation" (Medeiros, 2014). He has performed more stand-out film roles using motion capture than any other actor, including in "King Kong" (Jackson, 2005), "The Planet of the Apes" trilogy (Wyatt, 2011; Reeves, 2014; 2017), "The Adventures of Tintin: The Secret of the Unicorn" (Spielberg, 2011), "The Hobbit: An Unexpected Journey" (Jackson, 2012), "Star Wars: Episode VII – The Force Awakens" (Abrams, 2015), "Star Wars: Episode VIII – The Last Jedi" (Johnson, 2017), and "Mowgli: Legend of the Jungle" (Serkis, 2018). His fame is due to both his acting talent and the controversies he has generated within the animation community. Many visual effects artists believe Serkis minimises or elides the roles that animators play in the production of the performances of his digital counterparts. In an aptly-titled article "Where Does Andy Serkis End & Animation Begin?", Hannah Shaw-Williams writes:

> [i]t's easy to see why Serkis might become a target for the frustrations of under-appreciated animators and visual effects artists. He appears on talk shows, has been interviewed by countless websites and magazines […] – a reality that was only made possible by tireless animators. (2014, para. 10).

The origins of this controversy emerged during the lead up to the 2002 Academy Awards. In the "The Two Towers" (Jackson, 2002), the second movie in the "Lord of the Rings" trilogy, Serkis' performance was recorded via motion capture technologies and digitally mapped onto Gollum, a character whose physiology bears no resemblance to Serkis. "New Line

Cinema launched an aggressive campaign for Serkis to be considered in the Best Supporting Actor category" (Freedman, 2012, p. 44), and employed creative *legerdemain* in an attempt to convince the Academy of Motion Picture Arts and Sciences to accept for a live-action award what would normally be considered an animated performance. To accomplish this, New Line implied that Serkis was solely responsible for Gollum's performance, and refused a visual effects award from the Broadcast Film Critics Association so as to not run the risk of Serkis being denied "a straight acting nomination in the regular category" (Debruge, 2006, para. 8). According to the studio, Gollum "was pure performance, and motion capture was simply the means of recording this performance." (Freedman, p. 44). But how true is this assertion? Is an actor who uses motion capture when performing a role really the sole author of the screen performance of his or her digital counterpart? Or do animators deserve credit for crafting parts of the finished performance – and if so, how much? And why does this still remain a contested issue today, nearly twenty years later?

In order to answer these questions, it is necessary to ground ourselves in the unique nature of performance within motion capture. To do this, specific creative, theoretical, and technical understandings of key terms and concepts need to be established, especially for language related to performance that is often more "personally understood" than rigorously defined. This is the first of three chapters within a larger literature review section of this thesis and provides an overview of the state-of-play of the filmmaking industry as it interfaces with motion capture. Additionally, a basis for understanding the nature of ontology and indexicality within motion capture performances is explored. This provides a foundation for understanding the role motion capture plays in acting (Chapter 3) and the means by which acting is conveyed through motion capture (Chapter 4).

2.1.1 Gollum's Digital Makeup

Ultimately, Serkis was not nominated for a Best Supporting Actor role for his portrayal of Gollum in either "The Two Towers" or "Return of the King". The Academy struggled with the notion of what they considered a hybridised and heavily mediated performance

– whether Serkis' presence in the film was equivalent to a live-action role or whether the digital ontology of Gollum was too problematic for the performance to be considered "live" in the traditional sense (Burrill, 2005, p. 492). Serkis may have been Gollum, but Gollum was not Serkis – at least not in the sense of the character and the actor being one in the same body on screen (a relationship between actors and characters that has generally been taken for granted). The Academy considered this ontological difference to have potentially problematic repercussions should it allow consideration for Gollum's performance within the live-action category. Serkis expresses his own difficulty in coming to terms with this type of performance: "I had the first dawning realization that unlike any character I'd ever played before, one of the major challenges would be the fact that I didn't totally own the role." (2003, p. 18). Writing during the production of "The Two Towers" and "Return of the King", he credits himself with the voice, emotions, and movements for Gollum, but he expresses trepidation about the fact that his final onscreen body will be "taken care of by many very talented people" and "replaced by digital ones and zeroes" (p. 19). Despite the glib cliché, Serkis' situation raises an important distinction about how we understand and award performances. Can an actor's performance be recognized by the performance alone, or does Serkis' performance need to be performed through his body in order to be recognised as an "Andy Serkis performance"? To be more exact, how are we to understand the nature of acting in light of technologies that can separate an actor's performance from the actor's body?

The fact that Gollum and Serkis do not share the same body created one type of problem, but the Academy was also aware of ambitions within the visual effects industry to create digital clones of living and deceased performers. Using motion capture, other actors could *perform through* the body of the digital simulacrum of another person, leading to problematic performances that challenge agency and body autonomy. Likewise, the same motion capture data used to create one performance could be used to drive the performances of other digital bodies (Auslander, 2008, p. 170). The legal and ethical ramifications of such actions were mostly unexplored in 2002. How much can an actor's MoCap performance be manipulated by animators and directors before the actor is no longer seen as the creator or author of that performance (Pessino, 2004, p. 99)? According to Dan Lemmon, visual effects supervisor for the recent "Planet of the Apes" trilogy, such performances are driven by actors – but there is a significant distinction between a performance *driven* by an actor and a *solo* acting performance (Shaw-Williams, 2014, para. 12). The latter implies full control of the acting process from conception to completion; the former provides room for other creatives to enhance or attenuate the source-actor's original performance when it is applied through a digital counterpart. While popular industry rhetoric promotes a one-to-one performance mapping between source-actor and digital counterpart, numerous production realities contradict this assertion.

Due to the ontological uncertainty of where the actor and digital character begin and end, director Peter Jackson and Serkis adopted a new way to talk about the relationship between the source-actor and the character in motion capture. They promoted the concept of "digital makeup" to create a comparison between prosthetic effects as mere extensions of an underlying actor (such as John Hurt playing John Merrick in "The Elephant Man" [Lynch, 1980]), and the less familiar translation of performance from a human actor onto a digital character. However, what started as an attempt to reframe concerns about the legitimacy of Serkis' performance soon became an interpretation endorsed as fact.

Frustrated by reviewers who see him as merely lending his voice and movements to, or inspiring the emotions of, a performance, Serkis simply wants to be recognised as an actor performing a role (Lederer, 2012, para. 8). Serkis often says that he doesn't "see the difference between flesh and blood and motion capture", and that acting using MoCap is no different to acting by any other means (Fischer, 2005, para. 26). But how true is this? Certainly, motion capture provides the potential for a performance to be adjusted on many more levels than acting on stage or in front of a film camera. Without a rigorous understanding of the technical processes underpinning the translation of source-actor performance to digital counterpart, how can anyone be certain that a motion-captured role is fully the product of a single human actor or the combined performance of multiple creative parties? Balcerzak expresses doubt about this comparison to flesh and blood performance and the characterisation of MoCap as "simply a new form of acting – a kind of inevitable evolution for the art of performance in cinema" that preserves "the supposed autonomy of the actor" (2013, p. 195). He cites several critics who view Serkis' CG-rendered form as a digital prosthesis added in postproduction that preserves the actor's choices (pp. 195-196), but he views the comparison to prosthetics as lacking. Returning to Jackson's comparison with "The Elephant Man", Balcerzak believes that the viewer recognises the plausibility of someone existing beneath the "inches of prosthetics", whereas MoCap is a process of removing the actor's physical body and replacing it with something digital. Balcerzak counters the promotion of MoCap as a means to digitise an actor's performance, and instead frames the process as a means "to make the special effect perform realistically as opposed to […] digitally enhance the actor." (pp. 196-197). This draws into question the concept of "digital makeup" that Jackson uses to promote a direct linkage between Serkis' Gollum and Hurt's Merrick, whereby we are asked to believe that all of Gollum's performance is one-hundred percent faithful to the source-actor. According to Randall William Cook, the animation supervisor on "The Lord of the Rings" trilogy, "this is, frankly, a misrepresentation of the facts." (Amidi, 2014a, para. 8). A critical point of contention is the degree of "liveness" between a profilmic performance in prosthetics versus a performance recorded and re-mediated as motion capture data. Can the latter example be considered "live", and to what degree when compared to the former? Acting through prosthetics is a pro-filmic and additive process that preserves the actor while hiding parts of the actor's body; acting via motion capture is a subtractive process that removes the actor while encoding the actor's motion as data (generally at a vastly reduced level of detail compared to the original performance). The former produces an immediate result on screen and provides a scene partner with the ability to act against the final representation of the character, whereas the latter requires much more imagination on behalf of the scene partner to "see" the final character design in action. This requires a great deal more imagination when the final character design greatly varies from the source-actor, especially in terms of size, proportion, and species.

2.1.2 Liveness and Digital Performers

The history of the cinema star is also a history of the body within cinema, and specifically the "focus on the real biological individual, who has agency and, like us, is subject to mortality." (Bode, 2010a, p. 69). There is an immediacy to the corporeality of an actor as we watch him or her age and transform from one role to the next. There is a sense of the actor being in a real space smelling the air, feeling the breeze, and generally experiencing all of his or her senses. However, we don't normally associate these types of experiences with digital characters because we recognise that digital characters lack bodies in the most basic sense. How does this ontological uncertainty affect our understanding of the liveness of digital counterparts even when they believably pass for human actors?

Walter Benjamin cautioned that the act of filming "live" performance has the effect of separating the actor from the timeline of the performance (resulting in a series of performances edited together into the appearance of continuity), and requires the actor to play to the needs of a camera rather than participate in a relationship with a given audience. Benjamin believed that in both instances the critical element of "liveness" – the "aura" of the performance – was reduced or lost (1969, pp. 9-10). In his ground-breaking doctoral thesis on the nature of performance within motion capture, Matthew Delbridge argues that recording devices almost guarantee the loss of a performance's aura, and that within motion capture especially, performance neither "exists outside of, or separate from, the device" used to record it (2014, p. 159). Stage acting has historically been understood as more "live" than film acting, but now film acting seems more "live" by comparison to acting through motion capture. This suggests our understanding of liveness is historically-dependent and subject to the emergence of new technologies (Auslander, 2008, p. 169). While traditional film acting engages in a mimetic process with the actor, acting via motion capture disposes of the mimetic image altogether to instead record a trace of physical performance. How we describe the relationship between the "flesh and blood" performer and the digital performer is a subject of debate explored further in this thesis.

Performance within motion capture not only impacts our understanding of the relationship between acting and liveness, but also our understanding of what constitutes animation. Although the Academy struggles to recognise the hallmarks of live-action acting in MoCap performances, a fully motion-captured film could, in theory, be recognised for the Best Animated Feature Film category. However, as a response to an increasing number of animated films created using performance capture (a specialised application of motion capture discussed in Section 2.6), the Academy created a clear delineation for what it considers to be animation. Since 2010, the only entries considered animations are those "in which movement and characters' performances are created using a frame-by-frame technique." (Broadway World, 2010, para. 4). This excludes motion/performance capture on the very basis of how the performance is created and recorded, and underlines the struggle to legitimise MoCap performances as either 1) equivalent to live-action acting, 2) stylistically consistent to animation, or 3) a combination of both. The Academy pushed further and declared that "[m]otion capture by itself is not an animation technique" (2010, para. 4). As a result, motion capture is seen as not quite live action and not quite animation, leaving actors and films produced using this technology in a performance territory that is unrecognised at the highest level of achievement. While stellar motion capture performances may be recognised for their computer-generated imagery as part of the Best Achievement in Visual Effects category[2], the actual acting and animation that underlies those images are unrecognised within their own categories.

This would seem to call for the creation of a new awards categories to recognise achievements via motion capture. As far back as 2002 there was a call to create an Oscar for Best Synthespian Performance to recognise the fully digital incarnations of three popular fantasy characters from that year (Freedman, 2012, p. 41): Yoda from "Star Wars: Episode II – Attack of the Clones" (Lucas, 2002), Dobby from "Harry Potter and the Chamber of Secrets" (Columbus, 2002), and Gollum from "The Lord of the Rings: The Two Towers".

---

[2] The Best Achievement in Visual Effects category effectively recognises the gestalt of all postproduction contributions, including 3D modelling, texturing, rigging, animation, rendering, simulation, compositing, motion capture, and more.

However, the Academy expressed reluctance to add new categories (Thompson, 2008, para. 9), and has not done so since the addition of the Best Animated Film category in 2001[3] (The Telegraph, para. 2). As Freedman states: "Clearly, the question of how to view motion capture is an important one, as it illustrates the difficulties of defining different forms of film in the digital age." (2012, p. 38). Indeed, our relationship to actors and their screen images is undergoing a conceptual shift (Bode, 2010a, p. 69). While discussing the performance-captured Na'vi characters in "Avatar", Richard Brown describes this as the first film of the 21st Century: "It's unfair to take performances as good as these and not designate them as actors." (Block, 2010, para. 12).

In recent years, some alternative forms of recognition have started to emerge at smaller awards bodies. For instance, in 2018, the Austin Film Critics Association expanded its list of categories to include what may be the first of its kind: an award for Best Motion Capture/ Special Effects Performance. Five films were nominated for this award, each of which were attributed to a single actor deemed responsible for the MoCap performance. Actor Josh Brolin took home the award for his portrayal of Thanos (Whittaker, 2019) in "Avengers: Infinity War" (Russo & Russo, 2018). Recognising an actor for this type of performance is a good start, but it does little to resolve whether it is fair to present an award for the performance of a motion-captured character to just the source-actor. Who is responsible for the totality of the digital character's acting we see on screen? Is it just the actor who produces the original performance while wearing the MoCap suit, or does it also include the MoCap clean-up artists and animators who interpret the performance data later in the production pipeline, and who may also have a role in subtly (or sometimes heavily) altering the acting in a given moment of the performance? While there are numerous texts that discuss a theoretical understanding of the relationship between MoCap, animation, acting, and authorship (Furniss, 1999; Creed, 2000; Bode, 2005; 2006; Butler & Joschko, 2007;

---

[3] A new category for Outstanding Achievement in Popular Film was proposed in 2018 for inclusion in the 91st Academy Awards (in 2019), but was subsequently postponed while the Academy sought further input regarding the category (Kilday, 2018, para. 2). What effect, if any, this category will have on recognising motion capture performances is yet to be seen.

Bode 2008; North, 2008; Power, 2009; Sobchack, 2009; Bode 2010a; 2010b; Aldred, 2011; Allison, 2011; King, 2011; Freedman, 2012; Pallant, 2012; Balcerzak, 2013; Sito, 2013; Allison 2015; Bode, 2015; Bestor, 2016; Bode, 2017), these concepts remain largely unexplored from a practical perspective of the actual production of motion/performance-capture. By drawing from my experience as both an actor and an animator, this thesis attempts to unravel these issues and provide a foundation for how to conceive of the unique complexities within motion-capture acting.

2.2 The Technology and Production of Animation

Despite tensions within the film industry regarding the perception of motion/performance capture as genuine forms of animation, the finished products of animation and MoCap/PeCap[4] have more in common than what separates them. Bode discusses MoCap in terms of a hybridity between performance and animation (2008, p. 2), and indicates a discursive shift from PeCap being understood as "actor-assisted animation" to "technologically-assisted performance" (2008, p. 7). In this context, it is important to clarify what is meant by animation, especially *animation* within a *realistic* context.

Hosea defines animation as "mediated, moving images of a manipulated, artificial construct that could not have been photographically captured by a camera in real-time."[5] (2012, p. 22). She indicates that older definitions of animation that referred "to a hand-made, frame-by-frame process" are no longer meaningful in light of computer animation, in which "each frame no longer needs to be created by the animator." (2012, p. 22). Rather, the animator can map out "keyframes" and the software can interpolate shape and motion changes

---

[4] Hereafter referred to exclusively as MoCap unless discussing a distinction between the two. However, in general discussion, PeCap can be considered a subset of MoCap, and therefore referring to MoCap encompasses both terms. An expanded description of the relationship between both terms is found in Section 2.6.

[5] Presumably this definition does not include the cameras used to record motion capture. It could be semantically argued that the reference to a photographic process is intended to include traditional photography and videography methods.

between those frames. Animated motion can also be programmed through a variety of interfaces, including coding languages and procedural/node-based workflows within various software packages.

It is important to interrogate what Hosea means by moving images "that could not have been photographically captured by a camera in real-time." If we are willing to accept MoCap as a form of animation, it is arguable that some motion-captured performances could have been filmicly captured in real-time but were not due to other considerations (generally stylistic or safety). That is, in certain instances, the motion that drives the finished on-screen performance, regardless of the digital form such a performance may take, *could* have been recorded using standard live-action practices. However, such performances were recorded via MoCap to enable them to be retargeted onto a digital performance body. Such is the case with much digital stunt replacement work, whereby a source-actor may be recorded performing actions within a safer MoCap stage environment rather than on a profilmic set. The definition comes down to whether we understand animation based on the *process* of its recording or how the motion is *derived*. Hosea acknowledges that the definition she provides is a working definition for the purpose of her doctoral dissertation, and explains that her definition is limited in that "it continues to define animation reductively in terms of technical processes rather than to consider it more broadly and conceptually." (2012, p. 159). In a broader context, she embraces a definition where animation could be considered within the larger territory of performance (2012, p. 159). She presents "the case that animation is a form of performance by proxy that extends the body of the performer […] [through a] substitutive performance", and "[s]imilar to wearing a mask or manipulating a puppet, animation is a performance that is displaced onto the animated character" mediated through technology. (2012, p. 163). The act of displacement of performance is central to this definition, and provides a better fit for encompassing motion capture.

Norman McLaren, a pioneer of independent animation starting from the 1940s, provides one of the most durable and widely used definitions of animation:

> [a]nimation is not the art of drawings that move, but the art of movements that are drawn; what happens between each frame is much more important than what exists on each frame; animation is therefore the art of manipulating the invisible interstices that lie between the frames. (Thain, 2016, p. 168).

While McLaren's definition seems to privilege drawing, animation theorist Paul Wells sees the definition as a metaphor for any "activity that has taken place between what become the final frames of the film", such as adjusting the position of a stop-motion model into a new pose (1998, p. 10). McLaren's emphasis is on what happens between the recorded frames, both in terms of the historical production of each frame, and how the brain constructs movement and meaning between each discrete frame as it is witnessed in the theatre of the mind. Martinez identifies such "illusory movement" as the basis for a set of criteria to define animation regardless of the techniques used to produce it. Martinez defines animation as "moving image with produced and solely illusory movement. Its illusion of motion results from the creation or generation of a set of graphic positions with enough differences from each other to produce an intervallic illusion of motion." (2015, p. 44). The key point of difference in this definition is the requirement that the illusory movement does not reproduce real-time movement. That is, animation can be derived from records of real-time movement "as long as the resulting illusion does not reproduce the movement's real-time rate of change of positions but rather generates a new rate of positions" (2015, pp. 42-43). This definition would again challenge the status of MoCap as animation whenever the motion-captured movements are retargeted to a digital character at their originally recorded speed. However, based on this definition, if the rate of movement is altered even slightly, the motion-captured movements satisfy the criteria of animation. Such alterations to the speed of MoCap is common, especially when adapting the recorded motion to a different sense of scale between the source-actor and the digital counterpart (Serkis as King Kong, for example). Due to the fact that the audience is not privy to the original MoCap data to compare whether the movements of the digital counterparts exactly match the speed of their source-actors, this distinction becomes somewhat arbitrary in the case of motion capture.

For the purpose of this thesis, I combine elements of the definitions presented by Hosea and Martinez to form my own definition of animation:

> *Animation is a form of performance by proxy that extends and displaces the body of the performer onto a series of mediated, moving images of a manipulated, artificial construct set at positions with sufficient differences from each other to produce an intervallic illusion of motion.*

This definition importantly combines both aesthetic and performance-based consideration and, for the specific needs of this research, provides sufficient room to encompass motion capture. I will now briefly discuss an overview of performance as it relates to animation, and I provide an expanded understanding of performance in Section 3.1.

2.2.1 Animation Performance Styles

Animation offers perhaps the greatest range of performance possibilities of all visual media. One could discuss performance within abstract animation, such as in Oskar Fischinger's "An Optical Poem" (1938), which features early experimental performances of geometric objects to a symphonic composition. Or, on the opposite end of the spectrum, one could focus on the performance possibilities within hyperrealism, such as the character Neytiri from "Avatar" or Gollum from "The Hobbit: An Unexpected Journey". *Realism* and *hyperrealism* are used here to describe visual styles, as opposed to styles of performance. Within the context of visual style, *realism* refers to characters that appear true to life, whereas *hyperrealism* refers to characters that are both realistic and "exaggerated in comparison to reality" (Oxford Living Dictionary, 2017).

Hosea differentiates between *cartoonal* and realistic approaches to character animation. Cartoonal character animation involves a "heightened and exaggerated type of animated performance […] in which animated characters squash, stretch, exaggerate and otherwise defy the conventional laws of physics and human biology." (2012, p. 54). Cartoonal movement is often accompanied by a more cartoonal style, in which the character's sign system (e.g.: visual style) is simplified "in comparison with the organically integrated sign systems of embodied performance" (Silvio, 2010, p. 430). Animators frequently refer to the cartoonal style of movement as *believable* (rather than *realistic*) animation. In the context of visual style, a believable character is "one that provides the illusion of life, and thus permits the audience's suspension of disbelief (Bates, 1994, p. 122). The mention of the "illusion of life" within an animation context generally refers to the principles of animation developed at the Walt Disney Animation Studio during the 1930s. These so-called "12 Principles of Animation"[6] are considered the axioms of believable cartoon character movement. Such movement is derived from the observation of real-life movement, but which is presented in a caricatured or exaggerated manner. The greater the degree of this caricature, the more cartoonal the style of animation (Thomas & Johnston, 1981). By contrast, realistic character animation adheres to the laws of physics and biology, whereby every motion is distinguished by subtle, unrepeatable nuances based on a performer's "behavioural pattern" (Joon, 2008, p. 2). Realistic character animation results when an animated character's performance is indistinguishable from that of a profilmic actor, animal, or creature, and is best applied to visually realistic characters (like-style movement for like-style design). Since this research uses a live-action performer to generate realistic body and facial performance data, this thesis is primarily concerned with realistic character animation. Hosea cautions that the ubiquitous use of Disney's 12 Principles of Animation means that animators often indiscriminately apply the principles to their work when a closer observation of the natural world may yield better results. This is especially true when creating realistic character animation, as a reliance on principles alone becomes a "mechanistic technique that is learnt by rote rather than grounded in a critical analysis of nature." (2012, pp. 55-56). The practical work in this thesis is especially concerned with close observations of realistic movements and invokes the 12 Principles of Animation only in situations where a strict adherence to reality produces an unbelievable or uncanny result.

---

[6] For reference, these principles include: 1) timing and spacing; 2) squash and stretch; 3) exaggeration; 4) overlapping action; 5) arcs; 6) pose-to-pose and straight ahead animation; 7) easing; 8) anticipation; 9) staging; 10) secondary action; 11) solid drawing; 12) appeal (Thomas & Johnston, 1981).

2.2.2 Differentiating Animation from Visual Effects

The fields of animation and visual effects (VFX) frequently cross into each other, which can confuse how they are delineated. For instance, both fields make use of animated elements; however, based on the definition I provide above, animation *extends and displaces the body of the performer onto a series of moving images*. In this sense, animation is a form of embodied movement, even though the moving animated body doesn't have corporeal substance. This is different from the animated elements in a visual effects sequence that may take the form of moving particles or fluids, such as air, water, or fire. In these instances, the animation is meant to simulate real-world senses of particular phenomena, rather than to extend and displace the body of an animator as a performer. Dan North further differentiates "special effects" from "visual effects" in order to obviate another common source of confusion. He indicates that special effects are entirely pro-filmic (that is, recorded directly by the camera on set), whereas visual effects are "created or added to the image in post-production" (2008, p. 5).

Visual effects are generally understood to fall within two broad categories: "invisible" and "visible" effects (Fleming, 2012, p. 195). Invisible effects are designed to not be noticed, such as the insertion of background details to extend sets, alter the *mise-en-scène*, or otherwise compose an image without calling attention to themselves. Visible effects take centre stage and call attention to themselves, such as spaceships, spectacular explosions, digital body augmentations, and graphic/HUD elements. Creature effects such as the dinosaurs from Jurassic World (Trevorrow, 2015) are often considered visible effects and therefore a form of visual effects, but I would argue that creatures are still a type of performance by proxy, and therefore better understood as examples of animation.

2.3 Animating in 3D

During the late 1990s and early 2000s, the rise in popularity of 3D-animated films combined with the gradual decline in box office success of traditional 2D animated features created an imperative for many animators to upskill their hand-drawn toolsets to the digital realm. However, the tools for creating CG animation were far removed from the intuitive hand-eye relationship of a pencil or paintbrush. Instead, “artists had to accustom themselves to working second-hand through keyboard, graphics tablet, and computer screen”, the technical and incremental nature of which often disrupted the creative flow of ideas (Anzovin & Anzovin, 2005, p. 8). Working in 3D animation meant that artists needed to adapt to a structure of working posed by the software, which required “think[ing] successively like a designer, a sculptor, a painter, a puppeteer, and animator […], a lighting technician, and a camera operator.” (2005, p. 8). These steps roughly correspond to the following stages of 3D production, respectively: concept development/planning, modelling, texturing, rigging, animating, lighting, and cinematography/rendering. These are briefly explained below.

2.3.1 Production of CG Animation

The concept ideas for a computer-generated character are planned out either on paper or using a 2D software such as Adobe Photoshop. The character is then created in a 3D software through a process known as modelling. Modelling can take various forms, whereby the character is either created using digital sculpting software, such as Pixologic ZBrush, or is assembled from a series of polygons in a software such as Autodesk Maya. Both methods are commonly combined in modern practice. Texturing involves the application of reflectance models (shaders), colours, and textures to the model in order to provide appropriate surface details. Likewise, texturing also provides the ability to simulate how light interacts with a material to produce a heightened sense of realism (e.g.: subsurface scattering, a shader which calculates how light is absorbed and reflected by skin). Rigging is the process of creating a series of controls (called joints or bones) to form a hierarchical structure (called a skeleton or rig) that articulates the movement of the model, a practice that is vaguely analogous to creating a metal armature to support and move a stop-motion puppet. The animation stage is perhaps the most self-explanatory but the 3D animation process is more technologically-determined than its traditional 2D counterpart, as will be detailed in the following sections. To create illumination and mood, digital lights are added to the animated scene. In the case where the animated character interacts with live-action footage, lighting is critical to produce a believable integration. Virtual cameras are created to provide adequate framing and camera movement, and finally the scene is evaluated as a finished image through the process of rendering. As the scene advances to each subsequent frame, a new image is rendered, which results in the illusion of movement when all of the images are played back in quick succession.

The transition from 2D to 3D animation brought with it a number of other differences, as well. Within 2D animation, it was common for a lead animator to oversee a single character throughout an entire film. The lead animator was responsible for the main performance choices of the character, drawing out the keyframes and any other meaningful poses, while a team of assistant animators assigned to the same character created the in-between poses. In this way, the lead animator's role was more akin to an actor playing a character throughout a film, building the character's arc from narrative beginning to end. However, with the shift to 3D animation, animators are often assigned to a character on a shot-by-shot basis, eliminating any sense of ownership of the character's performance across the whole film. One or more animators may be assigned to a single character throughout the film, and the challenge is to maintain a consistent sense of character as the role passes through many hands. Animator Paul Naas indicates that "It's the most efficient way with respect to the production pipeline, but it makes it hard (if not impossible) to tell who did what." (Hooks, 2017, pp. 113-114). This fractured method of constructing a singular performance repeatedly plays an important role in determining the nature of acting in CG animation throughout this thesis.

Not only are the tools for producing 3D animation significantly different from those used in traditional 2D animation, but traditional animators often experience difficulty in achieving the same kind of plasmatic freedom[7] with software that they can easily achieve in drawings. Character technical director Javier Solsona explains that the 2D artist "can make use of squash and techniques that are easy to draw in 2D but hard to duplicate in CG. In CG you are bound to physical restrictions. You can only work on an enclosed, controlled environment." (Jones & Oliff, 2007, p. 87). One of the benefits to working in 3D is also one of its chief drawbacks: it is difficult for a 3D character to go "off model". A character that remains "on model" stays true to the size, shape, and volume of its original design, whereas going off model means that the character somehow diverges from the original design. Skilled draughtsmanship is needed to ensure a hand-drawn character remains on model, whereas a 3D model is a sculpture in space that requires a complex series of controls to achieve plasmatic flexibility. However, it is through deliberately going off model that animators can introduce added charm and appeal to their poses (2007, p. 90). This is a more significant issue when working with stylised characters, as more realistic characters are less likely to require off model posing.

Aylish Wood, a film professor at the University of Kent, provides accounts of animators experiencing compromised agency when using 3D animation software such as Autodesk Maya compared to traditional 2D methods. CG animators must "[work] against and with the out-of-the-box or automated qualities of the toolsets" as they "[negotiate] or even battle with the software." (2014, p. 325). While the software is capable of automatically interpolating motion between keyframes, it typically does so through the quickest and most efficient route, resulting in animation with a mechanical feel. The animator must intervene every few frames (and sometimes every single frame) in order to properly guide the motion to create a meaningful sense of timing, weight, and other relevant animation principles. (2014, p. 325). Whereas 2D animators are only bound by their drawing skills, 3D animators reap the benefits of automation while being subject to the technological structures and limitations of the software (2014, p. 326). Just like the 2D animator, the 3D animator is strongly encouraged to think through every frame of the animation and to make all of the decisions for the computer, rather than rely on perceived conveniences

---

[7] Sergei Eisenstein coined the usage of "plasmatic" in relation to animation when discussing the style of early Disney animated shorts. In this sense, plasmatic refers to a "rejection of once and forever allotted form, freedom from ossification, the ability to dynamically assume any form." (1988, p. 21).

afforded by automation (Kennedy, 2017, pp. 304-306).

2.3.2 Production of Photorealistic CG Characters

The culture of computer-generated animation, especially in the context of live-action integration, has arguably made the pursuit of realism its primary goal. Animation theorist Patrick Power argues that the terms "realism" and "naturalism" are frequently equated with each other within animation "to define a style of visual or audio-visual mimetic representation that aspires to photorealistic or cinematic verisimilitude." (2009, p. 108). This push toward realism is in part grounded in the historical uses of the 3D software to visualise engineering, design, architecture, flight simulation, and other predominantly technically-oriented solutions – as Power describes, "a legacy of ideologies of objectivity as opposed to subjectivity." (2009, p. 109). Scientific research is enabled by, and in turn feeds back into, 3D software, resulting in a science-minded approach in the development of new software features, especially in terms of achieving more "realistic" results. Fortunately, iterations of 3D software have also become more aesthetics-focused and artist-friendly in recent years, enabling a democratisation of the technical and artistic potentials of 3D animation, including the pursuit of realism (2009, p. 111). However, I argue that cinematic verisimilitude is at best a fleeting goal: even when realistic results are achieved, the perception of *realness* is easily undone by technological and artistic advances in subsequent films that set a new standard for what we accept as visually realistic.

Creating a fully photorealistic digital human capable of believable acting has long been the so-called "holy grail" of computer graphics (Ravikumar, 2017, p. 14). However, managing to produce both the physical and performance detail of a digital human face has presented the greatest number of challenges, leading some theorists to doubt whether such a goal is even achievable. The most promising approach to tackling the issue is to deconstruct the complexity of the face into a series of ever-smaller parts (Fleming, 2012, p. 103), from eyelashes to lips, to eyes, skin pores, expression-based blood flow, micro-displacements, and a host of other details. For example, when developing just the eyes for the character of Thanos

Figure 1. The face of Thanos. From Avengers: Infinity War, Russo & Russo, 2018.

(Figure 1) in "Avengers: Infinity War", digital artists at VFX company Digital Domain created extensive data around the eyelids, eye tissue, skin folds, eye wetness, conjunctiva, and layers of tissue that cover the eye in order to approach a photorealistic result (Seymour, 2018, para. 45). As I explore through my practical work in this thesis, this process of molecularization also extends to breaking down the movements of the face when animating realistic character acting (Sections 6.2 and 6.5). Bermano et al. note that "[subtle] details such as the twitch of an eyelid and the formation of small wrinkles significantly contribute to the realism of human faces and perception of emotions." (2015, p. 1). Realism demands that attention is paid to small details that often escape notice by the audience, but which would degrade the illusion of realism if they were not present. For instance, stubble and vellus hair (or "peach fuzz") was added to Thanos' face and head in order to enhance the realism of how highlights are formed on the skin, as well as to provide sufficiently high detail for closeups. In other projects, Digital Domain has even simulated expression-based blood flow (EBBF) to regions in the face, a primarily texture-based solution that helps to increase the realism of expressions like anger (Seymour, 2018, paras. 42-44). Creating such details both in terms of accurate modelling, texturing, rigging, and animation is a significant challenge that requires highly skilled artists. However, any one of these details is often insufficient on its own to elevate a character's sense of realism. Rather, each detail must work in concert with other details to enhance the perception of realism. For instance, as I explore in Chapter 4, the digital recreation of Peter Cushing in "Rogue One: A Star Wars Story" was one of the first uses of EBBF but was significantly less believable than the digital recreation of Paul Walker in "Furious 7" that did not implement EBBF.

2.3.2.1 Production of CG Animation

One of the most common techniques to enable robust facial animation is the creation of blend shapes (Kozlov et al., 2017, p. 76). Blend shapes are a set of 3D models that encompass a range of movements particular to a given face[8] (e.g.: raising or lowering an eyebrow, or opening an eyelid). The blend shapes all have the same topology as the base mesh (the 3D model's default pose in a neutral expression) and vary only in the arrangement of vertex[9] positions, which are generally confined to a basic expression or localised face region. (Zollhöfer et al., 2018, p.527). Using this technique, facial animation is achieved by linearly combining the neutral expression with the blend shapes in order to produce new facial shapes, generally via the adjustment of intuitive and semantically-meaningful animation parameters (Ravikumar, 2017, pp. 47-48). Blend shapes are sometimes derived as user-specific expressions, but they can also be "generated via statistical analysis from a large facial expression database, or can be hand-crafted by animation artists." (Zollhöfer et al., 2018, p. 527). Blend shapes tend to be more art-directable than other facial animation solutions, such as dynamic muscle rigs. However, in order to achieve a high level of realism, hundreds of separately modelled blend shapes are required. Due to the flexibility and

---

[8] Technically, blendshapes don't have to be confined to the face. However, when blendshapes are used throughout the rest of the body, they often take the form of corrective shapes that help to properly deform a model in a given region within a specific range of motion. In these instances, such blendshapes are often referred to as pose-space deformers.
[9] A vertex is a 3D coordinate in space that indicates a single point on the surface of a model. Each polygon in a character's geometry is defined by at least three vertices (triangles), but using four (quadrilaterals or quads) is common. A plane is drawn between these vertices, which is referred to as a face.

intuitive nature of this approach, blend shapes are especially helpful when applying PeCap data "to digital faces that differ strongly from the source model" (Li, Weise, & Pauly, 2010, p. 1), but they must work in concert with, and not inhibit, each other.

While PeCap technologies record many facets of facial movement, the granularity of the capture is not fine enough to record all of the subtleties of the deformations and timing. Such facets are paramount to fully replicating (or honouring) the original performance, but they are often difficult to reproduce through manual animation. As a solution to this problem, Bermano et al. propose a technique for adding fine-scale details to lower resolution PeCap data. They do this by extracting higher-level spatial and temporal details from 3D scans of a performer's face, including skin pores, wrinkles, and expressive deformations (2014, p. 3). The models extracted from these scans are added to a high-resolution performance database, which is then used to pair the art-directable low-resolution PeCap animation with its high-resolution counterparts (2014, p. 1). In a separate paper, Bermano et al. introduce a method for preserving the details of both static eyelid shapes and in the temporal formation of eyelid wrinkles (2015, p. 8). Garrido et al. make a case for the importance of accurate lip motion in virtual characters, as nuances in mouth expressions "strongly influence interpretation of speech and intent, and exact mouth motion is essential for the hearing-impaired relying on lip reading." (2016, p. 2). However, they note that both monocular and multi-camera PeCap systems have difficulty in accurately capturing the full range of possible lip motions, and they detail an approach to fully reconstruct complex lip shapes, such as a kiss and rolling lips (2016, pp. 1-2). While solutions such as these papers are freely available online, the actual implementation of these techniques is limited to proprietary software owned by various large VFX companies and is not clear if or when they will become available for artists who rely on standard commercial 3D animation packages (Ravikumar, 2017, p. 14).

An additional layer of photoreal complexity is the secondary motion generated by the muscles, fat, and bones in the face. Manually animating such motions is a tedious task and is often better accomplished through physical simulation. Layering this secondary motion on top of an existing blend shape-based performance helps to overcome some of the drawbacks of blend shape animation, such as "simple linear interpolation [that] does not handle rotational deformations, self-collisions, [or] physical interaction with external objects." (Kozlov et al., 2017, p. 76). However, the manual creation of the anatomical model of digital muscles, tendons, and tissues is a difficult and time-consuming process that requires specialised skills. Even so, such rigs are not always "capable of creating production quality results and require further artistic refinement" (2017, p. 76), and while some animators may be able to achieve flawless results, it is difficult and expensive to maintain consistent quality across large-scale productions (Ravikumar, 2017, p. 16). These further highlight the difficulties in approaching true realism when creating 3D character performances. This is additionally complicated by the fantastic expense and labour-intensive nature of achieving a seamlessly real result – a complication which may provide solace to flesh-and-blood actors who are anxious about being imminently replaced by digital performers. As Tekla Perry notes: "[it's] cheaper to hire even George Clooney than it is to use computers to generate his state-of-the-art digital double." (2014, para. 5). Regarding the production of "Gemini Man" (2019), director Ang Lee indicates that Will Smith's cyberstar counterpart Junior was twice as expensive as the A-list actor himself, and required a full year for hundreds of artists to create one shot of Junior talking (Kleinman, 2019, para. 11).

2.4 What is Performance Capture?

Special feature documentaries on Blu-rays and DVDs have led to a greater awareness of motion capture, although often from a cursory and heavily curated perspective. Typically, motion capture performers are seen wearing form-fitting Lycra body suits covered in Velcroed tracking markers (Kennedy, in press). Freedman believes that motion capture "exists as an unprecedented amalgam of both recorded and synthetic cinema" that draws into question "what constitutes animation in the digital age" (2012, p. 38). He asserts that motion capture has become its own mode of expression. This aligns with the views of Robert Zemeckis, a pioneering director of MoCap films such as "The Polar Express" (2004) and

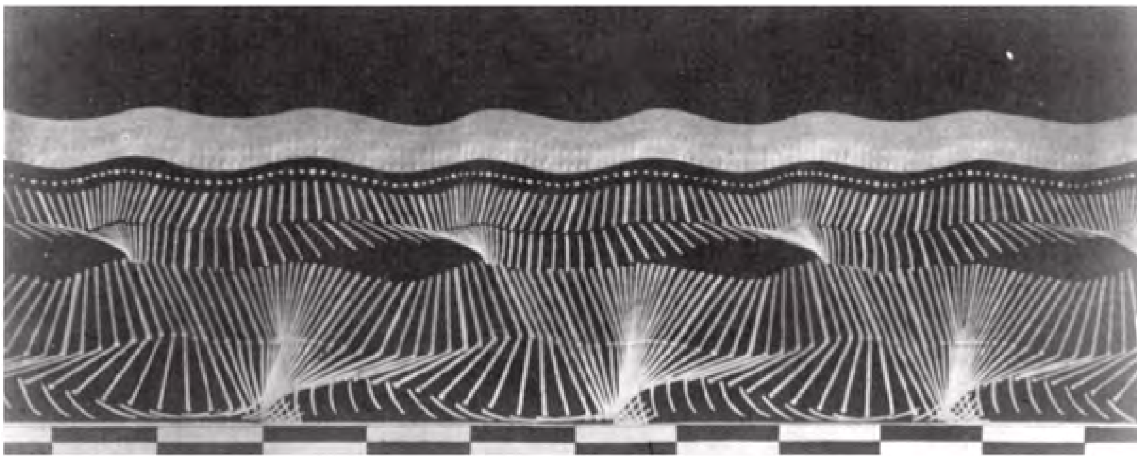


Figure 2. "Walking Man", by Étienne-Jules Marey, 1884.

"Beowulf" (2007), who believes that attempting to categorise MoCap using our existing understanding of synthetic versus recorded cinema potentially stifles a new art form (Freedman, 2012, p. 38). Motion capture finds its origins in the photographic experiments of Étienne-Jules Marey in the late 1800s, who developed a method for the creation of multiple exposures of a moving subject on the same photographic plate (Jensenius, 2013, p. 53). In "Walking Man" (Figure 2), a human subject wearing a black suit with reflective markers taped to his joints is recorded taking several steps. The result is an image that provides movement analysis of the temporal and spatial characteristics of a performed action (Jensenius, 2013, p. 53). Significant to this example is that the full body of the performer is hidden and reconstituted as an abstracted version of itself, presaging motion capture's conversion of human movement to data nearly a century later.

MoCap owes its origins to another early film invention – the rotoscope process. Developed by Max Fleischer in 1917, early rotoscoping involved projecting a frame of a film into a light box and tracing the actor in that frame. This process was repeated frame-by-frame in order to convert the profilmic recording into a moving graphic that possessed more fluidity than other animation techniques available at the time (Sito, 2013, p. 201).

Stemming from the Latin word *translatus*, translation refers to a process of carrying over, such as converting one language to another, moving people and objects from one location to another, or even changing a substance from one form to another. Rotoscoping and motion capture can both be understood as forms of translation. (Allison, 2011, p. 329). Rotoscoping was the first attempt to treat human motion as a form of data input that could be translated onto an entirely different graphic form. The result did not exactly copy the original image (a transcription) but rather retained the essence of the motion, creating a spectral effigy of the live performer in the animated result (a translation). This sense of performance translation and auratic transfer is central to motion capture, which has led some animators to derisively refer to MoCap as “the devil’s rotoscope” (Furniss, 1999, para. 1).

A range of definitions of motion capture exist and are primarily focused on its technological underpinnings. For instance, Menache describes MoCap as “the process of recording a live motion event and translating it into usable mathematical terms by tracking a number of key points in space over time and combining them to obtain a single three-dimensional (3D) representation of the performance.” (2011, p. 2). This definition is similar to descriptions of motion capture found in the narrow range of literature published for the benefit of people learning MoCap, such as in Kitagawa & Windsor (2008), Liverman (2004), and Hayes & Webster (2013). However, such past definitions were bound by the technical specifics of motion capture production at the time of their creation and do not necessarily fully embrace all possible application of MoCap today. Delbridge extends and modifies these definitions to produce a broader understanding of MoCap as “the capture of movement for a given purpose facilitated by an input motion system.” (2014, p. 18). Later in his doctoral thesis, Delbridge provides an expanded definition of MoCap that considers both technological and performance elements: “the process of digitally recording movement in 360 degrees and translating that movement onto a model in projected or screen-based 3D space.” (2014, p. 70). The trouble with creating an expanded definition that relies on technical specifics is that the specifics themselves are subject to rapid change. Although only a few years old, Delbridge’s definition of MoCap involving 360-degree recording is already ambiguated by the invention of 360-degree video, an omnidirectional camera process for simultaneously recording in all directions. A further issue with Delbridge’s definition is that MoCap data has the potential to be used outside of just projected and screen-based 3D space.

For instance, MoCap is used at the Stan Winston School of Characters Arts to drive the performances of animatronic characters (Caton, 2017). Also, the Pinoke Project at Deakin University involves the mapping of MoCap data onto robots to produce a transmedia dance performance that "[sets] out to answer the question of how a human and a non-human interact within a technological space." (Vincent, n.d., para. 5). Neither of these examples is screen-based but are still meaningful applications of motion capture.

While Delbridge's expanded definition works well for body performance, it lacks some of the nuance necessary to encapsulate all of the available methods for producing MoCap, such as markerless facial performance capture. As I detail below, performance capture is a subset of motion capture that is especially concerned with recording facial performance. An example of this technology is Faceware, which is a system that interprets monocular video of a performer's face without any tracking markers. The 2D video is imported into a software called Faceware Analyzer, which "[identifies] key features of the face and [samples] every pixel of it on every frame of video [… to] build a statistical model of [the performer]" (Faceware Technologies, Inc, 2017). Another software called Faceware Retargeter retargets (translates) this statistical model onto the face of a 3D character in order to drive its performance through a library of preconstructed facial shapes. Instead of recording 360-degree movement, the Faceware system solely interprets 2D data and reconstructs 3D movement from it. Considering this type of solution, and building on Delbridge's definition, I propose an updated definition of motion capture as:

> *the process of digitally recording movement in either 2D or 3D physical space and translating that movement onto a model/object in screen-based or physical 3D space.*

This updated definition accounts for how MoCap can be retargeted to both digital characters and physical objects, such as the animatronic characters and robots discussed above. However, it should be acknowledged that in the same way I am updating Delbridge's definition, I suspect that my definition of motion capture will need to be revised in light of future technologies.

MoCap incorporates a number of different systems, from optical motion tracking to non-optical methods such as magnetic, inertial, and mechanical tracking systems. These systems vary in how they capture data, as well as the purpose for how the data is used. In addition to the well-known use of MoCap in the entertainment industry (primarily films, television, and video games), the technology is also used to visualise and collect data for biomechanics[10], sports analysis, physiotherapy, dance, engineering, education, virtual reality (MetaMotion, n.d.; InspiredMag, 2013), and military applications (Delbridge, 2014, p. 70). The terms motion capture and performance capture are often incorrectly used interchangeably. While MoCap refers to any process of digitally recording and translating movement via an input motion system, performance capture exists as a subset[11] of MoCap that describes "the total recording of a performance without cuts using a motion capture system." (Delbridge, 2014, p. 11). MoCap sessions are generally segmented into producing clips of only a few seconds' length each, such as walk cycles. By contrast, PeCap sees an entire performance – such as acting or highly choreographed movement – "captured in one take, allowing traditional framing questions and dramatic devices to be employed post-performance." (2014, p. 11). As a result, PeCap reduces the need for multiple takes of a single scene. This greatly speeds up recording and allows actors to perform as if they were on stage – that is, continuously and within an ensemble – rather than in standard film practice where acting is performed through a series of starts and stops, sometimes with the actor isolated from the other performers in the scene. PeCap actors perform an entire scene in real time without standard device limitations, such as camera framing and lighting falloff (Delbridge, 2015, p. 4). Delbridge refers to this real time 360-degree frame as "The Omnisicient Frame", which is more akin to a 3D volume than a flat film frame. The Omniscient Frame is "a global frame, not hindered by the formally understood notion of the 'window' typified by the cinematic frame", which allows for performance and composition intentions to be interactively

---

[10] Among the first uses for MoCap technologies was to record and analyse human motion in biomechanics labs (Sturman, 1994, p. 1).

[11] *Subset* is used here to mean "a part of a larger group of related things". That is, all of the elements that comprise performance capture are contained within the larger category of motion capture.

experimented with throughout filming (Bancroft, 2014, p. 218) and continually adjusted in postproduction (Delbridge, 2015, p. 43). Film framing presupposes we are witnessing a "now" recorded from another time, whereas theatrical framing produces a fictious present. Delbridge explains that "PeCap reintroduces theatricality to the frame by extending it to include depth, capturing movement (not image) and presenting to us a captured and unmediated present." (2015, p. 45). MoCap is concerned with preserving the *physicality* of a performance, whereas PeCap is concerned with preserving the *theatricality* of a performance. In the context of motion capture acting, there is the live actor who performs and there is the actor's *double* – the on-screen transcendent performance of the animated character. The performance of a motion-captured animated character is always a construct composed of human intervention: the actor's *double*.

Current MoCap data is far less detailed than what can be immediately captured by a photograph or on film. The data itself must go through multiple layers of digital modifications before it can be viewed. Freedman indicates that MoCap "requires additional animation to succeed as a work of recorded cinema" (2012, p. 39), leaving much of the creative process to occur during postproduction (Furniss, 1999, para. 10). Performance-driven facial animation can involve either active (marker-based) or passive (video-based) methods to record and extract relevant motion data with maximum available fidelity to apply to an underlying animated face representation through a process known as retargeting (Ravikumar, 2017, pp. 18-19). Retargeting involves "[adapting] the parameters obtained from the capture stage and animating the virtual target character" (2017, p. 19), which can involve using the input data to drive a target character with very different physiology from the source-actor. In Ravikumar's examples in Figure 3, the human source-actor's performance is mapped to a stylised human head, an anthropomorphic lizard head, and an irregularly-shaped alien head. The nature of retargetting to each of these models varies based on how closely their physiologies match the source-actor. That is, mapping the source-actor to the stylised human head is a more straightforward process due to the physiological similarities between the two. Even though the stylised human character has different facial proportions to a realistic human, it still possesses a similar layout of shapes

Figure 3. The performance of an actor retargeted onto multiple characters of varying physiologies. Reprinted from *Performance Driven Facial Animation with Blendshapes* (p. 20), by S. Ravikumar, 2017, Doctoral Thesis, University of Bath. Used with permission.

and features. By comparison, the lizard head contains human-inspired features, such as a mouth that can produce human-like expressions, forward-facing eyes, and flexible ridges above the eyes that serve as eyebrows. However, retargeting the source-actor performance onto the lizard head requires much greater artistic judgment on behalf of the MoCap artist as the similarities between actor and character are more suggestive than exact. For example, greater creative liberties may need to be taken with a given expression on the lizard's head to achieve the same "feel" as the original expression on the source-actor due to these differences in physiologies. By greatest contrast, the alien head is so unlike the physiology of the source-actor that it is unclear exactly how the performance should be retargeted. The alien contains numerous eyes, all of which point in different directions, as well as a circular mouth directly in the centre of its orb-like body. The MoCap artist would need to perform a great deal of artistry and experimentation to provide a meaningful performance mapping between the source-actor and this creature design. These examples provide a good demonstration of how the more a character design differs in physiology from the source-actor, the more the MoCap artist must invoke artistic interpretation when retargetting, resulting in less of an exact (one-to-one) mapping of the original performance.

Motion capture comes with its own lexicon of technologically-derived terms, of which the most relevant to this thesis are the following: marker set, template, T-pose, range of motion (ROM), and capture volume. A marker set is used for optical passive motion capture and is composed of a series of reflective rubber ball markers attached via Velcro to a full body

Lycra suit. Each of the markers is attached to a specific location on the performer's body in order for the MoCap system to discretely recognise each marker and generate an accurate 3D representation of its location in space and time. A marker set is usually composed of between 35-50 markers (Delbridge, 2014, p. 71). With the marker set in place, the performer completes a series of basic body movements to calibrate the system's ability to recognise and accurately represent all of the markers, which results in a *template* specific to that performer and MoCap session.[12] The first of these movements is *T-pose*, which requires performers to stand with their legs directly beneath their shoulders and their arms laterally extended at 90 degrees to their torso (resulting in a T shape). The T-pose ensures that all markers are at a recognisable relative distance to each other and provide optimal visibility to the MoCap cameras. The T-pose is the standard neutral (or default) pose that a performer uses to begin and end each take. Bookending a shot with T-poses helps to realign any markers that become wayward due to the gradual degradation of the template during a shoot. Prior to filming, the performers also complete *range of motion (ROM)* exercises that help teach the MoCap software how markers on particular parts of the body move. Delbridge explains that "[e]ach ROM is unique to the individual performer and, while initially based on a prescribed set of movements, it adapts to match the individual performer's movement style and the specific characterised movement required of the character" (2014, pp. 74-75). Once the system understands the movement of a particular marker set, each marker can be given a unique name and other relational characteristics to form a cohesive set that is easily understood by a MoCap artist. This cohesive set is then used to control the motion of a proxy humanoid avatar, which itself is used to drive the motion of a 3D character (2014, p. 75).

The *capture volume* is the physical space in which the MoCap performance can be recorded by the system. The physical dimensions of the capture volume are determined based on the placement and settings of the MoCap cameras. Generally speaking, more MoCap cameras

---

[12] A new template needs to be created any time the marker set changes, which is generally at the start of each day's capture session. A previous session's template will not work because it is specific to the exact location of the markers as they were laid out in a given session, as well as the performer's posture and the capture volume variations in that session (Delbridge, 2014, p. 75).

result in a larger capture volume but also require additional space. If the performer moves outside of the limits of the capture volume, any markers that exceed those limits become "lost" to the system, but should snap back to their marker set when the performer re-enters the volume (Delbridge, 2014, p. 76).

2.5 Overview of MoCap Technologies

While a number of different motion capture solutions exist, this thesis is concerned only with approaches to MoCap that are common within the film industry. This section provides an overview of these specific approaches, their applications, and limitations.

As discussed in the previous section, the terms MoCap and PeCap are frequently, but incorrectly, used synonymously. All PeCap systems are a type of MoCap, but not all MoCap systems involve PeCap. In order to disambiguate the technologies discussed in this section, I will observe the following distinctions: whenever I refer to MoCap, I specifically mean a system that captures full body movements but which does not necessarily include facial capture; whenever I refer to PeCap, I specifically mean a system that captures both body movements and facial performance. To date, film-based motion capture has generally relied upon off-line (non-real-time) methods for recording and retargeting performance data. In situations where accuracy is less critical, real-time motion capture is often used to drive coarse-scale deformations of lower-detailed avatars. By comparison, off-line MoCap and PeCap methods provide greater reliability and performance fidelity when retargeting to photorealistic character rigs, especially in terms of facial performance capture. More recently, real-time data-driven methods have produced high-fidelity facial performance by "either learning a linear mapping between user expressions and high-quality target rig deformations or by constructing a target database of detailed subject-specific deformations that correlate to generic facial expressions performed by the user." (Zollhöfer et al., 2018, p. 537). These types of solutions, especially when coupled with advances in real-time rendering, are increasingly able to produce more believable photorealistic performances.

A significant barrier to widescale MoCap adoption and training has always been the cost associated with the equipment and its installation, especially in the case of optical systems. For instance, the 25-camera MotionAnalysis system used as part of this research was an investment of approximately US$320,000 when it was purchased in 2013 on behalf of Auckland University of Technology. Writing in 2017, Ravikumar indicates it is common for smaller optical systems to cost in the vicinity of US$80,000, which prices them outside the consumer level (p. 15). However, in recent years cheaper and easier-to-use alternatives have emerged, such as the Nansense inertial MoCap suit also used in this research, which retails from US$12,100 for its indie-level body suit and gloves (Nansense, 2019a).

While standard filming results in a specific encoding of information (either as an exposure of emulsion or the data of a digital image), the framing, focus, and frame rate remain impartial in a MoCap recording. This is achieved through both the nature of The Omniscient Frame and the high frame rate of the capture. By comparison to film, which records a bounded window between 24-30 frames per second (fps), MoCap systems record the entire capture volume at a standard rate of up to 250fps[13] (Delbridge, 2015, pp. 42-43). PeCap systems vary in their frame rates (although they are usually well in excess of standard film frame rates) which helps to generate more precise motion analysis and preservation of performance. Also, by natively recording at a high frame rate, this provides room for the MoCap data to be re-timed to a variety of film frame rates, from 24-120fps. Likewise, the high frame rate also greatly helps producing slow-motion sequences for analysing fast actions.

2.5.1 Marker-Based Systems

Marker-based optical motion capture systems involve a series of markers that are affixed to a performer's MoCap suit. These systems are classified as either passive or active depending on the way the system generates and interprets near-infrared light. Passive optical systems

[13] The MotionAnalysis Raptor-4 cameras used in this research capture at 200fps.

are the most commonly used due to their high accuracy in simultaneously tracking multiple objects, performers, and static props. Passive systems are wireless and traditionally rely on an array of 35-50 spheres (markers) coated in a reflective material and attached to the performer's suit at specific locations on the body. The optical system is composed of an array of up to 300 cameras pointed toward a given space, each of which emits a near-infrared light source that bounces off the markers and back to the emitting cameras. The cameras are calibrated to understand their own XYZ Cartesian spatial coordinates and to detect only the light reflected by the markers within the capture volume. A three-dimensional representation of a marker is produced in the tracking software whenever at least two cameras can see that marker (Delbridge, 2014, p. 71). When the software detects all of the markers as points within the virtual 3D space, it couples them with corresponding locations on a digital humanoid skeleton. This data is recorded and can be translated onto any virtual character to be animated, so long as there is a direct correspondence between the recorded markers and the character's rigging (Pizzo, 2016, p. 61).

VFX company Industrial Light & Magic use a proprietary adaption of the optical approach called "IMocap". Dubbed "faux cap", the IMocap suit was originally outfitted with a number of reflective markers similar to the typical optical setup, but also featured a series of rigid bands of a known size and series of patterns attached to the actor's head, arms, chest, waist, legs, and feet. More recent iterations of the suit feature fractal triangular patterns directly printed onto the suit fabric, as well as the bands being directly woven into the suits with updated colours and symbols. Instead of relying on infrared-detecting cameras within a MoCap volume, the IMoCap system uses image-based capture to calculate movement through two additional witness cameras. This allows the actors to work directly on the film set using standard lighting, even in difficult shooting conditions (Failes, 2019).

By contrast, active optical systems do not rely on the reflection of light. Instead, LED markers that each emit their own specific light signal are stitched into a MoCap suit, and the light is detected by cameras mounted in arrays similar to the passive optical setup. This method allows each marker to convey its own ID and therefore reduces the probability of the system confusing different markers. Since this system does not rely on a carefully managed infrared light setup, active markers can be used in an open environment with on-set lighting or practically no lighting at all. Since the markers emit near-infrared light, their signals are invisible and therefore do not affect the lighting on set in any way. A battery pack must be worn by the performer in order to power the active markers, so it is necessary to have a charger and multiple battery packs available in order to ensure uninterrupted recording (Pizzo, 2016, p. 61).

2.5.2 Markerless Systems

While marker-based systems offer both flexibility and accuracy, they also require greater setup and calibration time, as well as providing potential encumbrances for the performers wearing the MoCap suits. By contrast, markerless systems are less invasive and do not require specialised suits, instead relying on video sequences of the performer as the basis for tracking motion. However, this method poses its own difficulties, especially in terms of accurately estimating body poses from two-dimensional video along the missing depth axis. Likewise, the quality of the capture is limited by the lower available frame rate and resolution of video devices, and therefore both high speed movements and extremely subtle movements (varying by only a few millimetres) are less likely to be recorded with equivalent precision (Tinwell, 2015, p. 11). An early and popular consumer-level example of markerless motion capture was the Microsoft Kinect device, which tracked the full body silhouettes of video game players and translated their motions onto game characters in real-time for the Xbox 360 console. The advantage of the Kinect was its usability in a home environment in varying light situations, but as it was limited to a 640 x 480-pixel resolution at a 30fps frame rate (Tenney, 2012), the quality of its tracking was suitable more for home entertainment than serious film production. Another example is Simi Shape 3D, which is used for motion analysis of sports players. The advantage of this solution is that the performer is not limited to a specific capture volume. Instead, the performer is simultaneously recorded by eight cameras in a circle, the size of which can vary depending on the needs and nature of the performance. These cameras provide different angles to extract the performer's silhouette from the background and optimally fit the performer's motion along all three axes. As a result, this solution is especially well-suited for recording sports performances in a variety of environments, from football to ice hockey. (Simi Reality Motion Systems, 2014). It is important to note that markerless systems work best when the background is uniform, such as the grass on a playing field or the ice in a hockey rink, as it is more reliable for the software's algorithms to track the performer in these situations (Pizzo, 2016, p. 62).

Inertial motion capture systems are solutions that are both markerless and non-optical. Instead of markers, these systems rely on sensors placed on the performer (often gyroscopes stitched into the fabric of a MoCap suit) that are capable of transmitting their individual positions to the software. Similar to optical systems, the position data is translated onto a virtual skeleton, with better accuracy achieved through a greater number of sensors. Inertial systems, such as Nansense and Perception Neuron, are designed to be lower-cost alternatives that provide three benefits over marker-based optical systems: portability, no requirement for a specially equipped studio/lighting, and the ability to capture large areas without a specified volume. Also, the Nansense system provides high-resolution finger tracking via inertial gloves, which is a significant improvement upon the limited and lower-accuracy options for finger tracking among marker-based systems (Nansense, 2019b). A core drawback of inertial systems is that they often result in "motion-drift", in which the performer's virtual body may arbitrarily move in an unpredictable direction even when the performer remains still. This frequently results in footsteps that hover above or dip below floor level, hips that suddenly begin to float in an unexpected direction, or hands that unexpectedly interpenetrate other parts of the body.

2.5.3 Performance Capture Systems

Just as with motion capture systems, performance capture systems are divided between marker-based and markerless solutions. The first facial PeCap system used in large-scale film production was a marker-based system deployed by Sony Pictures Imageworks for

Robert Zemeckis' "The Polar Express" in 2004 (Gray, 2014, para. 12). This early system was specifically designed to overcome a production limitation in previous MoCap work, whereby the body was motion-captured in a normal capture volume and the face was recorded in a separate session while the actor tried to keep his or her head as motionless as possible. Without a system to directly record the facial actions, animators worked directly from footage of the actors to manually keyframe the facial animation to complement the motion-captured body. This approach led to difficulties in matching the performance intentions between the face and body sessions. The solution developed for "The Polar Express" provided the first example of a system that could simultaneously record both face and body performances (cpn_admin, 2012).

Modern PeCap recording devices generally take the form of a helmet with one or more arms positioned to either side of the face. Visual effects supervisor Dan Lemmon (personal communication, September 19, 2018) indicates that at Weta Digital, in order to provide a good fit to the performer's skull, a helmet is manufactured to precisely fit a 3D scan of the actor's head. This reduces the amount of unwanted secondary motion (viewpoint variance) that can occur when a performer moves his or her head too quickly with a poorly fitted helmet. Between one and four small, high-definition cameras are mounted to the helmet arms and oriented toward the performer's face. The cameras are optimised for either RGB-colour video (in the case of markerless capture) or for high contrast black-and-white footage (in the case of reflective marker-based capture). The facial markers are smaller versions of the same retroreflective markers used for body capture and are attachable to the face via spirit gum or a similar compound. Generally, between 100-200 such markers are glued onto a performer's face, with a greater density of markers around areas of high deformation such as the brows, eyes, and mouth regions. However, despite using even a large number of markers, the overall amount of facial performance data that can be extracted from the markers alone is merely a fraction of all the nuances occurring between the markers in the physical face itself. For example, during the development of "The Curious Case of Benjamin Button" (Fincher, 2009), visual effects supervisor Eric Barba lamented that the previous attempts at marker-based performance capture resulted in tracking that looked “stretchy and rubbery” (Duncan, 2009, p. 83). Animation supervisor Steve Preeg explained that more markers provides more detail about what the face was doing, “but they wouldn’t have necessarily told us what the character was feeling – and that’s what we need, a rig that would enable us to get emotion out of the character quickly.” (Duncan, 2009, p. 83). In order to provide a better solution for creating a believable likeness of the eponymous character in the film, which involved both digitally aging and de-aging actor Brad Pitt, Barba and director David Fincher opted for Mova Contour, a markerless system that they dubbed “e-motion capture”. Instead of markers, Contour makes use of fluorescent makeup applied to the entire face, and utilises stereo triangulation and computer image analysis to produce a “volumetric capture of a library of Brad Pitt’s facial expressions, which was then subdivided into thousands of ‘micro-expressions’ […] that enabled Benjamin to express infinite subtlety using facial movements that were unique to Brad.” (Duncan, 2009, p. 83). It is worth questioning terms such as “infinite subtlety” as these are often an exaggeration of the production realities that MoCap artists and animators actually face when working with particular systems. For instance, whatever “infinite subtlety” was achievable using Mova Countour in 2009 was qualitatively *less subtle* by comparison to the approach used a decade later on “Gemini Man”. Weta Digital used its proprietary “Deep Shapes” technology to simultaneously capture the shape and motion of wrinkle deformations with unprededented accuracy while calculating expressions starting within the deeper facial layers and working outward to the epidermis skin layer (Keane, 2019, para. 10). By comparison, the “infinite subtlety” of “Benjamin Button” was only surface-level, whereas Deep Shapes in “Gemini Man” provided greater subtletly from the inside out.

Both marker-based and markerless approaches to PeCap remain common alongside the development of new image analysis techniques to better understand and retain the nuanced details of a facial performance (Bermano et al., 2014; Bermano et al., 2015; Garrido et al., 2016; McDonagh et al., 2016; Kozlov et al., 2017; Zollhöfer et al., 2018). The following subsections provide a breakdown of the most common types of PeCap systems in film production today.

2.5.3.1 Performance Capture System

Monocular PeCap systems incorporate a single camera attached to a helmet rig. While this reduces the weight of the rig and occludes less of the performer's field of view, the single camera approach has traditionally provided less accuracy and completeness of the obtained reconstructions compared to multi-camera rigs. A common problem with 2D monocular systems is a missing depth axis, which can lead to an ambiguous solve compared to the true 3D shape of a face (Ravikumar, 2017, pp. 128-129). Without accurate depth information reconstructed from multiple views, "[m]ost monocular reconstruction approaches heavily oversimplify the real-world image formation process", relying on the tracking of a tight inner face mask region rather than the reconstruction of a full head (Zollhöfer et al., 2018, p. 541). Within this mask region are distinctive facial features including the eyes, eyebrows, pupils, lips, and jaw contours, but producing dense correspondences between these features across frames is limited by both uncontrolled lighting environments and the fact that most faces lack sufficient medium-scale texture details (Ravikumar, 2017, p. 128). For instance, the missing depth axis makes the movement of the lips both away from the mouth and toward the throat impossible to accurately calculate. Therefore, retaining such movements depends on the skills of MoCap artists to interpret and assign a meaningful degree of depth when retargeting to the 3D character, even though the degree of depth is not visible from the video data. Similar problems arise when the brows move toward the *sphenoid* bone (below the temple) and the cheeks pull around the *zygomatic* (cheek) bone. However, recently Garrido, Valgaerts, Wu, & Theobalt (2013) and Ravikumar (2017) have proposed solutions for monocular video input that overcome these limitations and generate high-quality animation with fine-scale details when applied to a blend shape rig, but such advancements remain proprietary and have not been incorporated into commercial-level software.

Due to the limited range of view inherent in a single-camera system, monocular systems rely on markerless approaches to recording facial data. Some of these systems are truly markerless and rely on pixel-tracking fine facial details preserved through high-definition video (e.g.: Faceware), while others make use of tracking dots drawn directly onto the skin. While only some monocular systems evaluate these dots as part of their tracking solution, in all cases the dots enhance the reference quality of the recorded performance by providing distinct landmarks for better human perception of small-scale facial movements. PeCap artists and animators benefit from these details when retargeting and adjusting the performance data on a 3D character.

2.5.3.2 Multi-Camera Systems

Multi-camera systems are often marker-based and provide the advantage of quickly producing accurate performance captures. However, there is a trade-off with the speed of preparing the actors, who need to have many small tracking markers affixed to their faces. This is a tedious and time-consuming process for both the actors and the PeCap team (Ravikumar, 2017, p. 137). The facial markers are also prone to falling off the actors' faces, which requires reattachment and recalibration whenever this occurs.

The raw motion data recorded by marker-based facial systems has a feeling of diminished expressiveness that can lead to character performances looking robotic or uncanny. Fincher compares this phenomenon to sandblasting the edges off a performance, and this issue played a key role in his decision to embrace a markerless strategy in the production of "Benjamin Button" (Duncan, 2009, pp. 83-84). While marker-based systems can record facial movements at a high temporal resolution (2000Hz), the markers provide only a low spatial resolution that doesn't account for perceptually significant details such as wrinkles and bulges (Ravikumar, 2017, p. 95). One commercially-available marker-based PeCap system is the Vicon Cara, which utilises four synchronised video feeds of the face. The four cameras provide a wide view of the face from different angles to improve facial tracking and 3D reconstruction. However, based on my own experience of working with the Cara, the system is bulky, intrusive, and heavy to wear, which can lead to head and neck fatigue after even short durations of use. Likewise, the markers are finicky and slow down production if they fall off and need to be reattached (sometimes requiring new templates and ROMS to be created). The markers can also be distracting to the actor, especially when placed on and around the eyelids, as the markers may feel noticeable or impact how the actor moves his or her eyes.

Monocular and multi-camera PeCap systems often rely on one or more “head lights” to brighten the actor’s face in order to get powerful and even illumination, especially if the set is quite dark. Lemmon indicates that it is best if the head lights are infrared and therefore invisible to the production camera, but this also requires that the PeCap cameras are sensitive to infrared light (such as no IR-cut filter on the sensor) (personal communication, September 19, 2018).

2.5.3.3 Photogrammetry and Mesh Propagation

While facial tracking relies on capturing and retargeting motion data, mesh propagation is concerned with capturing actual physical deformations over time. In this case, an actor’s face is recorded via high-speed photogrammetry at a very high resolution. Photogrammetry involves the near-simultaneous capture of photographs of a subject from a variety of angles. Subsequently, a photogrammetry software extracts 3D measurements and texture details from the 2D photographic data in order to construct an accurate 3D digital model of the subject. When these simultaneous captures are performed at sufficiently quick intervals, a “3D snapshot” of the performer’s head can be recorded once every 1/24 of a second in order to match the frame rate of film. When viewed in sequence, these 3D snapshots reveal the head in motion over the recorded interval, a process known as mesh propagation. This approach provides the greatest amount of performance accuracy from the recorded actor, but the data is hard-coded as digital geometry and cannot be easily animated or modified later. The data also does not provide a parametrisation of the face, which limits its usability in retargeting it to other facial models (Ravikumar, 2017, p. 47). Lemmon indicates that in his experience at Weta Digital, mesh propagation typically was not done live on set but rather in a separate session with the actor seated in a chair (personal communication, September 19, 2018). This approach appears to revisit the same problems seen in early performance capture when an actor's face and body were recorded separately, which often led to difficulties in matching the physicality and performance intent between the two takes. Despite this concern, Lemmon advocates for this approach as an alternative to multi-camera head rigs[14] (personal communication, September 19, 2018). Seymour indicates that while head-mounted camera units provide an actor with a more flexible and interactive acting environment, this comes at the cost of significantly less finely detailed tracking data. By comparison, a mesh propagation approach requires an actor to be seated within a controlled lighting rig that is surrounded by an array of high-quality computer vision cameras. An actor must produce solely facial performance in this setup while keeping his or her head and body still (2018, para. 19).

One of the best-known examples of the mesh propagation technique is the Medusa scanning rig developed by Disney Research Zürich. It has primarily been used to scan actors for blockbuster films such "Warcraft: The Beginning" (Jones, 2016), "Ready Player One" (Spielberg, 2018), "Avengers: Infinity War", and "Gemini Man". Dubbed "4D capture", the key information gained through this approach is a clear visualisation of how the face moves between key poses. This varies from blend shape solutions that focus on producing the key beginning and ending poses of an expression, but which can lack finesse in the temporal and spatial expression details between those key poses. VFX supervisor Dan Lemmon describes how the movement between poses, rather than the start and end poses themselves, is what actually produces a recognisable signature performance (personal communication, September 19, 2018). However, while blend shapes are easy to edit, mesh propagation provides limited opportunities for fine-tuning a performance after it is recorded.

In the production of "Infinity War", VFX company Digital Domain relied on a hybrid

---

[14] Lemmon indicates that there are advocates for more multi-camera head rigs who believe it is possible to get the same level of performance detail as mesh propagation by attaching more cameras to the head rig. However, due to the current camera sizes and the constructions of the rigs, he believes this approach is too distracting for actors and compromises their performances. He indicates that it is his preference to work with a monocular system over multi-camera setups because a single small face camera can be kept out of the performer's direct eyeline, allowing the performers to look directly at each other without obstruction (personal communication, September 19, 2018).

approach that married the best parts of a head-mounted camera with mesh propagation. Actor Josh Brolin's facial performance was recorded with a stereo camera rig, which was mapped onto a low-resolution 3D construction of the actor's face. Using artificial intelligence, Digital Domain transferred this performance to a higher-resolution version of the actor's 3D geometry. From this, Masquerade, a proprietary software developed by Digital Domain, used machine learning to extract high-resolution tracking data from the Medusa capture in order to convert the head-mounted camera's 150 facial data points into approximately 40,000 points of high accuracy facial motion data. (Seymour, 2018, paras. 15-20). This hybrid approach allowed Brolin's on-set performance with other actors to be converted into a high-resolution mesh that is adjustable by animators. Furthermore, the 4D data provided realistic transition shape data between blend shapes when compared to the standard non-linear combinatorial effect of combining digital expressions (2018, para. 29).

2.5.3.4 Performance Enhancement Models

In addition to the hardware solutions discussed above, a number of algorithmic and software-based solutions exist to better solve performance data. As noted above, head-mounted PeCap systems have difficulty with capturing fine-scale expression details. Bermano et al. note that "[a]s the desired level of realism increases, animators must spend increasing amounts of time to incorporate the nuances of deformation that are characteristic to a particular actor's performance." (2014, p. 1). Some of the subtlest expression details can elude even skilled animators (or the character rigs may lack sufficient flexibility to recreate such expressions), which can result in animation that is perceived as eerie or lifeless (see subsection 4.2.5). Bermano et al. propose a data-driven technique that encodes high-resolution representations of an actor's facial expressiveness and transfers those details to lower-resolution model of the actor (2014, p. 1). These details include subtle features such as small wrinkles, skin pores, and other expressive deformations, as well as temporal re-timing to produce greater fidelity of the source-actor's expressive changes (2014, p. 3).

One limitation to the method above is that it does not correctly evaluate the eye region due to a lack of accurate data within this region (Bermano et al., 2014, p. 10). Algorithmic solutions have been proposed to deal with region-specific difficulties in accurately reconstructing facial performance shapes, including the eyes (Bermano et al., 2015) and lips (Garrido et al., 2016). Human lips exhibit a large range of motion and deformation – from the stretching shape of a smile to the compressed and rolling shape of a kiss (2016, p. 1). The inner regions of the lips often overlap or become hidden by occlusion – for these reasons, lips are challenging to accurately reconstruct and track. An accurate reconstruction of lip shapes is essential to the creation of realistic and believable performances for digital counterparts, in part due to the fact that small differences in mouth shapes affect our ability to interpret speech and conveyed emotion. For instance, the so-called McGurk effect suggests that a difference between the expected mouth shape for a particular vocalisation can result in the viewer perceiving an entirely different vocalisation (2016, p. 2). While accurate reconstruction of lip shapes is difficult from even multi-view head-mounted camera rigs, Garrido et al. present a method for reconstructing detailed lips shapes from just monocular RGB video. They record a database of high-resolution 3D lip shapes through four cameras specifically recording the lips while six additional cameras record the rest of the face. They apply patterns to the lips via temporary tattoos in order to provide sufficient surface variance to ensure accurate tracking. They create a correspondence between the base mesh (default head geometry) and each lip shape reconstruction, and the resulting dataset consists of transitions in and out of each of these complex shapes (2016, p. 3). However, while the algorithmic solutions provided by both Bermano et al. and Garrido et al. are freely available as white papers online, they are solely implemented in proprietary software used by large-scale VFX companies. This is typical of cutting-edge performance enhancement models: similar solutions have yet to be incorporated into industry-standard 3D animation software such as Autodesk Maya and therefore remain out of reach for most 3D artists.

Figure 4. Motion Analysis motion capture cameras in the AUT Motion Capture Laboratory. Photo by Stefan Marks, used with permission.

2.6 MoCap at Auckland University of Technology

The research in this PhD is specific to a cinematic approach to MoCap and was be conducted using motion capture equipment available through the Auckland University of Technology Motion Capture Lab. This included a MoCap system composed of 25 Motion Analysis Raptor-4 cameras (Figure 4) to track the body and a Faceware ProHD PeCap system to track the face. The Motion Analysis system operated at 200fps with a resolution of 2352×1728 pixels. The capture volume of the Motion Capture Lab was 6×6×3 metres and the system could accommodate up to six simultaneous performers. Faceware ProHD is a monocular head-mounted camera system that analyses and retargets a facial performance from RGB video at 59.94fps and 720×1280 pixels.

Additionally, motion capture researchers Farmehr Farhour and Arash Tayebi at The University of Auckland (New Zealand) lent me a Nansense Indie full-body motion capture suit and Nansense Pro gloves for one of my dramatic performance capture session (see

Section 6.5). The Nansense suit produces inertial MoCap data at 90fps and the gloves capture accurate motion of all fingers, which is a superior solution to what is available in terms of hand-capture using a marker-based optical MoCap system like Motion Analysis.

The Motion Capture Lab also possesses a photogrammetry setup composed of ten Canon 700D DSLR cameras. While this is not a sufficient number of cameras to produce a full 360-degree simultaneous capture of a seated performer, the cameras are arranged in an array that allows for meaningful coverage from a front perspective. Once a front perspective is captured, the subject's chair is rotated 45-degrees and the capture is repeated. This process continues at 45-degree intervals until a full circle is completed. The performer is asked to retain his or her expression as well as possible throughout the full capture. However, a performer may succumb to stress or muscle fatigue, which can lead to a diminution of an expression. This process is performed as quickly as possible to minimise the risk of this happening.

2.7 The Ontology of Digital Performers

Within film, human performers are generally understood to be the sole creators of their own performances. However, when we attempt to understand the ontology of digital counterparts, the flesh-and-blood performer becomes one mere component within a digital performance matrix. The performance is encoded as data and becomes part of a virtual environment (Delbridge, 2014, p. 159), resulting in a type of screen performance that is not entirely human nor entirely digital in origin, and where the human is no longer "the sole locus of enactment" (Salter, 2010, p. xxxii). What implications does this have for our understanding of the indexicality of the performances of digital counterparts? Are digital performances any less *real* than flesh-and-blood performances? What is the nature of the digital character's existence and how does this relate to the corporeality of the human actor responsible for the source performance?

2.7.1 To Be or Not to Be (Human)

The concept of a digital performer immediately evinces a distinction between a living (real) performer and a synthetic performer, but how do we conceive of the synthetic performer? Is the fact of its existence any less real than or qualitatively different from a living performer, and if so, how? Salter writes that reliance on a dualistic framework between the human and not-human in the context of performance is complicated by how and where we locate agency. He suggests the anthropological concept of *entanglement* as a substitute, whereby we understand that "human and technical beings and processes are so intimately bound up in a conglomeration of relations that it makes it difficult, if not impossible to tease out separate essences for each." (2010, p. xxxii). Indeed, it is not safe to say that humans are solely perceived as flesh-and-blood organisms: there is a history from Aristotle onwards of the body envisioned as a form of technology itself – a musculo-skeletal, mechanical-organic system. The early twentieth century greatly expanded the idea of the technologized body as a machine, and the twenty-first century embraces the prostheticized, computationally augmented body as its successor. The body is

> [n]o longer conceived as a machine in the traditional sense of an organized heterogeneous assembly of parts, performance theorists and practitioners now see the contemporary body as something incorporated into larger than human systems—as something to be transcended through implants, prosthetics, sensors, actuators, and even genetic invasion. (2010, p. 221).

In director Andrew Niccol's film "Simone" (2002), Victor Taransky, a director who discovers a software that can replace his problematic actress with a digital counterpart named Simone, remarks: "Our ability to manufacture fraud now exceeds our ability to detect it." While this proclamation was arguably premature in the history of digital characters, it demonstrates a key anxiety about the performances of digital counterparts. The fictional actress Nicola Anders is replaced by Simone in each of her scenes in Taransky's movie. Simone is so successful that her audiences never recognise her as a synthetic character, and she goes on to develop a successful acting and singing career. When Barbara

Creed coined the term "cyberstar"[15] in 1999, this was the type of character she envisioned: a digital performer that can wholesale replace carbon-based actors (2000, p. 80). Many actors worry about their shelf-life and whether they could one day compete against cyberstars that are impervious to aging, work continuously in any conditions, perform dangerous stunts, never shy away from nudity, seamlessly interact with other CG characters, fully change their appearance as needed, and never expect compensation for their work. Speaking in 2005 regarding "King Kong", Peter Jackson comments that

> digital doubles will never replace actors. […] And I think it's a lot of old nonsense really when people say, 'Oh, well, we won't need actors anymore, we've got digital people.' But digital people don't have hearts and souls, and they can't provide everything an actor can provide in a performance. (Allison, 2011, p. 333).

To date there is no cyberstar who can fully pass for a living human, nor a digital actor who can believably perform on its own without human input. In short, the threat of the cyberstar never emerged because true cyberstars do not (yet) exist. Instead of cyberstars, today's cinema is awash in digital characters that look believable and produce compelling performances, but this is always the product of the contributions of many human artists. What has changed is how an actor's work contributes to the final film. Today it is common for live-action shots to pass through layers of visual effects postproduction, even in films which seemingly have little if any visual effects.[16] This renders both the actor and the live-action *mise-en-scène* as merely individual components in the postproduction of the final image (Kiwitt, 2012, p. 18), but which, when done well, still allows the audience to emotionally bond with the mixture of carbon- and silicon-based elements (Pizzo, 2016, p. 41).

2.7.2 How Real Are Digital Characters?

Owing to its more obvious indexicality, live-action presumes a more privileged connection to

---

[15] While the term *cyberstar* has become outmoded in academic discussions, I wish to reclaim it as a specialised category of digital counterpart performance alongside other terms such as *vactor* and *synthespian.* I disambiguate these terms in Section 4.4.

[16] Such films often utilise invisible visual effects to replace the filmed sky with a more mood-appropriate alternative, create season-based variations to the time of year of a filmed shot, or produce digital set extensions, among countless other possibilities.

the “real” than animation. Some film theorists perceive the coded reproducibility of a digital character to be at odds with an understanding of performance as a creation inextricably linked to the present moment (Shacklock, 2016, p. 76). By contrast, Joanna Bouldin believes that in animation (and, by extension, visual effects elements that are wholly computer-generated) traces of the “real” are maintained through “a complicated, morphed and multiplied connection. […] [T]he animated body is able to draw upon multiple originals – from models to voice actors to the animators themselves.” (2001, p. 48). She argues that not only does the animated body exist as a particular kind of hybrid for the digital artists responsible for its creation, but it also exists as a hybrid from the perspective of viewers who experience it as an element in film/television, comic books, video games, amusement parks, toys, and manifold other incarnations. Bouldin compares the animated body to the Body without Organs (BwO) (2001, p. 49), a concept developed by Gilles Deleuze in “The Logic of Sense” (1969) to reflect upon the experiences of playwright Antonin Artaud. Deleuze and Guattari describe the BwO as:

> […] opposed less to organs as such than to the organization of the organs insofar as it composes an organism. The body without organs is not a dead body but a living body all the more alive and teeming once it has blown apart the organism and its organization. […] The full body without organs is a body populated by multiplicities. (1987, p. 30).

Linking animation to the BwO, Bouldin describes the animated body as a “complex and constantly shifting assemblage formed from connections with a wide array of cultural phenomenon, biological bodies, technologies, and media.” (2001, p. 49). Historically, character performances on stage and in film are portrayed through the actor’s body. In MoCap and animation, the BwO challenges the established structure of the actor as the sole author of a character’s performance; the character we see performing – the actor’s double – is a simulacrum of something that, in most cases, does not exist, but which can still be understood as real. Bouldin argues that we experience animated bodies as real bodies, even if the animated bodies lack the indexicality of live-action (2001, p. 49). However, does the knowledge that an animated body is not a *living* body have any impact on the degree of realism we attribute to it – and more importantly, does this knowledge present a barrier to emotional engagement with the animated character? Creed imagines that if viewers become aware that a character on screen is synthetic rather than organic, the knowledge that the emotions are produced via a synthetic character will result in an emotional roadblock (2000, p. 84). Creed can be forgiven for this view as she authored her critique at least eight years before there existed any examples of digital characters whose performances (and the methods used to create those performances) could challenge her argument. For instance, Fleming believes that in Fincher's creation of Benjamin Button, any sense of the character's "false" ontology is subverted through "endowing his digital actor with a human element that grants the character a 'real' psychological depth and complexity" that attains a spiritual dimension to the performance (2012, p. 197). In this case, the human element is captured and preserved through the care taken to preserve the detail in Brad Pitt's performance through the PeCap process. Benjamin Button's body is further grounded in reality through the filming of age appropriate stand-in actors on set, onto whose bodies Pitt's digital head is superimposed. Still, the question remains whether the mere knowledge that a character is synthetic (in whole or in part) has any effect on how viewers perceive its "realness". Since the indexicality of live-action actors is considered self-evident, how do viewers reconcile the link between source performer and digital counterpart when the two do not share evident indexicality?

2.7.3 The Indexicality of Digital Characters

What is the nature of the relationship between a source-actor and a digital counterpart? While an actor's image on screen has traditionally been seen as an index of his or her real-world performance at a given time and place, the digital counterpart's presence is not constructed by such an immediate or apparent linkage. Fleming suggests that live action's privileged relationship to the real reflects a traditional binary opposition between the human (indexical live action) and posthuman (digital/animation). He proposes that a Deleuze-and-Guattarian lens challenges this paradigm of fixed identities and encourages us to see the human and posthuman as an artistic assemblage of expressive relations, whereby the source actor is in a process of becoming-digital while the posthuman (digital counterpart) is in a process of becoming-human (2012, p. 197). By viewing the human and posthuman as poles along a continuum that locates sites of performance, we can appreciate animation's ability to create protean possibilities for the live-action actor who is becoming-digital, while the digital character is grounded within a more indexical reality as it undergoes becoming-human. This ties in with Steve Dixon's assertion that performance artists "explore and enact their holistic autonomies and interiorities (gendered, spiritual, emotional, and political), not simply their bodily corporeality", and that if this occurs within a digital environment, "it is the medium that is virtual, unreal or disembodied, not the human performer within it. […] [T]he medium is not the message (and never has been); the performer is." (2007, p. 215). Dixon cautions against a tendency he perceives within cybertheory and digital performance studies to relate the fragmentation of the human body within the digital realm as true corporeal transformation or disembodiment. Instead, he sees the virtual body as an indexical trace and representation of a living, physical body. He suggests that the performing virtual body is just as authentic as the body of the live performer. Rather than the barrier to identificatory engagement that Creed imagined would doom the cyberstar, Dixon suggests that audiences detect the indexicality of a virtual body and are able to cognitively and empathetically engage with it as if it were "already embodied material flesh." (2007, p. 215).

It should be noted that Dixon specifically theorises within a performance artist context here. However, I suggest that his assertion is also applicable to performers within a motion capture context (whether they be *auteurist* performance artists, actors, or otherwise). That is, while performance art is not necessarily grounded in realism, actors working in a realistic context may also engage with their interiorities and autonomies (whether whole or in part) as part of enacting a character. These processes are not exactly identical but the two share much in common. However, a significant point of difference is that within a performance art context, the virtual counterpart's performance is generally seen as the product of just the performance artist. That is, the process of becoming-digital and becoming-human are two sides of the same coin, equally weighted and complimentary in process. However, within a MoCap context, a single performer may produce the performance that drives the digital counterpart (becoming virtual) while multiple MoCap artists work to interpret, apply, and enhance the performance as it is enacted through the virtual body (becoming human). This creates a lopsided relationship due to the collaborative nature of the VFX and animation industries. Regardless, within a realistic-mainstream-narrative-entertainment context, it is common practice to map the source performances onto the digital counterpart as accurately as possible. In such instances, the audience should be able to recognise an indexical source within the digital counterpart and engage with it as a coherent construct.

Identifying a virtual human performer with a flesh-and-blood source-actor may offer the clearest link to indexicality, but the question remains how an audience suspends its disbelief when confronted with a plausible creature that does not exist in the real world (Allison, 2011, p. 325). Within this context, digital creatures create a more obvious overlap between live action and animation. Drawing from "*A Estética do Filme*" by Jacques Aumont, Eliane Gordeeff suggests that live-action fiction movies consist of two non-realities: "the fiction itself and the way it is represented (images, objects and actors)." (2016, p. 281). Animation consists of an additional layer of non-reality: "the fiction of history, the representation of it, and the life that its elements seem to have" (2016, p. 281). British animators John Halas and Joy Batchelor (as cited in Hoffer, 1981) believe that the role of the live-action film is to present physical reality, whereas the animated film presents a metaphysical reality more concerned with what things mean than how they look (Wells, 1998, p. 11).

Animation, in the form of photorealistic creature visual effects, weaves an additional layer of interpretation within the live-action film, emboldening the audience to see the digital representation as not merely a cunning visual trick but as an actual body onto which sites of meaning can be ascribed and understood. The more that a digital creature's physical construction and performance are based on real world references – that is, the more becoming-human it is (or becoming-crocodile, becoming-chimpanzee, etc.) – the more an audience perceives recognisable signs of representation, which bolster a stronger indexical link to all such referents. This moves beyond the traditional view of film's indexical nature being almost exclusively derived from its photographic aspects (Gunning, 2007, p. 29). This view of the index is rooted in philosopher Charles Sanders Peirce's article "Prolegomena to an Apology for Pragmaticism", where he includes the index as one of a triad of signs (icon-index-symbol). For Peirce, the index is a sign existentially linked to a referent "by being really and in its individual existence connected with the individual object" (1906, p. 495). Gunning acknowledges the photograph as a frequently cited example of an index, but he believes that solely linking the index in film to an analogue photographic process has reached the limits of its usefulness when applied to digital media (2007, p. 30-31). Gabriel Giralt, a proponent for the non-indexicality of digital media, believes that computer-generated imagery lacks the same "resonance of reality" that traditional filmmaking is able to resolve (2010, p. 13). For him, digital characters lack the full complexity of flesh-and-blood performers, which reduces them to mere caricatures. Visual effects produce technical solutions to represent a non-existent reality, whereas

> representational realism does not create what does not exist but rather transposes existing reality onto the film plane, like one holding a mirror. Nothing could be more opposite to realism than a realistic computer-generated object that pretends to have an indexical referent (2010, p. 14).

Precisely what Giralt means here by digital characters lacking the "full complexity" of flesh-and-blood performers is unclear, especially in light of several recent visual effects characters and creatures that successfully pass for the real thing with the appearance of full complexity intact (e.g. Junior from "Gemini Man"). Giralt wrote his position in 2010 when the best examples of realistic computer-generated characters included Benjamin Button and the Na'vi, which, while stunning VFX achievements for their time, certainly lacked some of the visual subtleties of their human counterparts. However, more recent examples, especially when modeled from real-world subjects, are produced through an intensive process of high-resolution digital scanning, performance analysis, and physically accurate simulation to avoid any sense of caricaturisation to their flesh-and-blood counterparts. Likewise, I postulate a likelihood that given more frequent exposure to digital counterparts within live-action films, a viewer will more readily accept their ontology as equivalent (or at least near-equivalent) to flesh-and-blood performers. That is, increased familiarity with realistic digital counterparts may have a positive effect when a viewer recognises any moments of uncanniness (subsection 4.1.5) in terms of a digital character's performance, design, or integration into a scene. Such familiarity is enhanced by improvements to the graphical realism of digital counterparts, but this may also have the effect of reducing one's familiarity to earlier characters that no longer appear as "real" by comparison, which may in turn inhibit a consistent sense of empathy with those characters.[17]

Gunning argues that it is not meaningful to think of the digital and the indexical as opposed terms (2004, p. 39). He believes that this confusion arose from the traditional understanding of the photographic index arising from the relationship between a physical object being photographed and the image that appears within the light-sensitive emulsion. By contrast, a digital image is not recorded by the exposure of an emulsion but rather through the encoding of light colour and intensity as data within a matrix of numbers, which can only be resolved as an image through the intercession of a computing device. Markos Hadjioannou describes the analogue photographic process as producing "direct and isomorphic transcriptions", whereas digital photography converts "patterns of light and shade into abstract numbers that are assigned formal relationships on the basis of a preformatted numerical grid." (2008, p. 125). The key point of difference between these two processes is that in the digital photograph, "the image's internal continuum is disrupted by becoming a set of separate, discrete, self-contained units (pixels) that can be described using a mathematical function" (2008, p. 125). Analogue filmmaking also succumbs to a form of discreteness: the grains of silver halide in film emulsion produce a limited definition at which photon information can be exposed and stored, effectively producing the same "discrete, self-contained units" in the form of film grain rather than pixels. The film grains within an emulsion are inseparable and permanently encoded through the exposure of the image. However, while film has a permanent and discrete granularity, the emulsion only records a direct continuity between real-world objects and lighting. This varies from the digital image in that the recording medium, a series of digital bits, is a set of numbers that could just as easily be reordered to create an image of a non-existent, fabricated image (that is,

[17] A further question along this line of thought specifically regarding digital creatures is whether a viewer is more or less likely to empathise with a believable but fictitious creature (e.g. Na'vi, trolls, dragons) than a realistic replica of a living creature. These questions are ripe areas for future research.

discontinuous) as one linked to the real-world (continuous). The concrete materiality of the analogue process produces a claim to indexicality that the digital image, stored across a hard drive as an intangible series of highly mutable bits, does not possess (Hadjiannou, 2008, p. 125). Yet, while the analogue process may claim greater indexicality than the digital process, it is still subject to the direction of human choices, such as in the nature of framing the image, the choice of what is included or excluded (*mise-en-scène*), the exposure, and the style of colour grading. However, once those choices are made, they are permanently inscribed as part of the finished product of the image; this contrasts with the impermanent, variable nature of digital processes (Manovich, 2001, p. 36)

Film facilitates a form of documentation, whereas the computer (and by extension, the digital image) transmutes reality through a process of symbolisation and quantification (Allison, 2011, p. 334). Manovich believes that the digitisation of live-action footage severs its indexical relationship to pro-filmic reality and converts it into a form of animation, which he categorises as a sub-genre of painting (2016, p. 22). He goes on to define digital film as:

> digital film = live action material + painting + image processing + compositing + 2D computer animation + 3D computer animation (2016, p. 28).

Thus, for Manovich, "[d]igital cinema is a particular case of animation which uses live action as one of its many elements." (2016, p. 29). Allison refers to this blend of computer-generated images and profilmic material as "digital indexicality", and she believes this blending "reflects the heterogeneity of digital visual culture, which both draws from the traditions of older media and transforms them with the computer's enhanced capacities for automation and manipulation." (2011, p. 326). As I explore later in this thesis, the heterogeneity of digital visual culture likewise leads to a heterogeneity in the production of performance, whereby the digital character on screen is an amalgamation of the work of multiple artists beyond just the flesh-and-blood actor(s), simultaneously challenging and upholding an indexical link to the source perfomance(s).

For Gunning, the indexicality of the photograph is not linked to the storage medium, but rather in the medium's ability to accurately represent the object photographed. While the way that light information is stored and captured between digital and non-digital cameras is wholly different, both forms of photography are capable of being manipulated, which is why digital images can be submitted for legal documents and evidence in lieu of film-based photographs. Gunning also raises the point that prior to the invention of digital photography, there were plenty of devices that recorded indexical information as data, such as heart rate monitors, thermometers, speedometers, wind gauges, and barometers, among many others (Gunning, 2004, p. 40). As a result, the fact that information becomes digitised does not preclude it from being an index. Digital photography's ability to absolutely transform the originally photographed subject is a major source of anxiety for people who argue against the indexicality of the digital. However, while the digital process undoubtedly facilitates this process, film-based photography is also capable of transformations that can attenuate, ignore, or undo the indexical, such as combinations of retouching (airbrushing), filters and lenses, framing, exposure time, specialised developing chemicals, and overlaying elements through multiple printing (2004, pp. 40-41). As a result, there exist multiple challenges to the notion that the film-based image is the sole province of indexicality within photography.

Hadjioannou argues that the corporeality of digital bodies is impossible to reconcile through indexicality, and that we should instead search for a symbolic system specific to digital imagery that reveals its materiality according to its own set of relations. Allison argues that unlike well-known examples of the index – such as a photograph, a footprint, and a knock on the door – motion capture in its "pure" form (that is, as just the capture of movement data before it is applied to a digital puppet) exists solely as binary data that cannot be understood by the human senses alone, thus forming "a record of movement without movement." (2011, pp. 335-336). It is in the process of matching the movement data to a digital model that reveals the essence of the indexical/corporeal relationship in motion capture. Specific to MoCap, Hadjiannou embraces Deleuze's concept of the "*geste*" (the body as interval) to describe the seamless flow of a human performer's body as it is captured and translated onto a digital performing body (2008, p. 136). He describes the *geste* as

> a category of transgressing boundaries so that corporeality is the very procedure of a body becoming a role – or a digital image – without yet losing the rudimentary elements of physicality that connect that body with a certain space and time as that which s/he exists (2008, p. 135)

Instead of attempting to locate an indexical relationship between the image of the analogue compared to the digital, we may find a more meaningful discussion in what is captured and retained in terms of an actual performing body. Bode underscores this point by positing that "the history of rotoscoping and motion capture show that the index exists in the registering of gesture, not merely in reflected light." (2010b, p. 1). That is, if a source-actor's performance is directly used to drive a digital character's screen performance[18], we more aptly understand the linkage between the human performer and the digital performer by registering the *geste* than searching for the index. If within Hadjiannou's description of the digital image we understand a digital image as an individual pose and the *geste* as the movement between poses, this recalls Lemmon's observation that a recognisable (i.e. *indexical*) performance is produced through the movement between poses rather than the individual poses themselves. By validating performance above image, we continue to celebrate the metaphysical reality that animation brings to live-action.

The *geste* inherently possesses a temporal aspect, which mirrors André Bazin's observation that the key difference between photography and cinema is the recording of duration – or, as he described it, "change mummified" (1960, p. 8). While Bazin's description encompasses both image and temporality, motion capture forgoes the image altogether to produce what we may consider "movement mummified". Balcerzak refers to this as "a ghosting of the actor" (2013, p. 206). He argues that the onscreen image of a digital character does not fully negate the human actor's presence, but rather preserves the actor's "most externalised aspects of performance" (2013, p. 206). Although it is possible for a

[18] In Section 4.4, I distinguish between different modes of vactor performance, including synthespians, digital doubles, and further subcategories. This example specifically refers to synthespians, protean doubles, and digital doppelgängers. Digital fantoccini are excluded on the basis that the input of their motion is keyframe animated rather than being directly driven by a source-actor's performance.

motion capture performance to produce *movement mummified*, I challenge this assertion on the grounds that the MoCap performance is not necessarily *embalmed* (to continue Bazin's analogy). That is, while a film actor's original location, direction, and temporality remain intact as part of the filmmaking process (thereby *mummifying* it), the MoCap performer's movements may be subject to temporal and physical editing, preserving the original movements to greater or lesser extents. Therefore, MoCap does not necessarily result in just *movement mummified* but also provides possibilities for what I consider "movement reconstructed". However, in these situations, the reconstructions are heavily based on the underlying original performance, a type of spectral presence that preserves the original performance intent while resisting the wrappings of mummification.

Despite a process of encoding and translation to a digital body, as well as the ministrations of animators, Balcerzak argues that

> the records of the actor's movements (and moments) suggest the aura of a performing body allowing for a streamlined kinetic 'ghosting' of the performer [resulting in…] a type of spectral kinesis – a movement-based, yet always artistically compromised, possession within a digital image. (2013, p. 206).

This process harkens back to Edward Gordon Craig's concept of "the über-marionette". Writing in 1907, Craig suggests that the development of the theatre requires new modes of performance, the most promising of which he suggests takes the form of an über-marionette. However, the über-marionette has remained a paradox for theatre scholars because Craig coined the term without providing a meaningful description for it (Le Beouf, 2010, p. 103). Perhaps the clearest indication of its meaning comes from a passage in an article written by American actor Harvey Grossman, a draft of which was sent to Craig and received his implicit approval. Grossman describes the über-marionette as not a marionette at all, but rather an actor who is concealed head to toe behind a mask, and who fully conceals his or her physical and emotional self from the performance while exclusively inhabiting the mask and costume (Le Boeuf, 2010, p. 112). It is in the description of the actor who hides his or her whole physical being and personality within a full-body mask that Delbridge sees a direct correspondence between motion capture and the über-marionette. Delbridge's own understanding of the über-marionette is as a type of automaton that behaves like an obedient puppet. Within the context of MoCap, the obedient puppet becomes a character in virtual space driven by an actor who abandons both personality and outward form (2015, pp. 37-39). However, while the MoCap puppet can legitimately be seen as a type of full-body mask that replaces the actor's physiology with a virtual representation, traces of the source-performer's physiology remain – for instance, a characteristic gait, a hunch in the shoulders, or a stiffness in the back. That is, while MoCap may provide better opportunities to approach the goal of the über-marionette, the source performer is unable to fully abandon his physical or emotional self. Outward form may be subverted, but Grossman describes the über-marionette as an actor who "no longer exhibits *himself* upon the Stage [sic]" (Le Boeuf, 2010, p. 112, *original emphasis*); I contend that part of exhibiting oneself includes the physicality of one's movements in addition to the form of one's body.

2.8 Conclusion

Motion capture presents numerous challenges to our understanding of what constitutes acting on screen, including issues of authorship, presence, and indexicality. Who deserves credit for the screen performance of a motion-captured character: the source-actor(s) alone or in combination with the MoCap artists and animators? What happens to the source-actor's aura when his or her performance is encoded as MoCap data and translated onto a digital puppet? What is "real" within the context of screen space and how does this affect our understanding of the ontology of digital performers? This chapter does not presume to reconcile these issues but rather presents the scope of the problems, which will be analysed in further depth in the following chapters. An overview of key animation production considerations is presented to provide the reader with a clear understanding of the imperatives of animation most closely related to this study. Precursors to motion capture are discussed in order to contextualise current technological and production considerations of the MoCap process. Finally, the chapter explores a discussion of the nature of existence of digital characters, and specifically how we understand the indexical link between a digital counterpart and a source-actor. Due to the (at best) tenuous nature of this link, MoCap prompts us to consider an alternative to the image-based index of photograph to subject, whether as the preservation of movement rather than image, or the reconstruction of movement that adheres to a spectral presence.

Now that I've identified fundamental aspects of motion capture, including controversies surrounding its relationship to screen performances, it is important to establish precise definitions and contexts for terminology used throughout the rest of this thesis. The following chapter explores our understanding of what constitutes acting and performance within specific traditions relevant to motion capture. Using this as a basis, Chapter 4 discusses language specific to the actor's experience, extending it to fit a motion capture context.

3
THE LANGUAGE OF ACTING

3.1 Defining Performance

This thesis is concerned with how we understand the nature of performance and acting within the context of motion capture. This chapter explores the most common forms of acting within film and performance capture within a Western context. To understand how our perceptions of *acting* and *performance* are affected by motion capture, we must first agree to meaningful definitions for these and related terms. However, I find that *performance* and *acting* "are slippery terms that obdurately resist fixed definitions." (Kennedy, 2019a, p. 299). It is not the purview of this thesis to provide an exhaustive study of all possible definitions of these terms, but rather to lay a foundation with a set of definitions that clarify and serve its purposes. The performance language established in this and the following chapter is derived from the Method acting tradition and its descendents. This chapter lays a brief history of Western acting specific to the development of the style of emotionalism and considers its application in motion capture performances. This is considered through how we understand the relationship between the creation of a performance from a source-actor and the enactment of that performance through a digital counterpart that fully replaces the source-actor on screen. Likewise, this chapter examines how the technological mediation of a performance affects our understanding of how acting is produced and who is responsible for the acting witnessed in a digital counterpart.

Richard Schechner, one of the founders of the field of Performance Studies during the late 1960s and 1970s, promotes a definition of performance that privileges inclusivity over specificity: performance is "all the activity of a given participant on a given occasion which serves to influence in any way any of the other participants." (2002, p. 29). This definition allows for performance to occur anywhere so long as the performed actions are "done *for* someone, even if that person is the performer him- or herself." (Landay, 2012, p. 130, *original emphasis*). We generally conceive of a performance requiring a performer and an audience, but Schechner's definition allows for the performer to simultaneously be an active and receptive participant. Within the context of MoCap, the virtual avatar both affects and is affected by the performer whose movements are captured and mapped onto the digital character, producing a back-and-forth relationship through performance (as part of the Leib-Körper-Doppelgänger relationship of the actor's experience discussed in Section 5.2).

In Chapter 2, I explored the privileged relationship to "realness" that live-action holds above animation. A live-action film presupposes a recorded connection to some past *now*, whereas animation is often considered to have no tie to a historical past and is instead composed of manufactured graphical images on film or video that are played back in a linear form. According to this view, a performance requires three primary factors: time, space, and physical presence (Heathfield, 2004, p. 7). Performance studies scholar Peggy Phelan argues that performance only ever occurs in the present, making it an un-documentable and unreproducible phenomenon by its nature. For Phelan, recording live performance undermines its very essence and transforms the performance into something else (1996, p. 146). By contrast, Auslander challenges the traditional view of a binary opposition between the live and the recorded (2008, p. 2), indicating that if this binary exists it is due to cultural and historical contingencies rather than intrinsic characteristics (2008, p. 11). For instance, the category of the "live" was introduced as a response to the invention of reproductive technologies, and as new modes of recording are developed, we contrast the technologies based on their relative capacities for conveying liveness (2008, p. 169). Indeed, we colloquially refer to recorded events as performances all the time, celebrating an actor in film or television for a "good performance" while deriding another actor in the same production for a "weak performance". As a result, Auslander challenges Phelan's assertion that the *now* is the exclusive domain of live performance (2008, p. 44). The cultural predominance of television and film changes our experience of the world and our perception of liveness (2008, p. 2) – leading Hosea to comment that "[i]f we can see it happen 'now', it must be real." (2012, p. 95).

Phelan's assertion that disappearance is a unique aspect to live performance is also challenged by the interlaced scanlines of broadcast analogue television and the rapid succession of discrete, still frames in film and television (2012, p. 97). The act of watching a performer on screen can also be seen as a type of performance in itself, whereby the recording engages with the audience by re-enacting its contents. However, while the contents are repeatable (and therefore not "live" by Phelan's reckoning), the performance actually occurs in the relationship between the film and the audience – that is, in every aspect of how the film is transmitted and responded to in a particular location and time by particular individuals. The act of watching a given recorded performance is an event that is itself unrepeatable. And just as a live performance is subject to a viewer's perception and retention in memory, so too is the act of watching a recorded performance (Auslander, 2008, p. 159). This is the reason why fans of franchises such as "Star Wars" may fondly recollect the experience of watching the films in the cinema, which produced a type of performance that could never be replicated through repeated viewings at home.

While we may be willing to accept recorded actors as engaging in a type of performance, are we willing to extend the concept of performance to animated characters? And if not, how are we to conceive of the actions and emotions animated characters express while participating in a story? In our conception of what constitutes performance, is there an ontological distinction between a record of living human action versus a graphical construction of illusory movement and intention? If the animated human is visually indistinguishable from a flesh-and-blood actor, does this distinction cease to matter? In an attempt to reconcile whether an animated character can be considered to perform, Hosea proposes a meta-theory of performance

> in which a human body (or a substitute for it) assumes an identity beyond that of her everyday life and is displayed for an audience where she represents behaviours, which were planned and prepared. This is a time-based process, which takes place in the present and can be applied to theatre, live art and other performing arts [… including] animation. (2012, p. 32).

Hosea's distinction that a substitute for a human body can be the enactor of performance is critical. In this way we can consider how an animated character, regardless of visual style, can produce performance. It is important to note that Hosea's working definition doesn't include all possible animated "bodies" as potential performers, but rather only animated bodies that could function as a substitute for human bodies. The line, dot, and squiggle characters of "The Dot and the Line" are purely graphical without any semblance to human physiology or performance, and therefore do not count based on this definition. On the other hand, the movements and expressions of characters in feature-length 2D and 3D animated films are often based, at least in part, on reference performances that animators perform either in front of mirrors or record of themselves[19] (Kennedy, 2017, p. 302). For this reason, Hosea regards animation as a performative act that "enacts that which it signifies": as animators produce their own reference, they become animated by the character they are creating (2012, p. 163). Even if an animated character is not based on human reference, the animator can still be seen to perform *through* the character and therefore the animated body functions as a substitute for a human body. Motion capture provides a more immediate sense of body substitution: a human actor produces a live source performance, which is then digitised and retargeted onto an animated character. This definition is not reliant on live-performance as the enactors or living bodies as the enacted, and allows for characters from Bugs Bunny to Thanos to be viewed as performers.

If we decide to venture one step further and remove the requirement that performance is tied to a human (or substitute) body altogether, we could consider as performance the purely graphical elements of "The Dot and the Line". In this case, there is no relationship to live performance either directly recorded or by proxy. While animation is traditionally considered to not be "live", an animated character as an experienced phenomenon only occurs in the present moment. With the possible exception of MoCap, animation provides no record to a character's past because it has no past, and only exists in the now at the moment of projection. As soon as the projection stops, the animated character ceases to

[19] Animators also often work from videos of voice actors performing their lines in sound booths. However, in order to work with the microphone, voice actors tend to be limited by what motions they can produce with their bodies. Voice actors often provide a good reference for facial performance while animators often have to work out and record their own body performance reference for a character.

exist. As Hosea indicates, "this makes animation, like the theatre, also an 'art of the present'." (2012, p. 98).

3.2 Defining Acting

The spread of digital technologies has given rise to a diverse range of digital performance modes, which are well documented by Steve Dixon in his book "Digital Performance: A History of New Media in Theatre, Dance, Performance Art, and Installation" (2007). Dixon is considered a leading authority on the relationship between digital technologies and performing arts (LASALLE College of the Arts, 2016, para. 12), and he defines *digital performance* as "all performance works where computer technologies play a key role rather than a subsidiary one in content, techniques, aesthetics, or delivery forms." (Dixon, 2007, p. 3). For the purpose of this thesis, I am concerned with the practice of acting within digital performance, and specifically what constitutes acting within a MoCap context.

Due to the personal creative nature of acting, definitions of the craft vary between scholars and practitioners, as well as between practitioners themselves. Directors and acting coaches often discuss acting using concepts such as "being", "reaction", "imagination", "empathy", "pretending", "playing", and "mimesis" (Kennedy, 2019a, p. 299). The term *actor* is rooted in the Greek word *hypokrites* (sharing an origin with "hypocrite"), which refers to a stage actor or a pretender (Merriam-Webster, 2021, para. 2). The role of an actor today is paradoxically both obvious and mysterious. The actor strives to represent truth while committing a deceptive act of representation. The actor is simultaneously present and absent, participating in a double performance of revealing a character while attempting to hide any semblance to his or her personal identity (Worthen, 1984, p. 4). However, while flesh-and-blood actors often attempt to deflect reference to themselves, the animated character, which has no existence outside the film medium, can never pretend to be anything other than an animated character (Kennedy, in press).

Ed Hooks, a veteran actor who has made a career out of teaching acting to animators, believes the role of the actor is to "[behave] believably in pretend circumstances for a theatrical purpose." (2017, p. 27). He instructs that a theatrical purpose involves an action, conflict, and objective (2017, pp. 118-119). Hooks admonishes the actor (or animator) who expresses emotions without engaging with a theatrical purpose as incapable of producing good acting (2011, pp. 19-20). Thus, for Hooks, the process of representing truth is dependent on more than just the actor expressing emotions we believe are real (or truthful), but also on producing believable circumstances in which those emotions are enacted. However, what exactly is meant by the notion of "truth" is not consistently defined within a Method acting tradition.[20]

Bloch, Orthous, and Santibáñez-H's use the term *acting behaviour* to describe an actor's conscious effort to convey thoughts and feelings to an audience through voice, gesture, and posture in order to carry out an artistic motivation (1987, p. 1). Chris Pallant provides a more specific variation of this idea[21] through the term *actorly performance*, which he describes as "a physical performance, informed by narrative, that results in a nonsynthetic action and/or expression" (2012, p. 38). He explains that the performance may become mediated in some way after the recording event, but that this does not detract from its status as actorly performance. Importantly, this term does not "privilege filmic or televisual actorly performance over that which is delivered for any other screen media", such as actorly performances captured as PeCap data and retargeted onto digital counterparts within video games, television, and film (2012, p. 38).

For Michael Kirby, a drama theorist and colleague of Schechner, "[a]cting means to feign, to simulate, to represent, to impersonate." (1972, p. 3). He loosely defines acting as

> something that is done by a performer rather than something that is done for or to him. […] If the performer does something to simulate, represent, impersonate and so forth, he is acting. It does not matter what style he uses […]. No emotion needs

---

[20] This concept is defined and explored in greater detail in Section 4.1.

[21] It is important to point out that the literature does not suggest a link between the concepts of Bloch et al. and Pallant, and any similarity between these concepts is likely due to coincidence.

> to be involved. The definition can depend solely on the character of what is done. […] Acting can be said to exist in the smallest and simplest action that involves pretense. (1972, p. 6).

Although acting in Western film and television is embedded in depictions of realism, acting is composed from a variety of styles that are each informed by and dependent upon the socio-historical contexts in which they exist. As a result, a meaningful definition of acting must be able to adapt with time and embrace diversity (Kennedy, in press). Eli Rozik provides a theoretical understanding of what acting is, describing it as "inscribing on matter a description of an entity in a world and deflection of reference to it" (2002, p. 110). Citing Fiebach (2002), acting theorist Barry King writes that "the human actor is central to the process of dramatic signification" within acting, and especially within theatre (2011, p. 248). However, he points out that the formulation of the actor as the indispensable element of signification privileges the actor over the contributions of others involved in a complex collaboration process. Not all actors contribute the same amount to dramatic signification, such as the differences between a star actor and a walk-on extra who has no lines. King explains that even if "the term *actor* is being used as a metonym for the collective contribution of human agency, the question of non-human performance and agency still remains." (2011, p. 248, original emphasis). This is relevant when we consider that while the actor's body has traditionally been considered both the instrument and medium of acting (Kemp, 2012, p. 3), within MoCap the body's function as an acting medium is curtailed or eliminated. This impacts on how we understand acting is produced in a motion capture context, especially in terms of the roles that MoCap artists and animators play in the production in the performance of a MoCap-led virtual body on screen.

To articulate how actors produce different amounts of dramatic signification, Kirby describes a continuum of acting, from not-acting through to complex acting (Figure 5), that is concerned with quantifying degrees of performance rather than identifying particular styles (1972, p. 3). Simply put, the continuum describes the degree to which an actor assumes a character as opposed to playing him/herself. On the left end of the continuum, actors receive references of acting, as opposed to producing those references themselves (1972,

NOT-ACTING **ACTING**

Non-matrixed Performing	Non-matrixed Representation	"Received" Acting	Simple Acting	Complex Acting

Figure 5. Kirby's continuum of acting. Reprinted from *On Acting and Not-Acting* (p. 8), by M. Kirby, 1972, *The Drama Review: TDR*, 16(1).

p. 5). For instance, non-matrixed performing has the least capacity for acting and occurs when the performer "is merely himself and is not imbedded […] in matrices of pretended or represented character, situation, place and time" (1972, p. 4). Toward the right end of the continuum the capacity for producing references and representing character increases (1972, p. 6). "Received acting" occurs when we understand performers' roles in a story more by what they wear and how they're contextualised rather than by anything they actually do. Kirby elaborates: "Extras, who do nothing but walk and stand in costume, are seen as 'actors.' Anyone merely walking across a stage containing a realistic setting might come to represent a person in that place – and, perhaps, time – without doing anything we could distinguish as acting." (1972, p. 5)

Of concern to the experiments in this research are simple and complex amounts of acting. Kirby describes simple acting as "that in which only one element or dimension of acting is used. [… Emotion] may be the only area in which pretense takes place […] or only an action such as putting on a jacket may be simulated." (1972, p. 8). By contrast, complex acting is multi-dimensional and simultaneously engages in more than one area of pretense (1972, p. 9). Hosea indicates that complex acting is associated with "lifelikeness and authenticity in which the actors lose their own egos and become the characters that they portray." (2012, p. 168). She argues that acting is rarely authentic or spontaneous, and instead "draws upon a series of learned, planned and twice-behaved actions" (2012, p. 168). Schechner couches *restored behaviour* in personal terms as "'me behaving as if I were someone else,' or 'as I am told to do,' or 'as I have learned.'" (2006, p. 34). Since restored behaviour is never performed for the first time, it becomes *twice-behaved behaviour* (2006, p. 36), and can "[refer] to actions or speech acts that are prepared or rehearsed and then re-presented." (Hosea, 2012, p. 25). The process of creating animation leads to a specific type of twice-behaved behaviours, whereby an animated character's performance is (re) constructed through drawing from live reference, creating planning sketches, and refining drawings and poses over time. Unlike in stage or film acting, the twice-behaved behaviour in animation is never produced in real-time. This is complicated when we embrace a definition of animation that includes motion capture, as an actor performing twice-behaved behaviours in real-time through the MoCap equipment appears to also produce them live through animation. While this is true to an extent, what matters is whether the motion-captured performance is later edited to produce the final animation, as is most often the case. The less common instances of this include live projections as part of theatre or dance in which a director may believe that a given aesthetic choice is complimented by motion-captured actions that are not fully resolved.

3.2.1 Acting Before Stanislavski

The *realist* style of acting popular in North America and Europe is grounded in the concept of *believability*, "privileg[ing] mimesis above other forms of representation." (Kennedy, in press). Hosea indicates that such styles "[aim] for lifelikeness and authenticity in which the actors lose their own egos and become the character that they portray" (2012, p. 168). As explored in Chapter 2, the "realness" in live-action is derived from the assumed linkage between performance and flesh-and-blood performers. However, the audience must adjust its perceptual expectations in order to embrace realism in the context of character animation because animated characters have no fleshy basis: "[i]nstead, the audience must extend its suspension of disbelief in order to accept as authentic the behaviours and situations acted through animated bodies, despite those bodies having no physical connection in the real world." (Kennedy, in press). This type of perceptual leap is deeply precedented in pre-Socratic traditions of animism (Hammond 1895, p. 394), whereby the philosophy of hylozoism imputes life and intention to all forms of matter (Goldstein 2003, p. 295). Film historian Donald Crafton describes how audiences can perceive human intentions and mental states in nonhuman animals, and extends "filmed animals [as] analogues for human agency" to cover "animated characters [performing] as figurations of film actors endowed with their own agency." (2013, p. 61). Indeed, one standard definition of *animate* is to invoke life, as in the story of Prometheus who animated humankind from clay.

Rhonda Blair, whose research combines the fields of acting with cognitive neuroscience, details how Joseph Roach's "The Player's Passion" reveals paradigms of acting to be directly influenced by paradigms of science, and suggests that as science changes, so too does acting. While some aspects of what constitutes acting are relatively stable throughout time, many aspects of acting are highly susceptible to shifting cultural mores (Blair, 2008, p. 23). One of the most significant philosophical models to influence scientific and ethical appraisals of the self is a form of mind-body dualism developed during the seventeenth century by René Descartes. Known as Cartesian dualism, this theory connects the mind with consciousness and self-awareness, while understanding the mind to be separate from the material substance of the body (Kemp, 2012, p. 2). The origins of mind-body dualism extend as far back as Plato, but Descartes' particular formulation, whereby mind is split from matter and reason from feeling, proved especially significant in theories of acting (as well as the ethical considerations of acting) in the West throughout the seventeenth- and eighteenth centuries (Blair, 2008, p. 5). Moving from a Hume-inspired morality derived from feelings to a Kant-inspired morality based on reason (Wilson & Denis, 2018), popular ethical discourse of the time promoted the rational self as the paragon of virtue. Within Cartesian dualism, rationality is associated with the mind, whereas feelings (emotions) are associated with the body. Therefore, the mind is virtuous and trustworthy, whereas the body and all of its emotions are suspect and dangerous. For actors, this shifting sentiment produced two problems. First, actors utilised their bodies as their "instruments", through which they experienced (or represented) their morally suspect baser instincts, such as sexuality, passions, and emotions. Second, actors participated in an irrational act of pretending to be someone else while the audience was implicated in this falsehood through their suspension of disbelief. The seventeenth century political philosopher Thomas Hobbes writes that

> [a person] is he whose words or actions are considered, either as his own, or as representing the words or actions of another man […] whether truly or by fiction. When they are considered as his own, then is he called a natural person: and when they are considered as representing the words and actions of another, then is he a

feigned or artificial person. (2011, p. 101).

Hobbes outlines a clear distinction between someone who participates in sustained mimicry as being a feigned person, and someone who scrupulously adheres to maintaining one's own self-manifestation. The former is associated with pretence, deceit, and hypocrisy, whereas the latter is considered virtuous (Barish, 1969, p. 1). Blair describes how an actor navigates this falsehood on stage by simultaneously being and not being there:

> He is present as an actor, strutting his stagey stuff; but he is also absent, negated by the dramatic illusion he creates. In performance, the actor is engaged in two performances, a "double effort" that reveals him as an actor while it conceals him within his dramatic role. (2008, p. 4).

The actor's doubleness serves as a metaphor for the daily ambiguity of choosing to behave in ways that are either consistent or inconsistent with one's internal sense of self, which may threaten to "transform the self into an uncomfortable fiction." (2008. p. 6). Emotions were considered troublesome due to their "irrational" basis – a person could be susceptible to emotions that led to actions that were in poor judgment or even harmful. For this reason, rational pursuits were to be celebrated while diversions that appealed to the passions were spurned.

How an actor reliably navigated a range of emotions on stage became of critical importance to French philosopher Denis Diderot, who believed that emotion was separate from physical activity. In "*Paradoxe sur le Comédien*" ("The Actor's Paradox", written in 1773, but not published until 1830), Diderot expresses that the actor's role is not to experience emotions on stage, but rather to produce the appearance of being affected by emotion through conscious muscular control of expression (Kemp, 2012, p. 2). Diderot views emotions as a disruptive interference in daily life and that a person could only clearly express one's self when in a relaxed and unemotional state. In Diderot's estimation, by feeling nothing at all, a good actor triggers the strongest emotional reaction in an audience. Diderot also views a physical technique as superior to an emotional technique[22] because (in his estimation)

[22] It is important to note that emotionalism and anti-emotionalism were not the only considerations for

an emotional actor could never reliably repeat his or her performance with the same quality or success (Konijn, 2000, p. 22). Fearing that actors would become overwhelmed by emotional performances, Diderot writes: “If he were to be himself while acting, how would he then be able to stop?” (as cited in Konijn, 2000, p. 22). For Diderot, actors need to constantly monitor and adjust the physical features of their performances in order to maximise the emotional reception from the audience: “all [the actor’s] talent consists not in feeling, as you suppose, but in rendering the signs of the emotion so rigorously that you are taken in” (as cited in Bloch, 1993, p. 121).

Diderot’s anti-emotionalism sits in direct opposition to the emotionalism epitomised by the American “Method” school of acting developed during the mid-twentieth century, which asserts the pre-eminence of the actor’s authentic emotional experience (1993, p. 121). These two approaches sit at opposite ends of a continuum that describes how personally involved an actor is with feeling the emotions he or she physically expresses (Figure 6). Between these poles are a variety of techniques that prioritise a balance amid physical control and the creation of genuine emotion, including the Delsarte method (discussed below) and the Alba emoting technique (discussed in subsection 3.2.4). Animators are often advised to record and work from their own acting reference in the creation of their animated character performances. However, most animators lack any meaningful acting training or experience. As a result, their reference performances range the gamut between severely underplayed (non-emotional) to stereotypically histrionic (emotionally overacted), as well as an over-reliance on reproducing the canon of codified animation exercises from key textbooks such as “The Animator’s Survival Kit” (Williams, 2002) and “Timing for Animation” (Whittaker & Halas, 2002). In my experience of both teaching animation and watching the reference performances produced by professionals in the field, animators rarely appear to feel the emotions they portray. Instead, these animators do their best to “put on”, or mimic, an

---

the development of a new form of acting. Changes in the size and layout of new theatres led to alternative methods of performance. For instance, Delsarte’s codification of broad and iconic body language to physically represent emotion read well in larger theatres with gas lighting, poor sightlines, and unreliable accoustics. However, the development of the Little Theatre Movement (started 1912) and Chamber Theatre (1947) provided more intimate venues where naturalistic acting played better to an audience that witnessed the performances much closer up.

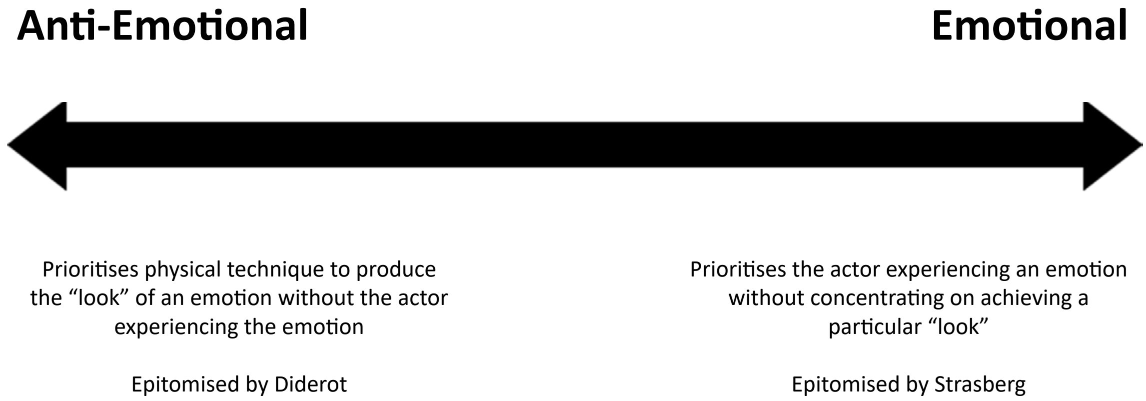


Figure 6. Continuum of emotional involvement on behalf of the actor, from Diderot's anti-emotionalism in "*Paradoxe sur le Comédien*" to the emotionalism of "Method" acting as described in "Strasberg at the Actor's Studio" (Hethmon, 1965).

emotion based on preconceived notions of how particular moments should look or play out. This draws into question whether emotionally affective character animation is dependent on an animator producing reference performances in which he or she actually experiences the emotions in a given sequence. Alternatively, can emotionally affective character animation be derived from physical technique alone, in which the animator does not experience genuine emotions during his or her reference performances? In this case, the animator works from the put-on emotions to develop caricatures of authentic expressions. Since exaggeration is a key principle within animation, it could be argued that even when working from reference of genuine emotional expression, the animated result is still exaggerated. The question is whether the degree of emotional affect, as well as the overall finished performance, is any different when working from reference that is grounded in the experience of genuine emotional feeling versus reference that relies on embodied acting techniques to produce a given expression. While this question is directed more toward keyframe character animation, it still holds relevance to animators who are responsible for modifying motion capture performances. This is underexplored territory within academic animation literature (Kennedy, 2014; Kennedy, 2015a; Kennedy, 2015b) and therefore a ripe area for future research.

A nineteenth century descendent to Diderot's method was a system of expressivity developed by Francois Delsarte. Delsartism promoted that "all inner emotional states could be codified and expressed by the correct combination of mechanical gesture, stance, and facial

Figure 7. Illustration of a Delsarte-inspired pose by Rose Sprague, characteristic of "teapot-style" acting. Reprinted from *An Hour with Delsarte: a Study of Expression*, by A. Morgan, 1889.

configuration[s]." (Bode, 2008, p. 5). This declamatory style of performance became popular across Europe and the United States and was used to train singers, composers, and public speakers, in addition to actors. The Delsarte method was detailed in a reference manual full of illustrations that link poses and gestures to particular emotions, providing would-be actors and theatre-going audiences with a library of consistent and recognisable expressions and body configurations. The catalogue nature of Delsarte's system later influenced industrialised animation studios, which developed model sheets of characteristic poses and facial expressions in order to maintain performance consistency across multiple animators working on the same character (Crafton, 2013, p. 29). Today, Delsarte's system is often derided as "teapot acting" (Figure 7) due to the rigidity of its codified actions, but it serves as a historical precedent to more recent attempts to produce taxonomies of expression, such as the Facial Action Coding System (FACS) (discussed in subsection 4.3.2) and manner of breathing (discussed in Section 4.2).

3.2.2 Figurative Acting and Gesture

A hallmark of cartoon acting is an exaggerated form of performance that is only achievable by means of a graphic medium. This style of acting is described by different names such as *figurative acting* (Crafton) and *cartoonal acting* (Hosea), and is epitomised in the Looney Toons cartoon films of the 1940s and 50s, especially under the direction of Chuck Jones.[23] The acting in these films is characterised by an emphasis on dynamic movement that produces easily legible signifying gestures, resulting in broad humour and slapstick comedy heavily influenced by Vaudeville and burlesque performances. What is gained through this style is an emphasis on clarity of broad intention and emotional response. As liberating as the graphical medium is, it provides its own particular limitations, as well. The graphical line can convey only so much detail, so gestures and expressions that lend themselves to this style of drawing are prioritised while subtler details may be avoided (Crafton, 2013, p. 23). As a subject for future research, I would hypothesise that in terms of graphical animation, emotion is better conveyed through individual strong key poses whereas in realisitic motion, emotion is better conveyed in the motion between poses.

Flesh-and-blood actors frequently rely on an embodied approach to acting, looking inward to develop a complex character; by comparison, figurative acting is completely extroverted. When lacking a clear development of nuanced, emotional personalities and expressions, animated characters often play recognisable "types" by "marshalling a small range of instantly identifiable movements and characteristic gags in film after film." (Crafton, 2013, p. 23). These types of performances are pure histrionics, with characters frequently "breaking the fourth wall" by addressing the audience or camera. Figurative actions are not always preceded by clear or consistent motivations, but rather whichever series of gags is most likely to elicit a laugh from the audience.

[23] Hosea identifies how cartoonal performance entered live-action acting approaches through the work of Tim Burton, Frank Tashlin, and Oleg Mironchnikov.

This is not to suggest that figurative acting is *ad-hoc* and lacking a system. On the contrary, figurative acting relies heavily on standardised character model sheets developed from the Delsarte tradition, which provide a blueprint of poses and facial expressions to guide all animators working on the same character. And just as theatregoers became literate in the somatic signs of actors performing using the Delsarte method, animators adapted this system to move a graphical character from pose to pose, often relying on nonverbal techniques to express emotions. In doing so, a character conveys a conventional range of recognisable gestures and expressions that could be reliably taught to future animators and audiences (Crafton, 2013, p. 29). Donald Graham, the Chouinard Institute art professor responsible for teaching drawing to the animators at the Walt Disney Studio during the 1930s, indicates that "each [gesture] must be clearly staged or presented to the audience. If a gesture, no matter how subtle, is not grasped by the audience the whole scene and conceivably the whole picture may be lost." (cited in Crafton, 2013, p. 43). Where the simplicity of a facial design limits its expressive potential, the character's body language needs to make up for it. (2013, p. 46). However, in more realistic animation, a character's face is more detailed and therefore capable of greater expression, thus requiring less need to fall back on full-body exaggeration.

3.2.3 Stanislavski's Influence

The Russian actor and director Konstantin Stanislavski had the greatest impact on American acting out of any other figure (Blair, 2008, p. 28). Stanislavski was chiefly responsible for developing a realistic style of acting in which the performed behaviour was based on that in ordinary life. This style of acting predominated the stage and screen throughout the twentieth century and remains the principle performance style in Western films and television today, regardless of genre. Schechner writes that "[r]ealistic acting assumes that the emotions of the character are like those of 'real people' even if the characters are dancing across treetops or living in a distant galaxy." (2002, p. 149). Characters acted in the realist style speak regular language in recognisable ways and experience the same emotions in ways familiar to ordinary human experience. Schechner notes that realist acting, unlike ballet or

Japanese Noh drama, requires no special knowledge of theatrical codes in order to decipher what the characters are thinking or doing (2002, p. 149). A key reason why this thesis focuses on realistic acting as opposed to other acting styles is because of the predominance of realistic acting within performance capture specifically.

In 1906, Stanislavski began to lay the foundation of his system while engaged with the Moscow Art Theatre alongside his creative collaborator Vladimir Nemirovich-Danchenko. Both men were appalled by the state of the theatre of the day – acting was full of artificiality, with little rehearsal time and poor effort put into creating meaningful scene design and costumes. Stanislavski especially took issue with the prevailing styles of false feelings (Diderot), declamation (Delsarte), and overacting. Anton Chekhov's play "The Seagull" aptly summarises this sentiment, with the playwright protagonist, Konstantin Treplev, decrying that the theatre needs new forms or there should be no theatre at all. Likewise, Stanislavski believed in the necessity of new forms, both in terms of the theatre and style of acting, in order to meaningfully produce the new Russian dramas of the day, including those by Chekhov (Carnicke, 2009, p. 28).

Stanislavski continued to develop his system over the next three decades. The lessons learnt by the Moscow Art Theatre were introduced to theatregoers in New York during the company's first American tour in 1923 (Carnicke, 2009, p. 20). The Stanislavski system posited a number of performance considerations that challenged the approaches to acting popular in the United States at the time, including:

- An actor who doesn't engage with emotions produces performances that feel artificial;[24]
- An actor should consider his or her performance in terms of objectives, obstacles,

---

[24] It is worth noting that while the concept of *realism* in Western theatre often refers to a type of performance that reflects everyday life and emotions in great detail (i.e. *emotional realism* or *aesthetic realism*), the German playwright Bertolt Brecht developed an understanding of realism based on how and why people behave according to their respective social relations (a type of *philosophical realism*). For Brecht, the theatre on its own was already realistic and a focus on emotional realism was merely fake and manipulative (Barnett, 2021).

and actions;

- An actor should work from his or her previous emotional experiences rather than attempting to mimic the emotions of others;
- An actor should work from within to get in touch with the point of view of the character he or she is to enact, to make the performance feel more *natural* without overacting;
- An actor should be adaptive and reactive to his or her environment and co-stars (da Silva, dos Santos, & Iurgel, 2011, p. 146)

Today, actors and directors still quibble over which approaches to acting produce the most *realistic* results, partly due to the fact that the concept of what we consider to be "natural" acting changes as social life changes. Schechner notes that "[w]hat was considered 'natural' in 1902 is not natural in 2002; and 2002 behavior will look unnatural when regarded by those living in 2102." (2002, p. 149). This is evident in Classical Hollywood cinema from the 1930s-1940s, in which the realist style of acting feels stilted and stylised by today's standards of what is considered natural behaviour (2002, p. 149). It is important to reiterate that realistic acting is a style and not "life itself", and it is suited especially well to film where close-ups and microphoning add extra degrees of intimacy. While realistic acting works well for realist and naturalist dramas,[25] it is less successful in poetic dramas such as those by Shakespeare (2002, p. 152).

While acting systems in the late 1800s prescribed rules for specific ways to perform in different situations (such as the Delsarte system), Lee Strasberg describes Stanislavski's method as the opposite of a system. Instead of setting rules to be followed, he believes it demonstrates a path for the actor to follow and how the actor can repeatedly find what he

---

[25] As a further example of the ambiguous language within acting, the meanings of *real* and *natural* vary depending on context. For instance, within the context of the realist style of acting, the two terms are generally used interchangeably without problem. However, within a larger dramatic context, the two terms are distinct. A realist drama involves characters behaving in believable ways as described in this section. By contrast, a naturalist drama is a heightened form of realism where stage time is equivalent to real time and no jumps in time or place between scenes is allowed. Naturalism often involves characters that are shaped by their circumstances and victims to their social or economic environments (Cash, 2020).

or she needs along that path (Hethmon, 1965, p. 41). Stanislavski encourages the actor to embrace both internal and external techniques – producing psychological and emotional groundwork while perfecting the physical apparatus of the body. In his later years he emphasises the importance of the external techniques more than the internal, advocating for the use of a type of physical sense memory. This theory rests on the idea that by connecting to even the smallest physical truth in an action, the actor's emotions sincerely respond to what the body is doing. Disagreement between American actors and Stanislavski regarding the importance of mastering a physical rhythm is what ultimately led to the development of the Method style of acting, with practitioners like Strasberg championing the pre-eminence of working internally from psychological memory (Balcerzak, 2013, pp. 198-199).

The Polish actor Richard Boleslavsky, who studied directly with Stanislavski in Russia, was the first teacher of the Stanislavski technique in the United States starting in 1923. A number of the great American acting trainers who came later were Boleslavsky's students, including Lee Strasberg and Stella Adler. Strasberg founded the Group Theatre in 1931, of which Sanford Meisner was a member. Stanislavski, Meisner, Adler, and Strasberg "collectively defined the primary terrain of American acting for the latter half of the twentieth century", and they all pursued the same goal for the actor "to 'live truthfully,' emotionally and physically, in imaginary circumstances." (Blair, 2008, p. 39).

Meisner's teachings prioritise what is happening now between one actor responding to another in the moment. Each actor's attention is to be directed toward his or her scene partner with an emphasis on reacting spontaneously to the dynamic between the actors, rather than being distracted from the moment through self-consciousness (Goldstein & Levy, 2017, p. 149). Meisner is famously associated with the mantras of "get out of your head" and "don't think – do", which give priority to instinctive reaction above thinking and planning (Blair, 2008, p. 43). However, Meisner's system of separating thought and action repeat an erroneous and problematic mind-body dualism, which Blair addresses through a cognitive science approach to acting (subsection 3.2.4).

Adler's approach is similar to Meisner but emphasises script analysis, given circumstances, and imagination to a far greater degree. All character actions are to be justified from a thorough analysis of the script, both in terms of how they are to be performed, why they are performed, and what the character desires to accomplish by doing them (Goldstein & Levy, 2017, pp. 148-149). Adler's approach focuses on sociological dimensions of performance, including how a character (and actor) engages with his or her environment through imagination and mental representations of the fictional space (Blair, 2008, p. 44) based on real-life observations (da Silva, dos Santos, & Iurgel, 2011, p. 147).

Strasberg, a co-founder of the famed "Actor's Studio", is principally known for the development of "The Method" style of acting. Similar to Stanislavski, Method acting relies on an actor participating in sense-memory exercises, in which an actor recalls objects and the sensory experiences associated with those objects. However, Method acting also engages with emotion-memory (or affective-memory) to recall emotional experiences from the actor's past. The actor is to build a character using a combination of sense- and emotion-memory (da Silva, dos Santos, & Iurgel, 2011, pp. 146-147). This approach to acting proves especially powerful on screen, with Marlon Brando ushering in the era of Method acting in "On the Waterfront" (Kazan, 1954) famously lamenting "I could have been a contender, I could have been somebody." (Pallant, 2012, p. 37). Strasberg's teachings were the first to provide the bridge between acting on stage and screen. Despite the popularity of Method acting, Strasberg's approach has been criticised for being emotionally indulgent and requiring narcissism from actors. Meisner and Adler reject Strasberg due to his emphasis on emotion memory, which is often misunderstood as an actor trying to remember a feeling or recapture an affect. Blair writes that instead, emotion memory requires the actor to "[remember] a situation in all of its vivid, sensory details, which evokes a rich nexus of images that then facilitates a feeling response; this builds directly on the actor's earlier practice in concentrating on physical and then internal objects." (2008, p. 44). To clarify the core differences between the acting techniques related to the development of the Method, Strasberg was principly concerned with psychological aspects of acting, Adler with the sociological aspects, and Meisner with the behavioural aspects (Krasner, 2010, p. 144).

3.2.4 Cognitive Science and Acting

Emotions and realist acting are intricately entwined, but what is often unclear to novice actors is how to reliably access emotions, especially emotions that may be difficult for a particular individual. In my experience of acting classes in New Zealand, male students often struggle with accessing the feeling of hurt, while female students tend to have difficulty producing anger. Bloch et al. indicate that drama schools are often good at teaching literary and physical aspects of acting behaviour, but that psychophysiological (emotional-expressive) aspects are usually dependent on a student's intuition, life experience, and imagination (1987, p. 2). The lack of technical support for accessing emotions is especially problematic for students who may be deficient in these areas. Likewise, for animators who may have no acting ambitions or experience, there is a dearth of techniques to safely and reliably access emotional terrain when producing animation reference performances. Artaud writes that an actor must become an "athlete affectif" – an athlete of the heart (1987, p. 2) – and like any athlete, good technique is necessary.

In the early 1970s, professors of neurophysiology Susana Bloch and Guy Santibáñez, and professor of dramatic art Pedro Orthous, collaborated through the University of Chile's departments of Physiology, Psychology, and Drama to apply cognitive science to acting by means of emotional effector patterns. They define an emotional effector pattern as "*a unique interdependence between a specific breathing rhythm, a particular expressive attitude* (both facial and postural) *and a given subjective experience.*" (1993, p. 125, original emphasis). They united the discoveries of emotional effector patterns into a coherent physical emotion regulation method (Bloch, 1993, p. 123) and named this the BOS Method (using the letters from their last names), which they later rebranded as the Alba[26] Technique (or Alba Emoting) (1993, p. 135). This technique provides the actor with exercises specific to creating voluntary control of the parts of the body involved in emotional and verbal

[26] Bloch et al. explain that "*alba*" is a Spanish word that means both "dawn" and "white", which was significant to the authors at the time due to a production of "The House of Bernarda Alba" they had been working on shortly before (1993, p. 135).

behaviour, techniques to control mental and emotional states (such as stress or anxiety) that may interfere with such exercises, and teaches the actor how to "reproduce at will the respiratory, postural and facial configurations which correspond to 'real-life' emotional behaviours." (Bloch et al., 1987, pp. 2-3). This goes beyond the assumption that specific facial configurations alone produce sincere emotional conveyance[27] and instead adds extra secondary components (i.e. breathing patterns and postural configurations) that may not be obvious at first glance, but which are critical at a subconscious level to reading an emotion.

In the realist acting style, actors are instructed to personally experience a given emotion, but acting out a scene often requires that an actor moves from one emotion to another, or shifts the intensity of a given emotion. While an actor may be good at producing the phenomenological experience of an emotion, he or she may struggle with Artaud's *emotional athleticism* to proficiently and nimbly move from one emotion to another in order to match the timing of the dramatic situation (Bloch, 1993, p. 122). Bloch indicates that the Alba Technique is based on scientific experimentation under laboratory conditions and helps an actor to access and control an emotion at will. This is achieved through the management and monitoring of an actor's breathing and expression patterns, which she suggests contributes to the actor's subjective experience of a given emotion (1993, p. 121). Konijn describes this in a similar way according to peripheral feedback theory, which states that an emotion can be aroused due to the imitation of the physical signs associated with that emotion. These physical signs are relayed to the brain as physical reactions, and the brain responds by producing the experience of the emotion specific to those physical patterns (2000, p. 98). That is, for each emotion, there is a unique emotional effector pattern that produces "*specific [patterns] of breathing, facial expression, degree of muscular tension, and postural attitudes*", among which the respiratory component is the most vital element (Bloch, 1993, p. 124, original emphasis). Recalling Artaud once more, Bloch cites the following quote from "*Le Théâtre et son double*" (1964) as a conceptual background to

[27] See the discussion of the Facial Action Coding System (FACS) in subsection 4.3.2.

her experimental demonstrations: "Breath accompanies the feeling, and one can penetrate into the feeling by the breath, provided one has been able to discriminate within the breaths *which one corresponds to what feeling*" (1993, p. 133, *Bloch's translation and emphasis*). While Artaud was arguing for a heightened and ecstatic form of performance completely separate from the style of realism, Bloch believes the Alba Technique helps to answer Artaud's inquiries into which manner of breathing corresponds to which emotion (1993, p. 133).

A significant result of Bloch et al.'s research is that by repeatedly initiating and interrupting an effector pattern, the sensory feedback phenomenon is avoided. That is, the actor is able to retain all of the expressive component of an emotion without subjective involvement. As I explore in Section 4.1, this means that the actor is able to engage with an emotion *sincerely* rather than *authentically*. Bloch et al. indicate that this is of benefit to actors because it allows them to freely move between expressions of emotion without being bogged down in navigating the actual experience of emotions – a psychophysiological process that may exceed the time restrictions of the dramatic moment (1987, pp. 3-4). As a result, the actor does not necessarily need to feel an emotion in order for it to appear *natural* or *true*, so long as the actor is able to produce the correct effector pattern for the given emotion (1987, p. 15). For animators who need to produce nuanced and believable reference performances, but who have limited acting experience, it may be ethically more responsible for them to engage with the Alba Technique than other techniques that rely on the subjective experience of emotions[28]. If the actor's expressive components of the represented emotion are accurate, Bloch et al. argue that this "is sufficient to evoke the corresponding emotion in the observers (public)." (1987, p. 17). Through independent theatrical experimentation, Roxanne Rix validates the effectiveness of the Alba Technique but suggests that the requirement to consciously reproduce facial manipulation specific to a given expression

[28] Bloch et al. indicate that an actor's correct execution of the effector patterns frequently produces the subjective feeling of a given emotion, but that with extended training in the Alba Technique, actors learn how to step-in and step-out of effector patterns in order to sidestep subjective emotional experience (1987, p. 17). If the Alba Technique is taught to animators, it is advisable to guide them through the process long enough to learn these step-in and step-out techniques.

is unnecessary[29] (1993, p. 140). Using just breathwork and minimal postural coaching, the actors in her experiments demonstrate "vivid and natural" facial expressions (1993, p. 144). The Alba Technique may be seen as a type of codified system of emotional responses devoid of context for reliable recall and repetition. In this way, it shares similarities with Delsarte's codified system. However, while Delsarte based his system on the legibility of specific gestures to convey particular emotions, those gestures were situated within the social context of 19th Century European and American cultures, and may not have been decipherable to anyone who was unfamiliar with those contexts. By comparison, the Alba Technique is based on theories of cognitive science to produce universally recognisable signs of particular emotional responses.

Blair sees that one of the most valuable aspects of the Alba Method is its ability to dispense with Descartes' mind-body split altogether[30], moving us beyond a discussion of how aligned an actor is with a role. Likewise, by shifting the focus away from whether an actor's feelings or a character's feelings are being authentically represented at any given moment renders Diderot's paradox moot. She explains that

> [a] new vocabulary based on neurocognitive research provides a specific, material way of talking about the phenomenon of acting rather as that of a single organism, keeping the psychoanalytic or psychological in its proper and useful place. "Taking on a character" or "imbibing the reality" of a role become recontextualized as strategies about performance, because the character becomes a set of choices and behaviors—a process, rather than a discrete entity, a motivated movement, rather than a gloss of feeling—, supported by what the actor brings to the role.

The Alba Technique reveals a clearer basis for understanding whether the act of animating can be considered a process of acting. That is, how (and when) are animators also actors? If the Alba Technique is considered a technique of acting (which I believe it should be), then

---

[29] However, I wonder whether initiating and interrupting an effector pattern to produce a sincere emotional response that lacks authenticity leaves out some nuance to a performance that would have otherwise been present if authenticity were engaged. For instance, would an actor participate in spontaneity and "happy accidents" when navigating an authentic emotional experience that would not be present in the exclusive sincerity of the Alba Technique? These questions serve as meaningful territory for future research.

[30] Zarilli suggests that Leder's account of the *absent body*, in which our experience of the body is most often in a form of absence, provides a phenomenologically informed understanding of how Descartes may have arrived at his conception of mind-body dualism (2004, p. 656).

if an animator engages with the Alba Technique to produce animation reference, regardless of whether the animator authentically engages with emotion, the animator functions as an actor. If we consider the animator's role in working from the reference to produce the animation as a performance by proxy (Hosea, 2012, p. 67), then there is a foundation for considering animators as actors.

3.2.5 How Acting for Animation Differs from Live-Action

Acting for the stage is typified by the experience of an actor's physical presence performing in the present moment. While the stage actor may be subject to a variety of performance enhancements, including makeup, lighting, and costumes, the actor principally controls how a performance is received. Unlike in film, there are no middlemen who edit the stage actor's performance prior to its reception (Wolf, 2003, p. 48). By contrast, acting for film results in the recording of a series of "present moments" that are later edited to produce the illusion of a continuous dramatic event. If we ignore the process of recording reference footage, animation forgoes the present moment altogether, and the animator works frame-by-frame to construct the illusion of a present moment. In this case, the animator doesn't perform for a live audience (including a film director) but rather for an audience in his or her head (Hooks, 2005, p. x). While actors risk vulnerability by using their bodies as the sites of performance, animators never do so. Instead, animators perform *through* the graphic bodies of their animated characters (performance by proxy), a process that Hooks believes can lead to an emotional disconnect between the animator and character. In such instances, this accounts for the preponderance of "very cerebral and overly-cute performances [because the] emotionally disconnected animator [imposes] the performance rather than justifying it emotionally." (2011, p. 49). In order to overcome reliance on the cute and cliché, the animator must find emotional grounding for any story point or action he or she creates for the character.

An animator must re-motivate an acting moment over and over for weeks on end, painstakingly analysing every motion and expression of the performance. While an actor often relies on internal feelings to manifest believable external movement and expressions, the animator must live through internal feelings for weeks on end, figuring out how a character thinks and feels, and producing external signs that meaningfully correspond to those motivations (Jones & Oliff, 2007, p. 177). When an animator works from pre-recorded audio clips of a voice actor performing lines for a character, the animator must also determine the motivations and feelings the actor used to produce his or her performance of the character. That is, the animator not only needs to understand the character but also how the voice actor understands the character, and how this intent can be effectively conveyed through gesture and expression. Animators are concerned with how to create a specific acting result, whereas actors (within the realist style) are taught that you cannot act results (2007, p. 177).

Voice acting for feature animation is most often recorded in a sound booth with only a single actor at a time. This is despite the fact that the voice actor may be recording dialogue for a scene that features characters simultaneously talking to each other. While actors in the same shot need to be physically together when recording scenes for camera, this is not an imperative for animation. Hayes speculates that actors schedule voice sessions around their live-action work, which allows for more stars to be included in an animated film than if they all had to work together at the same time. However, he advises that it is his preference to record actors together whenever possible because he believes that the back-and-forth interaction between actors results in more natural performances with better rhythms (2013, p. 176). Director Gore Verbinski utilised this approach during the recording process for "Rango" (2011), which included the voice actors dressed up in costume and performing on a rudimentary set. When asked why he opted to go this route, Verbinski replied:

> I guess fear, really. Fear of a microphone and a cold environment and nobody's reacting to anybody. We didn't want a performance that was too clean. If people are running, I want to hear them out of breath in the next line when they stop and talk. (The New York Times, 2011, para. 16).

Hooks also makes note of the benefits of the ensemble nature of voice recording for animation. When recording the audio for "Monster's Inc" (Docter, 2001), John Goodman and Billy Crystal performed most of their lines in the studio together. Hooks believes that this resulted in an improvement to the buddy-film energy of this movie:

> Speaking from my background as an actor, I can assure you that acting alone is far more difficult than acting with a partner. Acting is reacting. If the other actor isn't there, you either have to depend on something that was recorded earlier or on your imagination. (Hooks, 2005, pp. 52-53).

When actors work together, they have each other's presence to play from. Rather than having to rely on their imaginations alone, the actors provide each other with physical cues, which can add extra layers of believability and improvisation to the performance.

Hooks admonishes acting in Hollywood animation as often over-reliant on the spoken word, and he cautions that acting itself has little to do with dialogue. However, when animators work through a scene that contains dialogue, they often start from pre-recorded audio and attempt to make the animation fit to the words. Hooks advises that in the real world, intention and movement precede words: a person first thinks, then responds emotionally, then physically – often very quickly but always in that order. When working from pre-recorded dialogue, animators must "work backward from the way actual humans function, finding the inner impulse – the thought – that would be expressed by the recorded words." (2011, p. 28).

Dean DeBlois, director of the "How to Train Your Dragon" trilogy of movies, indicates that in the production of Hollywood feature animation, voice actors are rarely given the full film script. He believes that this is due to the production management teams desiring to limit how much knowledge any one creative individual has over the full movie, especially if the producers lack confidence in the quality of the material. DeBlois indicates that it would be his preference for the actors to read the whole script because it would better contextualise each character, helping the actors to more clearly understand their roles, motivations, and relationships throughout the full story, just as an actor in a stage play would do. He also advocates for voice actors to record their dialogue together, noting that while a clean recording of a lone actor may aid the manipulation of individual lines, when actors run their dialogue together and "step on each other's lines", it results in a more believable character dynamic when animated (Bancroft, 2014, p. 88).

Film presents a number of medium-specific acting considerations that are different from stage acting. Significant among these is the adjustable vantage point of the camera, which allows much greater intimacy with the actor through close-ups. Bode writes that the "European film-makers Jean Epstein and Béla Balázs saw the cinematic close-up as revealing the micro-physiognomy of the human face, cuing viewers to search for signs of psychological interiority in the twitch of a lip or an eye-lid." (2015, p. 97). To this day, the close-up retains its power to convey the actor's face as the "primary locus of performance" (2015, p. 98). Actor-director Jane Drake Brody is critical of the film close-up's role "as the model of truth for a performer" and indicates that an over-reliance on the close-up results in inactive introspection at the cost of active storytelling (2017, p. 30). Hayes writes that the close-up is often avoided in traditionally animated films due to a belief that a drawn character's face has less performance detail than a real person (2013, p. 9). This is one of the reasons why 2D animation is especially concerned with conveying emotion through full body character performances, such as in two animations by Michaël Dudok de Wit: "Father and Daughter" (2000) and "The Monk and the Fish" (1994). These films provide no close-ups of the character's faces whatsoever and instead rely on staging, body language, contrast through timing, and how the spatial relationships between characters.

A film actor learns the need to perform in different ways depending on the cinematic framing used for a shot. For instance, in a close-up, it is desirable for an actor to limit body movement entirely and try to focus all of the acting effort through his or her face. However, in a long shot, small gestures may not register at all and the actor is encouraged to employ a broader body language for visual clarity. This likewise applies to acting for animation that is cinematically considered (Hayes, 2013, p. 9). However, in animation, greater attention is paid to the clarity of a character's silhouette (known as the animation principle of "staging") and creating a powerful "line of action" that strengthens the force and attitude of a pose.

While the stage actor maintains a continuous performance from the beginning to the end of a play, the film actor's performance is recorded piecemeal over a variety of takes, often out of order from the final sequence of the film, which requires the actor to emote in fragments rather than experience a continuous emotional journey. Balcerzak indicates that the edited nature of this process places the film actor at a disadvantage to the stage actor in terms of determining his or her final performance. This discrepancy is even greater for the MoCap actor, who "is literally stripped of his physical body to exist as pure kinesis – a marker cloud to be employed as a tool by the filmmaker." (2013, pp. 210-211). With each iteration – from stage acting, to film acting, to motion capture acting – the actor's role shifts from autonomy over generating the final performance to a utilitarian element for a director to dissect and reassemble.

3.2.6 Acting in MoCap Versus Animation

Acting in motion capture is convoluted and combines facets from stage acting, film acting, and animation: an actor's performance is produced and captured live, which can then be edited into a sequence combined from different takes. Animators are then able to adjust elements of the performance in a frame-by-frame fashion. Even more bizarre is that elements of performance from one take can be combined and reconstituted with another take – for instance, a continuous swordfight sequence may be the product of multiple MoCap takes that are edited together to create a seamless result. MoCap also allows for the blending of performances such that the swordfight actions of the upper body may be mixed with a running sequence from an entirely different capture session to produce a running swordfight performance.

Andy Serkis writes about a particular insight he learned while performing Gollum for "The Two Towers". Serkis worked with Weta Digital animation design supervisor Randall William Cook to plan the performance for a shot, which needed to match the pre-recorded action of the character Frodo, played by Elijah Wood. Serkis soon learned that the performance that seemed most "truthful" was the "quietest and least fussy" version (2003, p. 36). He recalls how Cook, who had a career as an actor before becoming an animator, indicated that the temptation within the frame-by-frame nature of creating animation is to make the character busy all the time, but that a character who remains physically still is often more powerful[31] (2003, p. 36). In screen acting, as in MoCap, a character's emotions and thoughts are primarily communicated through the eyes, and stillness is an important factor for carrying a close-up and allowing the audience to connect with a character. Serkis acknowledges that this is difficult to achieve when a character is over-energised, and that "in the later performances [he] was more aware that [he] could internalize the feeling and still be confident that it was reading to the viewer." (2003, pp. 84-85). Speaking from my personal animation experience, as the style of a character approaches a greater sense of realism, the more that the performance must likewise appear "natural" or "real" in order to create a psychophysiological coherence between form and function. A character full of exaggerated antics and timing suits a more graphic, cartoon aesthetic, while a performance grounded in the familiar limits of human (or human-like) action and emotion will generally be more consistent with the motion-captured performance of a photoreal character. One important barrier to producing a fully realistic performance aesthetic has been creating human-like digital eyes, which often fall short in terms of physical and kinetic detail. While recent strides have produced much more compelling digital eyes, this part of the human anatomy remains a formidable challenge to conveying a source-actor's full performance through a digital counterpart. For instance, in the film "Gemini Man", Will Smith plays Junior, a 23-year-old synthespian counterpart to Smith's 50-year-old self. The quality of the synthespian's realism is among the best digital human work to date, and sustained mid-shots and close-ups suggest a certain confidence in the handicraft that encourages the viewer to search for flaws. The eyes are also of an excellent standard overall, but there are a few times throughout the film when they fall short of absolute believability, and in

[31] Serkis critiques the busy nature of traditionally animated performances as often being absent of reflection or repose (2003, p. 84). This is a debatable point, especially in terms of the nature of storytelling and genre, as well as feature-length animation versus Saturday morning breakfast serials. However, while his argument has flaws, it also carries some truth: animated characters are often over-animated for the sake of entertainment rather than producing compelling and believable performances. Serkis made this comment in 2003; it would be interesting for him to revisit this opinion in light of a more recent range of animation performances in the two decades since.

those instances the audience's total belief in the performance becomes compromised in a moment's uncanniness. Director Ang Lee's insistence on revealing Junior through close-ups is emblematic of an apparent compulsion among filmmakers whenever they believe a digital character sets a new standard in photorealism. This pattern dates back to "Final Fantasy: The Spirits Within" (Sakaguchi & Sakakibara, 2001), the first film to feature believable[32], motion-captured humans, in which the character Aki Ross is introduced through extreme close-ups of her eyes and face.

McNary writes how during the production of "Avatar", James Cameron was enthusiastic about how recording actors via PeCap allows the director to provide actors with his or her undivided attention. Cameron also believes PeCap requires less footage than conventional filming, where each take often needs to be shot from multiple vantage points and duplicated by the actor. A single PeCap take provides a full 360-degrees' worth of footage, and allows for framing considerations to be finalised in postproduction rather than at the time of filming. For Cameron, this provides opportunities for exploratory and creative takes because performances don't have to be duplicated (2011, paras. 8-10). For Robert Zemeckis, PeCap allows actors to work a scene in continuity with all of the other actors in that scene. This approach is similar to stage acting and rarely, if ever, exists when shooting a standard film; it generally results in a more immediate acting environment with better interactions between actors (Ressner, 2013, para. 34). Zemeckis lists a number of traditional filmmaking limitations that can negatively affect an actor's performance that are obviated when recording PeCap for a fully digital character, including: actors hitting their marks on the floor (to ensure they are at the proper locations for camera focus), avoiding shadows from boom mics, reapplying damaged makeup, and fixing the volume and placement of wigs (Kehr, 2016, para. 27). It is important to note that Zemeckis' comments are specific to actors working in a purely PeCap environment, such as a motion capture stage. When the actors have to be filmed as part of a live-action scene (to be later digitally removed and replaced, such as in "War for the Planet of the Apes"), many of these limitations still matter.

---

[32] "Believability" in CG characters is always time-dependent: what is considered believable in one year may be diminished by the standards set in subsequent years.

Performance capture is often performed on a motion capture set within a prescribed capture volume. While these volumes can range from a few cubic metres to the size of a football stadium, the PeCap actor's performance is bounded by the limitations of the available equipment. On the other hand, 3D software has virtually limitless parameters, and it is important that both the directors and actors understand how the physical limits of the real-world setup will impact the results within the virtual space (Delbridge, 2015, pp. 16-17). A MoCap stage generally utilises grids of coloured tape affixed to the floor to indicate the boundaries of the capture volume. These are imaginary borders beyond which the system is unable to accurately record the performance data.

Within acting for stage and film, there is often reference to the concept of "the fourth wall", which acting coach Ivana Chubbuck describes as "the dimension of the [place] that makes the space you're working in (stage, set, classroom, location, et al) private, separating the actors and the stage from the audience or camera crew." (2005, p. 133). The role of the fourth wall is an actor's imaginative device to provide a sense of enclosure, intimacy, and privacy within what is really factitious surroundings. Chubbuck advises that the actor should personally endow the imaginary fourth wall for a given space with specific details about its floors, walls, and furniture in order to make it psychologically effective (2004, p. 133). On stage, the fourth wall exists between the actor and the audience, and during filming the fourth wall sits at the limits of the set between the actor and the camera. However, a motion capture stage, similar to theatre in the round, is a fourth wall in every direction – there are no actual barriers between actor and audience, only imagined ones. The PeCap actor must employ a highly developed sense of imagination to mentally set-dress what sits beyond the limits of the capture volume in every direction. However, when a production makes use of newly developed LED wall technology to create real-time, fully immersive virtual sets (such as those used in the series "The Mandalorian" [Favreau, 2019-2020]), PeCap actors are able to see their factitious surroundings in rich detail, thus reducing (or eliminating) the need for a fourth wall.

Wolf indicates that the use of doubles is common in live-action filmmaking, and each type of double provides some kind of property or ability that the actor does not have (2003, p. 49). He outlines nine types of doubles based on the separable component they provide from an actor's performance:

• Body	(doubles)
• Voice	(dubbing)
• Face	(face replacement)
• Movement	(motion capture)
• Facial Expression	(motion capture)
• Timing	(temporal editing)
• Image	(scanning and remapping)
• Shape	(three-dimensional scanning)
• Behaviour	(algorithmic simulation) (2003, p. 55).

Likewise, motion capture frequently features layers of doubling as the MoCap actor's performance is "divided into separate components, each of which can be individually manipulated and recombined in postproduction." (Kennedy, in press, p. 6). Working from Wolf's list of performance doubles, I adapt each of these components to how they may be used within a motion capture context:

- **the body** (capable of being swapped with a performance double);
- **the voice** (capable of being dubbed);
- **the face** (capable of being replaced with another face);
- **movement** (editable via motion capture, including combining movement from one or more performers);
- **facial expression** (editable via performance capture to adjust an expression, alter a performance, or create a mapping between different physiologies);
- **timing** (retiming movement);
- **image** (texture detail, such as skin and animated blood flow);
- **shape** (3D-scanning an actor's physiology); and
- **behaviour** (algorithmic simulation, such as crowd simulation) (Kennedy, in press, p. 6).

In the movie "Deadpool" (Miller, 2016), the digital character Colossus provides an excellent example of how multiple actors, each lending discrete performance components, can be combined and reconfigured into the performance of a single digital counterpart. A video by Wired demonstrates how Colossus was constructed from five separate performers (Chitwood, 2016):

> a motion capture recording of a performer enacting Colossus' physical movements (**movement**); a performer in a tracking suit and platform shoes filmed onset interacting with other actors (**body**); a performer whose face was used as the basis for Colossus' facial features (**shape**); a performer who produced all of the dialogue (**voice**); and a performer responsible for producing the facial performance of the character (**facial expression**). (Kennedy, in press, p. 6).

The fact that a single unified consciousness on screen can emerge from the contributions of multiple performers was originally outlined by drama theorist Jiří Veltruský. He proposed that within the context of theatre "there are two components to any dramatic action: (1.) the intent and (2.) the performance of the intent", and that these two aspects could be enacted by more than one person (Veltruský, as cited in Hosea, 2012, p. 69). Traditionally in stage and film, the character and the actor are one in the same body. However, a digital counterpart is a virtual body onto which one or more actors' performances are translated. The nature of authorship within a MoCap environment is convoluted not just by multiple actors, but also by multiple animators who contribute to the final performance product. Wolf describes this as the "divisibility of performance" and indicates that performance "needs to be rethought due to the large number of people who may be involved in the creation of a single example." (2003, p. 48). An original contribution of Hosea's doctoral thesis is an application of Veltruský's theory to animation, and Hosea indicates that despite there not being "one unifying consciousness and embodied presence creating the performance, a coherent character can still be read from a unified creative intention" (2012, p. 68). Such a creative intention is mapped onto the digital counterpart's body in order to produce a performance by proxy – a substitution for the actors and animators alike (2012, pp. 89-90). Conscious of the constructed nature of doubled performances, Shacklock critiques the tendency in the Stanislavski school of acting[33] to:

[33] And, by extension, the Strasberg and Hagen schools of acting.

> reiterate notions of a coherent and autonomous self [… whereby] acting emerges from the actor's sense of self […]. Here, acting is the project of the single actor: the expression of a coherent, authentic, interior self and the work of the creative, autonomous agent. (2016, pp. 70-71).

The imperative to create the appearance of an animated character's single, unified consciousness requires certain production considerations for animators. For instance, animation director Derek Hayes indicates that on industrial-scale productions, all animators who contribute to a single character must produce their work as if it were all created by the same animator. That is, there is no space for an *auteur* approach[34]: the animated character must appear to consistently perform from moment to moment and from scene to scene (Hayes, 2013, p. 21). Within motion capture, the production considerations are slightly more complicated as (usually) a single actor creates the basis for the character's entire performance. It is then up to the animators to preserve and/or interpret that performance. Instead of individual animators working to produce the illusion of a unified performance from scratch, MoCap animators work together to maintain the illusion of a single author of the motion-captured performance, even when modifying aspects of the performance to alter what was achieved during the live capture.

Wolf writes that instead of MoCap actors being the sole authors of a performance, "they may become more a supplier of raw data to be combined into a performance and shaped by others. Performance, then, becomes a special effect, and nuanced emotion a technical challenge as well as an artistic one." (2003, p. 55). In her essay "Blackface, Happy Feet", which outlines racial politics within motion capture, Tanine Allison critiques a common refrain of directors who promote MoCap as a direct transferal of the actor's performance: "[t]his piecemeal characterisation flies in the face of those – like Serkis, Cameron or director Peter Jackson – who would like to see the digital character as a direct extension of the human performer, transmitting every nuance of expression and intention." (2015, p. 123). In a

[34] Hayes indicates that auteur approaches to animation are more common in smaller-scale productions, especially if the animator is responsible for the design of the character(s), story, or other creative aspects of the production (2013, p. 22).

similar critique, Balcerzak challenges the common Hollywood philosophy that MoCap is the digitisation of the human, and counters that MoCap is truly about the humanisation of the digital: "[i]t is a process developed to make the special effect perform realistically as opposed to, as suggested by many, digitally enhance the actor." (2013, pp. 196-197). Such a dominant understanding of acting as the product of the contributions of an individual rather than a collective (Dyer, 1986, p. 108) foregrounds the necessity for a paradigm revision for acting, especially within a PeCap context (Pallant, 2012, p. 38).

How are we to understand the nature of acting across the myriad manifestations of technologically constructed performances? Certainly, in a role where a visual effect forms a minor part of a character design and the actor remains visible in the role (such as the digitally augmented character Cable, played by Josh Brolin in "Deadpool 2" [Leitch, 2018]), there is generally little question as to who is responsible for the performance. However, when an actor disappears within the character design of a digital counterpart and the performance becomes subject to the creative decisions of multiple artists, individual responsibility for the performance becomes less clear. If the human source-actor cannot be seen beneath the digital character, is it legitimate to consider the digital counterpart to be "acting"? To demonstrate how acting is influenced by various technological approaches toward performance, the following section attempts to define and map out a continuum of screen performance categories based on their respective degree of technological mediation.

3.3 Technological Mediation of Screen Performances

As a critique of academic and industry discourses that intrinsically bind acting to biological individuals rather than animated characters, Bode suggests a screen performance continuum that embraces acting based on the degree to which a performance is technologically mediated. She proposes that the continuum ranges from acting-centred performances to technologically bound performances, in which an actor's recorded performance is "reconstituted or repurposed within new mise-en-scène and interactional contexts." (2010,

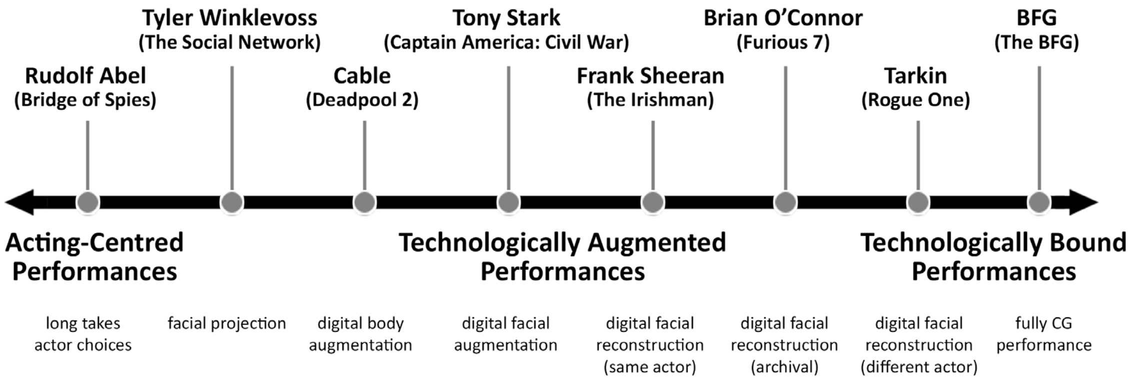


Figure 8. Expanded screen performance continuum, based on an idea proposed by Bode (2010).

pp. 69-70). Bode proposes this as an idea but doesn't actually draw out the continuum. However, I believe a visual represenation of Bode's continuum would be valuable as a means to better understand the nature of performance *vis-à-vis* technological mediation, and specifically how it may be used to locate acting within animated characters.

Expanding on Bode's idea, I present a detailed interpretation of this continuum that charts a variety of technologically mediated performance types within live-action films (Figure 8). Between the poles of acting-centred performances and technologically bound performances, I've added an intermediate territory that I refer to as "technologically augmented performances". This description recognises that within the middle-ground, performances can remain embedded within a flesh-and-blood basis while also being digitally altered.

At the left side of the continuum, I cite Mark Rylance's portrayal of the character Rudolf Abel in the film "Bridge of Spies" (Spielberg, 2015) as a prime example of a fully human performance grounded in an actor's choices. Rylance, who is described as "the greatest actor of his generation" (Caesar, 2016), has successfully navigated from a celebrated stage career to film acting. His performance as Abel feels effortless, as if there is no trace of acting whatsoever – so fully does Rylance live the role. His strength as a screen performer is in his restraint and composure, which he deftly uses to hold back a well of emotion that resides within his eyes. As Constance Grady writes for the news website Vox: "Rylance's secret weapon in any medium is the earnest, open vulnerability on his face" (2018, para. 10). There is nothing technologically mediated by his performance aside from the editing used to combine his takes.

As the continuum moves toward the right, an actor's performance gradually becomes more technologically augmented, eventually resulting in the actor's original performance being substituted by a fully digital replacement. In "The Social Network", identical twin brothers Cameron and Tyler Winklevoss appear to be played by the same actor, Armie Hammer. However, Hammer played the full role of Cameron while working alongside Josh Pence, who functioned as a stand-in for the role of Tyler. Both actors were recorded in the same scene together, swapping roles so that Hammer could play both characters. Lola Visual Effects subsequently used digital facial projection techniques to track Hammer's original performance over Pence's body in order to produce the appearance of Hammer simultaneously playing both roles. While this method erased the major focal point of Pence's performance, it preserved (while transposing) Hammer's performance (Lévêque, 2013).

Toward greater technological augmentation, we encounter roles such as Cable in "Deadpool 2", which fully preserves actor Josh Brolin's original facial performance but digitally augments his torso and one arm to produce the appearance of a cybernetically enhanced body. In "Captain America: Civil War" (Russo & Russo, 2016), Robert Downey Jr plays a teenaged version of the character Tony Stark, physically equivalent to how Downey looked in films from the mid-1980s. Instead of creating a synthespian double for this de-aging process, Lola VFX preserved Downey's original performance while using digital compositing to remove wrinkles, perform face-lifts, and adjust the appearance of blood flow to produce a more youthful glow. Visual effects supervisor Trent Claus describes this process as similar to using Photoshop, except that while Photoshop works on a single image, this process is performed on footage filmed at 24 fps (Giardina, 2016, paras. 4-6).

In "The Irishman" (Scorsese, 2019), Robert De Niro, who was 75 at the time of filming, plays the character Frank Sheeran at the ages of 24, 36, 41, 42, 47, 55, and 76. VFX studio Industrial Light and Magic (ILM) used proprietary techniques to digitally de-age De Niro's performances (Seymour, 2019, para. 4). This resulted in De Niro's facial performance being tracked and mapped onto a highly accurate, but thinner, 3D model of the actor. Skin texture information was also captured and modified to remove wrinkles and age-related pigmentation. This process recorded the underlying performance and reprojected it via digital facial reconstruction to replace De Niro's profilmic facial performance. This is critically different from the de-aged Tony Stark example, which adjusted the original footage through digital compositing but did not alter the fundamental basis of that footage.

When Paul Walker died midway through filming "Furious 7", the seventh film in The Fast and the Furious franchise in which Walker played the lead character Brian O'Connor, Universal Studios contracted Weta Digital to produce a photorealistic synthespian version of Walker that could seamlessly perform side-by-side the live-action shots of the actor that had already been filmed. However, unlike directly recording an actor's performance via motion capture (or via the infrared Flux camera rig used for the Irishman) and retargeting it onto a digital Walker, director James Wan opted to have animators work from footage of Walker from previous films in the franchise to ensure that the performances on screen matched Walker's nuances and mannerisms as closely as possible. This was meant to avoid having animators rely on artistic interpretation to reconcile how Walker might have moved differently from another source actor playing his character (Giardina, 2015). Creating an actor's performance from archival footage is a specialised case along the screen performance continuum, sitting between a source actor whose performance is retargeted onto his or her synthespian double (digital facial reconstruction of the same actor), and a source actor whose performance is retargeted onto the synthespian double of another actor (digital facial reconstruction of a different actor). As an example of the latter, when the production of "Rogue One: A Star Wars Story" required the digital recreation of Peter Cushing's character Grand Moff Tarkin from the original "Star Wars: Episode IV – A New Hope" (Lucas, 1977), ILM was tasked with digitally resurrecting an actor[34] who died in 1994 (Cavna, 2016, para.

[34] Even if Cushing remained alive at the time of filming "Rogue One", the production team would still have required a digital version of Cushing that matched his physiology from 1977.

2). In order to do this, director Gareth Edwards cast Guy Henry to play Tarkin due to Henry's similarity to Cushing in terms of stature and manner of speaking (Itzkoff, 2016, para. 10). Henry's facial performance was recorded on set via a PeCap rig and was retargeted onto the digital Tarkin character. In this way, Henry plays the role of Cushing playing the role of Tarkin. The result is a transferal of performance from one physiology to another, but the target physiology still matches that of a recognisable flesh-and-blood actor.

At the far right-side of the continuum is a full CG performance in which a flesh-and-blood actor produces a performance that is retargeted onto a fully digital character that has no real-world counterpart. Here I return to Mark Rylance, whose portrayal of the BFG (Big Friendly Giant) in "The BFG" (Spielberg, 2016) provides an excellent example. It is fascinating to note that all of the same qualities I admire in Rylance's performance as Rudolf Abel are also present in his synthespian counterpart in "The BFG": the effortlessness, restraint, and open vulnerability. However, at this far end of the screen performance continuum, Rylance neither performs through a digital version of himself or another actor, although the BFG's face does possess traces of Rylance's physiology. While his performance as Abel is unadulterated beyond standard live-action filmmaking techniques, Rylance's performance as the BFG is fully technologically mediated, with every aspect of his performance captured, digitised, quantified, retargeted, evaluated, and subject to revision.

3.4 Conclusion

This chapter established meaningful definitions for terms related to performance and acting within a motion capture context. A historical overview of the development of realism within Method acting was explored and laid a foundation for the most common style of acting within performance capture today. Importantly, this chapter discussed key differences between how acting is produced for live-action versus animation and performance capture.

When considering acting within a performance capture context, a number of concerns arise that are unique from live-action. For instance, how do we understand the divisibilty of performances in PeCap when multiple people are responsible for constructing the final product of acting for a digital counterpart? Acting is predominantly understood as the product of an individual rather than a collective. This is unsurprising in that the majority of the history of acting involves individuals on stage producing their performances in real-time; there never was a reason to question the origin of an actor's performance because it was self-evident. Even in the relatively recent invention of screen acting, an actor's performance was still considered the sole product of the actor despite any number of edits that might occur after filming. Method acting's focus on the actor's examination of self further underscores the sense that acting is produced by the actor alone. However, the technical realities of performance capture considerably draw into question how acting is produced and by whom. Since the source-performance is often never seen by the audience but is rather translated onto a digital counterpart – which may or may not resemble the source-actor – the viewer is left with a certain mystery about how exactly the character's performance is derived from the actor. A common thought may be: "a computer must be involved" – but exactly what that entails is unclear to the average person. Yet, even the people who oversee the most prestigious film awards seem similarly confused. This chapter attempts to provide a foundation for better understanding this nebulous transfer of acting between human performers and digital characters, and is further developed in the following chapters.

Academic and industry discourses often assume that acting is exclusively the product of biological individuals rather than animated characters. This chapter makes the case that animated characters can be seen to act. Likewise, the expansion of the screen performance continuum discussed in this chapter demonstrates a means by which to locate acting in animated characters. Through Hosea's application of Veltruský's drama theory to animation, we are able to see the contributions of multiple artists (beyond just the source-performer) as a single coherent character. As the degree of technological mediation of a character increases, we understand that the character's on-screen performance is the product of an increasingly large collective of artists. These artists generally try to uphold the original intent of the source-actor's performance, but this is at the director's discretion and in certain productions the final performance may significantly vary from the source performance. The more visually similar the digital counterpart is to the source-actor, the more likely the final on-screen performance will adhere to the source-performance; however, digital counterparts that are physically quite different from their source-actors may require much more animation intervention to perfect their performances.

When we consider what acting becomes in the context of performance capture, especially when a digital counterpart's physiology significantly varies from that of a source-actor, we start to see acting as a type of performance-by-proxy. This recalls Hosea's description of animation as a form of performance-by-proxy, as well. Perhaps it is reasonable to conclude that within a performance capture context, the human source-actor is merely one strategy (or component) for producing animation. Therefore, within performance capture, acting is a process carried out in several phases by multiple creative teams, each of which provides distinct artistic and technical skills necessary to produce a finished on-screen performance. The prevailing understand of acting as the sole product an individual's contributions seems insufficient in this case, and performance capture provides a clear example of how acting can be understood as the product of a collective.

With an understanding of how animated characters – and specifically performance-captured characters – are capable of acting, the following chapter discusses terminology common to an actor's experience and considers how such terminology is applicable within a PeCap context.

4

WHAT IS REAL?

4.1 *Authenticity, Truthfulness, and Aura*

Authentic, core, organic, real – these words are used in every one of my acting classes. Their meanings are tacitly understood within that context, but exact definitions for these terms are elusive. These terms are frequently used to describe "good acting" (Bode, 2010, p. 58), so any qualitative measure of acting needs to be grounded in clear definitions of these terms. Likewise, if we are to evaluate how elements of acting, especially *good acting*, are transferred from a source-actor to a digital counterpart, we have to be able to meaningfully define the nature of that exchange. Actors don't necessarily need such exactitude in order to hone their craft, but an academically rigorous inquiry demands our communication to be clear and fully considered.

This chapter provides an overview of the language specific to understanding how acting is produced within a Method acting perspective (and its descendents), as many of these concepts are value judgments from within the Method traditions. For the purpose of this thesis, the application of such terminology is primarily linked to Western films featuring motion capture performances. This overview includes how expressions and emotions are conveyed, interpreted, and understood, and how this creates a basis for empathy with animated characters. How are we to understand who is actually "acting" in the context of performance capture – the source-actors, the animators, or the digital counterparts themselves? As many digital counterparts are designed to look "real", this chapter examines how realism is conveyed through the design and performance of digital characters, as well as challenges posed by uncanniness. By reflecting on these concepts, this chapter suggests a categorisation of different performance types specific to a subset of digital characters.

4.1.1 Authenticity and Sincerity

Terms frequently used to describe acting – such as "real", "natural", imitative", "truthful", "character", "action", and "emotion" (Blair, 2008, p. 23) – are grounded in the subjective experience of the individual actor, the exact understanding of which may vary from person to person, which makes it difficult to demand precise definitions. Compounding this issue is the fact that the terms are often used inconsistently – sometimes interchangeably – resulting in different meanings for different people. A willingness among actors to appropriate pseudo-scientific or pseudo-psychological vocabularies and techniques for their own purposes is one of the most difficult aspects of understanding the subjective nature of the actor's process (2008, pp. 1-2). Likewise, how do we as an audience know when we are in the presence of authenticity, honesty, or the real? Blair relates a common experience among actors of feeling like one has fallen short of achieving authenticity or honesty only to receive accolades from the audience for an outstanding performance (and vice-versa) (2008, p. 14). If the actor doesn't feel like he or she performed in a truthful or authentic way, how is it that the audience perceives such truth and authenticity? How do the contexts of animation and motion capture affect the way we understand the notion of authentic performances?

I will start with an understanding of what comprises *authenticity*. "Authentic" is derived from the Greek word *authentes* (the combined form of *autos* [self] + *hentes* [doer]), meaning "one acting on one's own authority", from which *authentikos* (meaning "original, genuine, principal") also derives (Harper, 2020a). Professor of sociology Rebecca Erickson notes that different disciplines assign their own definitions to authenticity and commonly, but erroneously, conflate it with sincerity. She cites Trilling's (1972) understanding of sincerity as "whether a person represents herself truly or honestly to others" (Erickson, 1995, p. 122), and views authenticity as how consistently a person represents oneself to oneself. Sincerity is outwardly focused, whereas authenticity is self-referential and does not include any reference to "others" (1995, p. 125). As such, authenticity is a goal rather than a fully achievable state of being: according to Heidegger (1962), "one is neither authentic nor inauthentic but more or less so" (Erickson, 1995, p. 122). That is, at any given moment, an actor may closely represent his or her true feelings to him/herself; in the next moment, the actor may also believe he/she is representing true feeling to him/herself but is actually experiencing false feelings. To better understand these differences, I define these terms within an acting context as:

Authenticity: the degree to which an actor represents a character to him/herself.

Sincerity: the degree to which an actor represents a character to the audience.

When an audience watches an acting performance, are they more concerned with sincerity or authenticity – that is, does the audience value actors more who fully and consistently represent themselves to the audience or to themselves? It may be difficult to imagine how someone can truly and honestly represent oneself to others while not also doing so to him/herself. However, this may be possible in situations where a person unknowingly transmits honesty about him/herself without being aware of it or internalising it. For instance, following a recent breakup of a romantic relationship, if a person expresses outward signs of emotional pain (e.g.: through facial expressions, body language, behaviour) while inwardly denying that he or she feels hurt (e.g.: through denial as a coping mechanism), then this is a sincere emotional response with inauthentic self-awareness. By contrast, someone may be emotionally consistent with him/herself (authentic) while not clearly or fully expressing that emotion to others (insincere). Returning to the previous example, this could happen if someone feels anguish as a result of the breakup but doesn't want the other person to realise it, instead preferring to express indifference or even joy. The ability to be sincere without being authentic, and to be authentic without being sincere, helps to explain how an actor can convey a believable emotional state without fully experiencing it. Likewise, it explains how an actor who feels like he or she experienced a powerful emotional moment may not convey the depth of that experience through outward signs recognisable by an audience. Furthermore, film editing may also produce an emotion that an actor never directly experienced. In early cinema, Soviet filmmaker Lev Kuleshov demonstrated that a viewer

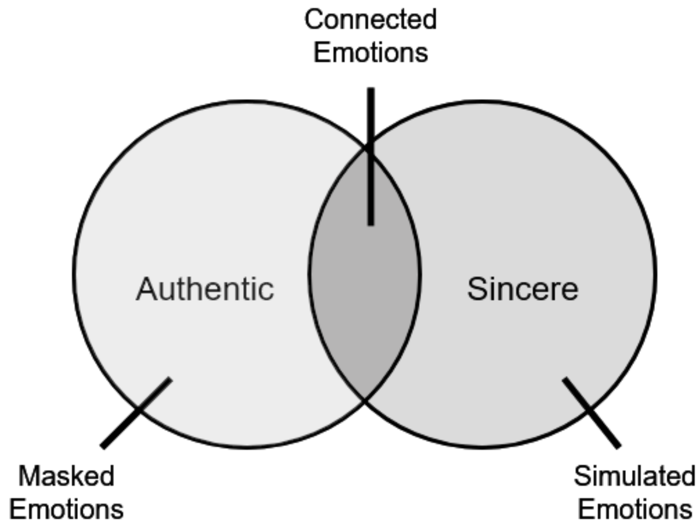


Figure 9. Relationships between authentic and sincere emotional experiences.

derives meaning from two sequential shots greater than from a single shot alone (Bode, 2010, p. 47) – such as a man staring off-camera followed by the image of a dead child. In this example, the so-called Kuleshov effect results in the viewer sensing the man's sorrow even if the actor portraying the man was emotionally neutral.

Authenticity and sincerity should not be seen as mutually exclusive, however. It is possible for an actor to represent his/her true feelings to him/herself while expressing those same feelings to an audience. I define an "emotionally connected" performance as one in which an actor truly represents a character to him/herself while also truly representing the character to an audience. By considering the permutations of authenticity and sincerity, we can define relationships between these emotional experiences (Figure 9). For instance, when an actor is both authentic and sincere, this results in *connected* emotions. When the actor is authentic but not sincere, this produces *masked* emotions, whereas being sincere but not authentic results in *simulated* emotion. The concept of simulated emotions provides a vehicle for how an "unfeeling" animated character is able to convey emotions to an audience. It is reasonable to speculate whether a sincere performance that lacks authenticity is visually and affectively identical to an emotionally connected performance. Certainly, it is possible that the addition of authenticity to sincerity may produce an emotional response that feels more nuanced, more real, or more interesting, although further research would be necessary to test this hypothesis.

While some actors and directors may argue that any sincere response would be enhanced by concomitant authenticity, ultimately the audience is only concerned with what it can perceive through sincerity. This has ramifications for a PeCap actor, as well: if the source-actor provides a sincere performance, the imprecise nature of the facial capture technology may not record all of the external signs that the human mind requires to recognise sincerity. If any of those details are lost or compromised, it becomes the responsibility of the PeCap animators to restore them to the best of their ability. However, this process is only as good as the flexibility of the digital character's facial rig combined with the skills of the animators at recognising and replacing what is missing from the captured performance. Fundamentally, PeCap animators are concerned with sincerity and not authenticity.

Animation professor Leslie Bishko is an expert in Laban Movement Analysis and has made a significant contribution to the field of animation studies by applying the Laban system to animation processes. She describes how, when portraying a healthy individual with no physical maladies, a live actor relies on his or her body to produce signs of sincere[35] characterisation (2014, p. 54). The only exception to this may occur if the actor must portray a character who is physically different to the actor in some way – whether through disfigurement, age, or disease, as examples. In such cases, the actor must enact outward signs of the physical differences in such a way that the audience believes the actor experiences them. While a less common concern in live-action, animated characters often considerably differ in size and proportion to human actors/animators. One of the trickiest aspects of any animated character's performance is the creation of a believable sense of weight. Since animated characters don't actually have mass and aren't subject to real-world physics, giving them a sense of weight is an illusion, but a necessary one in order for the audience to believe that the characters are (literally) grounded in their worlds. Disney's 12 Principles

[35] In Bishko original text, she refers to this as "authentic characterization". However, Bishko's understanding of *authenticity* is more similar to the description of *sincerity* in this thesis, and to avoid confusion I describe her concepts as "sincere" rather than "authentic".

of Animation were developed to produce a greater sense of believable movement within a cartoon style, but they do not provide complete movement concepts. Bishko describes how these principles are often relied upon by animators who do not consider how such overuse diminishes the sincerity of character performances (2007, p. 24). Sincerity allows the audience to suspend their disbelief and accept a factitious animation as a character who genuinely thinks and feels. Sincerity becomes the basis for an audience to empathise with an animated character as sincerity implies authenticity, even if the animated character is *ipso facto* incapable of feeling anything. Bishko cites the eponymous character of the film "The Iron Giant" (Bird, 1999) as an example of sincere character animation, whose emotions and movements are consistently expressed in appropriate measure relative to each other (2007, p. 25). By contrast, she finds the over-gesticulation of the character Ratchet in the film "Robots" (Wedge, 2005) to be an example of a performance that exceeds the qualities of the character suggested by the voice actor, resulting in inconsistency between character intention (as suggested by vocal inflection and tempo) and physical action.

> At one point during his speech to the employees, Ratchet does a short "dance" in which he is imitating what he perceives as childlike, whiny behavior, using a squeaky voice and shaking his hips back and forth with elbows flexed close to his ribcage and hands dangling limply. These movement qualities are highly inconsistent with what we have seen so far. Furthermore, the repetitive rhythm of his movement phrases is quite similar to the movement of other characters throughout the film. With all of these factors combined, Ratchet's dialogue says one thing, but his movement says something else. A style of movement has been used without attention to it's [sic] appropriateness for this character. (Bishko, 2007, p. 25)

Believability and consistency of movement is not just a performance concern for keyframe animation: even raw motion capture data based on real-world mass and gravity often feels weightless or "floaty". MoCap animators often need to "dial up" the sense of weight on digital counterparts, sometimes by ensuring that feet are more securely planted on the ground, adjusting the timing and spacing of actions, or increasing a sense of pressure on the hips and legs, among other techniques (Animation Mentor, 2013). The goal of this process is to produce *functional believability*, which Bishko describes as "character movement that has a high degree of biomechanical accuracy", including a "range of joint motion [that] falls within normal ranges for human performance", a sense of gravity through "weight shifts, dynamic alignment, and distribution of weight in motion", and the timing and subtlety of rotations in joints that "mimic the refinement in human motion." (2014, p. 51). Bishko indicates that sincerity in an animated character's performance occurs when the actions are both functionally and expressively congruent with the character as a whole: "[w]hen presented with well-crafted, [sincere] character performances, not only do we believe in the illusion of a virtual being, we believe there is an actual, specific being." (2014, p. 52). Bishko's argument is based on how well the character conveys congruency to the audience rather than to itself – as previously mentioned, an animated character has no inner self to compare its thoughts and actions against. A functionally and expressively congruent animated character is still bereft of consciousness and lacking in the psychological basis for authenticity. Until Creed's concept of a cyberstar emerges and can successfully pass a Turing test, this will remain a fundamental ontological difference between digital and flesh-and-blood actors – perhaps significant enough for digital actors to not be considered actors at all by some people.

4.1.2 Truthfulness

The relationship between truth and acting has been a source of theorising and disagreement throughout the history of the craft. Truthfulness is frequently conflated with authenticity and sincerity, and while it is related to these concepts, it is not equivalent to them. In defence of Method acting, which has often been criticised as "unreal" and "untrue", Hethmon notes that "truth" in acting has no unequivocal meaning. Indeed, each actor, director, dramaturge, and theorist may have a different idea of what truth in acting is, and be equally entitled to their positions. Strasberg advises that an actor is *real* (e.g.: *truthful*) "when his overt response neither exceeds nor falls too short of his imaginative impulse." (Hethmon, 1965, p. 197). Using the performance language developed in this thesis, we can understand Strasberg's definition of *truthful* to indicate when an actor's sincerity is directly aligned with his or her authenticity. However, the actor's job is not just to play an emotion as naturally as possible, but rather to convey the emotion based on the dramatic demands of the story. If the script requires the character to enact rage but the actor – out of a desire for naturalness

Figure 10. Jennifer Connolly's digitally inserted tear in "Blood Diamond".

and a concern for over-acting – summons only tepid anger, then the performance rings untrue (1965, p. 197). I suggest a provisional definition of *truthfulness* as "the degree to which an actor represents a character consistent with the needs of a story." This definition is functionally similar to the previous definitions of authenticity and sincerity; however, while authenticity is inwardly focused and sincerity is outwardly focused, truthfulness is *dramatically* focused.[36]

The film "Blood Diamond" (Zwick, 2006) provides a famous example of how truthfulness is sometimes manufactured in postproduction when a performance falls short of the dramatic requirements of a scene. In one of the film's climactic shots, Maddy Bowen (played by Jennifer Connolly) speaks on a mobile phone to her mortally wounded

[36] As an interesting example, when the Alba technique produces only outward signs of an emotion without the actor actually experiencing the emotion, then the performance is by definition truthful while being inauthentic yet sincere.

informant Danny Archer (played be Leonardo DiCaprio). This is a scene with intense emotions that Connolly ably performs, but despite conveying emotional sincerity, she did not cry. Director Edward Zwick felt that the emotional impact of the scene would be heightened if Bowen wept, and he instructed VFX artists to digitally add a tear rolling down her cheek (Figure 10) (Hoyle, 2007, para. 5). While the director was content with Connolly's performance during filming, he judged her expression as lacking sufficient truthfulness in postproduction. Hoyle records a leading technician on the film as saying: "Acting is all about honesty, but something like this makes what you see on screen a dishonest moment [….] Everyone feels a bit dirty about it." (2007, para. 6). By digitally adding outward markers of emotions, the resulting character performance is more truthful but less honest – that is, it better meets the dramatic needs of the script while altering the actress' original performance. I suggest a provisional definition of honesty as "the degree to which the sincerity of an actor's original performance is conveyed through the final screen performance". That is, how closely a character's final screen performance conforms to the actor's original performance. While digital alterations were noteworthy in 2007, they have become common practice in films today and include "opening or closing eyes; making a limp more convincing; removing breathing signs; eradicating blinking eyelids from a lingering gaze; or splicing together different takes of an unsuccessful love scene" to produce the illusion of better chemistry (2007, para. 7). Honesty may also be affected by editing, which has the ability to change the perceived quality of a performance.

While these changes may result in dishonesty, they are cheaper than reshooting a scene (Hoyle, 2007, para. 8) and they often rescue a scene that is otherwise full of sincerity and truthfulness, but which may have been compromised by a moment's mistake or unwanted reflex. In the context of motion capture, a director's intent may be to preserve a source-actor's performance, but the synthetic nature of a digital counterpart makes for acting that is infinitely editable. PeCap offers the temptation to further craft a performance in postproduction in the pursuit of dramatic truthfulness, but this always comes at the cost of impoverished honesty. Yet, as I mentioned in subsection 3.2.5, it is increasingly the role of the actor to serve as a mere utilitarian performance element for a director to dissect and reassemble. The question then becomes: if a performance can be so easily stripped apart and reconfigured, is there a fundamental aspect to the original performance that is lost when it is separated from its original *gestalt*? How are we to conceive of the aura of a source performance *vis-à-vis* the performance of the associated digital counterpart?

4.1.3 Aura

The concept of aura in film extends back to Walter Benjamin's critical essay "The Work of Art in the Age of Mechanical Reproduction" from 1935. Since then, nearly all discussions about aura in relation to recorded performance have been shaped by Benjamin's thesis. In his essay, he writes that "[t]he authenticity of a thing is the essence of all that is transmissible from its beginning, ranging from its substantive duration to its testimony to the history which it has experienced." (1969, p. 4). However, when the substantive duration of an object ceases to matter, the historical testimony of the object is displaced, which impacts on the authority of the object. Benjamin describes that which is eliminated in this process as the "aura" of the object, for which he indicates: "that which withers in the age of mechanical reproduction is the aura of the work of art." (1969, p. 4). He is especially critical of the technique of reproduction, which separates an original work of art from its unique existence whenever copies are produced (1969, p. 4). The aura intrinsic to a particular live performance is likewise affected by the recording of that performance. Benjamin notes that while a stage actor presents the performance to the audience using his or her body in person, the recorded version of the performance is presented to a camera. The camera "need not respect the performance as an integral whole", as the camera operator may choose various shot compositions and focal lengths, and the editor may combine, rearrange, or remove different takes in order to produce the final film that is presented to the audience (1969, p. 9). Film eliminates the live dynamic between a performer and the audience, which allows the audience to take a critical stance without regard for personal contact with the actor. For Benjamin, it is in this live dynamic that the actor's aura is experienced, and when the live dynamic is lost, the audience assumes the role of a camera as mere witnesses to actions (1969, p. 10). Benjamin suggests that the actor encountering a camera rather than a live audience results in a similar form of estrangement that one feels when looking at oneself in a mirror, except that the reflected image is no longer tied to its source and is transportable, editable – ready for public consumption at a personally convenient time and place (1969, p. 11).

Almost a century later, motion capture makes this phenomenon even stranger: the actor's double is instead recorded as kinaesthetic body data via a point cloud (a series of points that record the locations of parts of the actor's body in three dimensions over time) or a skeletal representation. Instead of a film actor responding to his or her spectral legacy captured on celluloid, the MoCap actor responds live by interacting with a point cloud that represents his or her movements. Sometimes the point cloud movements are mapped onto a low-resolution version of a digital counterpart for the actor to better envision the intended result in real-time, but the process of motion capture requires the actor to recognise aspects of his or her aura in an unfamiliar double. Bode writes that a digital character who fully replaces an actor's body and face onscreen "gathers its spectacular value from the ways it appears to retain a recognisable connection to the actor" (2015, p. 106). If the actor's aura is to be found in such a PeCap performance, it is not "'beneath' or 'inside' the digital character" but rather in the "kinaesthetic trace, through which [the actor] 'invests' or 'imbues' [the character] with personality, 'life' or 'soul'." (2015, p. 107). Such investment is made through the signs of the actor registered through the digital character, such as facial expressions, voice, gesture, body posture, and body movement (Dyer, as cited in Balcerzak, 2013, p. 205). The actor's aura is only detectable through the physical body, or the simulation of a physical body. If an actor's sincerity and honesty are compromised through the PeCap recording process (either by data noise or lack of detail), then the visible traces necessary to transmit the actor's aura are also compromised.

The multiple filmic incarnations of King Kong provide a variety of examples regarding aura. Balcerzak discusses how the special effects involved in the King Kong films evolved with each instance. In the first "King Kong" (Cooper & Shoedsack, 1933), Willis O'Brien conveyed a distinct sense of pathos and soul through stop-motion animation of the Kong puppet. Kong is a purely simulated creation, featuring no attempts to dress up a human in a gorilla suit to portray the beast, and therefore never hiding a human actor as in future adaptations. Balcerzak writes that

> the original Kong, as an onscreen character, seems to pride itself within its unreality and non-humanity by forgoing the actor completely. It is a cinematic presence, but one distinctly based in the illusion of movement onscreen and not the aura of a body. (2013, p. 208).

This contrasts with the corporeality of Andy Serkis' performance as Kong in Peter Jackson's 2005 adaptation. While both Kongs are constructions fully embedded in the techniques of filmmaking, Balcerzak believes that performance-captured version is imbued with a more complete sense of presence. This is due to an illusion of autonomy Kong received from Serkis' movements and breathing, effectively "humanis[ing] the special effects performance by 'ghosting' the actor as a tangible presence" and "affixing the aura of a body to Kong." (2013, p. 210). This is arguably true, but significant portions of Kong's facial performance and sections of Kong's body performance were created through keyframe animation alone and not necessarily based on any of Serkis' performances (especially in terms of action sequences such as when Kong fights the Tyrannosaurs). This begs the question: is the mere suggestion that a human actor is responsible for the entirety of a performance sufficient for the audience to "detect" an aura at moments when one may not exist?

Jordan Vogt-Roberts' "Kong: Skull Island" (2017) provides an interesting counterpoint to the argument that a human performer must underlie the motions and breathing of a digital character to generate a sense of aura. In stark contrast to the 25-foot tall Kong in Jackson's version, Vogt-Roberts' Kong is a towering 104-foot tall behemoth. ILM animators originally favoured a PeCap approach to generating the gargantuan Kong's movements, but Vogt-Roberts preferred the style of the department's keyframe animation studies. This was due to the massive difference in size between Kong and the human MoCap performers. ILM animation supervisor Scott Benza advises that when MoCap of humansized performers is applied to large creatures, the resulting creature motion feels smaller in scale than desired (i.e.: an issue related to functional believability). As a result, less than ten percent of Kong's movements were motion-captured, with the rest created via keyframe animation. (Edwards, 2017, p. 65). Despite most movements not being based on source-actor input, this Kong has a palpable aura. Vogt-Robert's Kong is produced from copious reference videos of apes and other beasts engaging in a variety of behaviours. The animators meticulously studied, combined, and applied those behaviours to Kong via keyframe animation to produce a range of performances that, at a sufficiently high degree of visual and performance realism, reproduced the auras of the animals on which they were based. These highly skilled animators also created and worked from their own reference performances to serve as a basis for both emotional and physical actions, which were further embellished with established animation techniques to produce a greater sense of scale, nuance, and vitality in motion (2017, p. 65). This example suggests that performance-captured human motion is useful but not necessary in order to imbue a digital character with a sense of aura, especially when the character greatly varies from human scale. This also suggests that aura is not something that must be immediately transferred from a source-performer to a digital counterpart, but rather is something that may be reconstructed if the motion that conveys the original aura is meticulously studied and adapted to the physiology and movement of the digital character. It is even conceiveable that a false-aura may emerge, in which the audience reads a digital performance that combines the motions of multiple source-performers within a single shot as a coherent and original aura – even though such a *gestalt* aura lacks a real-world counterpart and only ever exists within the moment of viewing that sequence. In such a case, aura is not something dependent on a live performing body but is rather a mental construct produced through a visual spectacle so convincing that a viewer believes it was produced by a live performing body.

While Benjamin viewed the actor's aura as inextricably tied to a live performing body, Balcerzak argues that motion capture allows for the performing body to be stripped from its physical presence, leaving behind its aura (2013, p. 210). This implies that the aura is directly retained within the recording process and ready for translation to a digital counterpart (whether the aura is successfully preserved through this translation is dependent on the quality of the performance capture, the flexibility of the character rig, and the skill of the PeCap animators.) Some actors prefer to be recognised for their work rather than their appearance: Lupita Nyong'o was eager to play the fully performance-captured character Maz Kanata in "The Force Awakens" following her role as Patsey in "12 Years a Slave" (McQueen, 2013), in which her naked body was the site of abusive spectacle (Carter, 2015, paras. 3-4). In a physical role, the actor is especially prone to hypersemioticization, which King describes as "the uncontrolled and unstructured incursion of meaning" – that is, any outward signs that the audience may use to construct fictional meaning that is not intended by the director (2011, p. 251). King offers as examples an actor perspiring in a "dead of winter" scene in a stage play, or a wig that is dislodged, or even an actor's sexual attractiveness in contexts where it should not be a factor (2011, p. 251). This latter form of hypersemioticization may have inspired Nyong'o's desire to forgo her own flesh altogether in "The Force Awakens", instead preferring to lose herself in another body and encourage the audience to pay attention to her acting abilities rather than her physical features. Bode writes that PeCap "allows the actor to escape the confines and strictures of their own face and body." (2008, p. 7). Instead of physical qualities such as the eyes, PeCap transmits an actor's "subjectivity, expressivity, and intentionality" through voice, gesture, and muscle groups (2008, p. 7). That is, actors' auras are embedded in the performance of their bodies (e.g.: in kinaesthetic data represented through point clouds or digital skeletal puppets) rather than in their appearances.

4.1.4 Animated Realism

The concept of realism within acting is generally used to reflect emotions, behaviours, and responses that are indistinguishable from our everyday experiences. Giralt refers to this process as "transpos[ing] reality onto the film plane, like one holding a mirror." (2010, p. 14). However, within an animation context, realism also refers to the degree by which a character is physically believable and human-like, as "judged on aspects [of] appearance, movement, behavior, sound, context and the relationship between all of these factors."

(Tinwell, 2015, pp. xiv-xv). If the character is not human, then realism refers to how closely the creature resembles its own kind (or if it is a purely fictional creature, how plausibly it resembles recognisable terrestrial analogues).

Unpredictability and spontaneity are considered key elements in producing a sense of performance realism (Hosea, 2012, p. 127), as if the actors are genuinely surprised and affected by a situation rather than reciting over-memorised dialogue. These qualities are expected of a good live-action performance, but they are nearly antithetical to the frame-by-frame construction of traditional animation. However, animators can approximate this quality if they are able to inject a sense of spontaneity into their reference acting and work this into their animations. Motion capture allows for the clearest blend between live-action spontaneity and animation, so long as the outward signs of a spontaneous response are accurately captured and retargeted onto a digital character. In the production of "Benjamin Button", director David Fincher discusses a Thanksgiving Day dinner scene that features an old-looking Benjamin laughing at a joke. The scene was constructed using a digital version of Benjamin's aged head (played via PeCap by Brad Pitt) attached to a wizened body double (played by Robert Towers). Fincher explains that during the laugh, Towers

> tensed his shoulders, as if he was intaking breath and then exhaling for the laugh. When that movement was in sync with what Brad was doing facially, it worked and looked like a singular performance. But if it was off even by a frame, it looked freakish. It was very important that the facial performance closely follow the body performance. (Duncan, 2009, p. 84).

This example highlights how when working toward realism, all elements of a shot must be integrated with exactitude in order for the illusion to be believable. The closer a human character approximates realism, the more likely that any deviation from our tacit expectations of how a real human looks and behaves will result in an uncanny response.

4.1.5 The Uncanny

Due to our disproportionately large mental capacity for dealing with faces, most digital constructions of human faces result in a viewer response of uncanniness. The concept of the *uncanny* was first psychologically explored by Ernst Jentsch in 1906, who described how lifelike wax dolls and automata can provoke *uncanniness* – a sense of not being able to "distinguish between what is imagined or real, alive or dead." (Tinwell 2015, p. 2). Over a decade later, Sigmund Freud engaged in a psychological investigation of the uncanny. He traced the etymological roots of the word to the German *unheimlich*, "the opposite of *heimlich*, *heimisch*, meaning 'familiar'; 'native'; 'belonging to the home'" (Freud, 2004, p. 76). Therefore, the uncanny is something which is both novel and unfamiliar, and possessing a *sinister* undertone. Tinwell extrapolates Freud's definition to include "everything that should have stayed secret and concealed but has now become visible. […] Uncanniness occur[s] when objects or situations [evoke] a sinister revelation of what is normally concealed from human experience." (2015, p. 4). The English novelist Nicholas Royle has written extensively on the uncanny, and condenses its definition to simply that which is "strangely familiar" (2011, p. vii). To illustrate the concept, he provides examples such as "homeliness uprooted", "curious coincidences", the fear of encountering a missing body part or prosthesis, or in "the sense of a secret encounter" (2011, pp. 1-2).

Within the VFX industry, uncanniness is most often discussed in relationship to the uncanny valley. The concept of the uncanny valley was developed by Japanese roboticist Masahiro Mori, and is the hypothesis that as the degree of human likeness of a robot increases (either in terms of form or motion), so too does its perceived sense of familiarity. However, as the degree of human likeness nears perfection, the perceived sense of familiarity precipitously declines, and familiarity only rebounds once a perfect likeness is achieved. This sudden drop in familiarity is known as the uncanny valley (Figure 11). If the robot is in motion, the effect of the uncanny valley is even more pronounced, including both the degree of positive familiarity and the depth of the uncanny valley. For instance, a corpse, while uncanny, is still less uncanny than a *moving* corpse (Mori, 1970, p. 33). While the uncanny valley was originally developed to describe human responses to robots, it was repurposed to gauge the believability of digital counterparts at the beginning of the 2000s. Since then, the uncanny valley has become a common talking point among animators and VFX artists working toward photorealism (Pollick, 2010, p. 71).

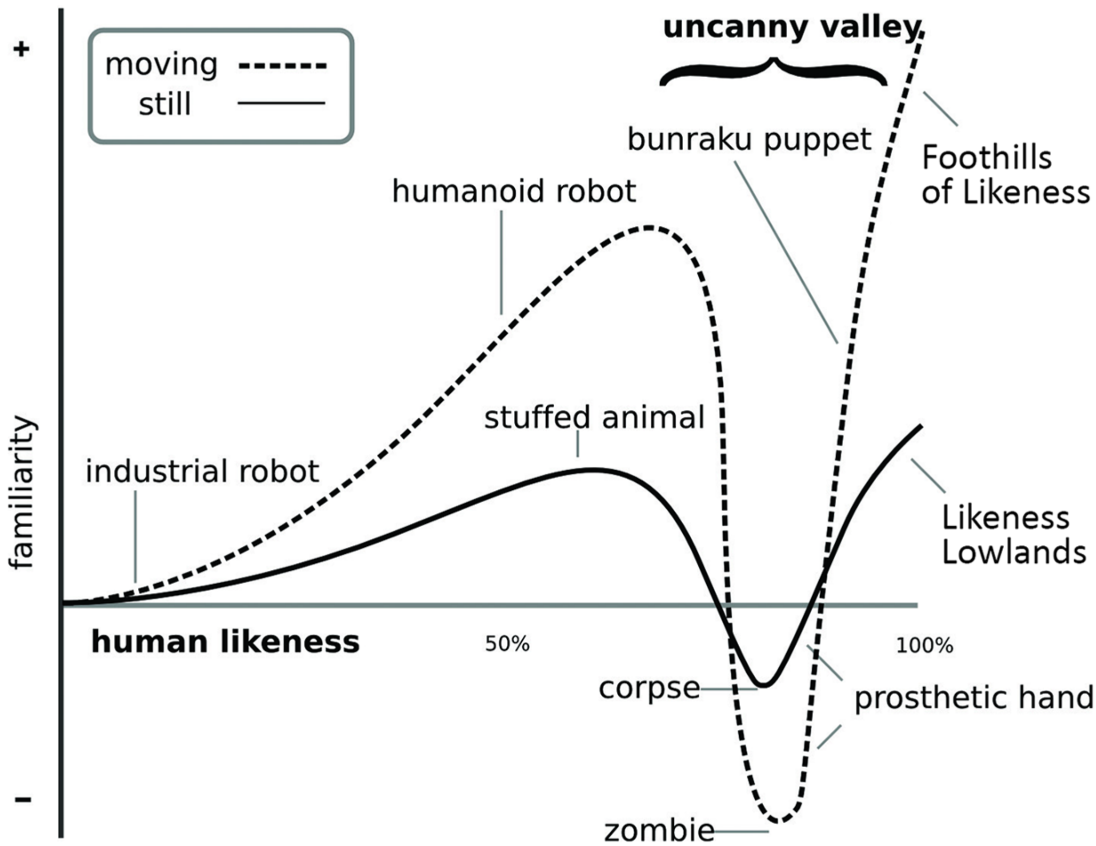


Figure 11. Graph of the uncanny valley.
Reprinted from "The Uncanny Valley", by M. Mori, 1970, *Energy*, 7(4), p. 33.

Uncanniness remains the principal challenge in the quest to achieve photorealistic digital counterparts. This is unsurprising, due to humans' large mental capacity for dealing with facial expressions, which are "conceived as an evolved […] component of emotion."[37] (Izard, 1997, p. 57). Animator Nancy Beiman notes that due to our inherent familiarity with ourselves, humans are the most difficult characters to believably animate. We are subconsciously attuned to any deficiencies in physiology, acting, or movement of an animated human (2010, p. 120), and this becomes more pronounced the greater the stylistic similarity is to a flesh-and-blood human. Many discussions about the trajectory of human uncanniness in feature films start with "Final Fantasy: The Spirits Within". This fully computer-generated film was "one of the first serious attempts at realizing digital human

[37] Izard is careful to point out here that "*observable expression is not viewed as a necessary component of emotion*" (1997, p. 57, original emphasis). However, since this thesis is in part concerned with the way emotions and expressions are read in film, I am principally concerned with the observable component.

characters for film" (Duncan, 2009, p. 83). The motion-captured digital counterparts were the first real test to see how a movie-going public would respond to human-like virtual characters (Taylor, 2000). The film's creators boasted about the realistic quality of its characters, but audiences felt otherwise (Sobchack, 2006, p. 177). The reception plunged deep into the uncanny valley: the characters were human-like but eerily not human enough, "present[ing] us not with people but with indifferent people-shaped objects with plastic skins and blank gazes" (Bode, 2006, p. 175) who "seem all slick surface, manufactured and insincere."(2006, p. 185). The film's characters seem to exist in an "awkward status in-between animation and live action cinema" (2006, p. 176), forcing the viewer to contend with whether the animated body is actually alive (Aldred, 2011, p. 4). This is a further example of unintended incursion of meaning, in which any dip into the uncanny valley results in a hypersemioticization of the digital flesh.

Writing in 2008, North indicates that "[t]he sophistication of a digital face able to convey facial expressions based on instinctive, even subconscious emotional communication rather than broad mimicry is at present far beyond the scope of any animator" (2008, pp. 152-153). Tremendous strides have been made in the ensuing decade, but by and large, digital human faces continue to "create a barrier to identificatory engagement with the audience" (2008, p. 153). However, through improvements in the fidelity of performance capture as well as advancement in 3D modelling, 3D scanning, shader/render development, and the skills of character animators and rigging artists, producing a digital human that surmounts the uncanny valley is within our grasp.[38] For instance, when recreating the likeness of Peter Cushing, who played the character Tarkin in "A New Hope" in 1977, the animation team for "Rogue One: A Star Wars Story" (Edwards, 2016) created the digital Tarkin with controls capable of producing extremely fine levels of detail. These included a deformer control that simulated lip compression and adhesion (digital lips are notoriously difficult to produce accurate interactions due to their wide range of motion and variations in shape), as well as other means for producing tiny imperfections, twitches, and muscle flexes across the

[38] It may be argued that this has already been achieved in the majority of the scenes featuring Junior in "Gemini Man".

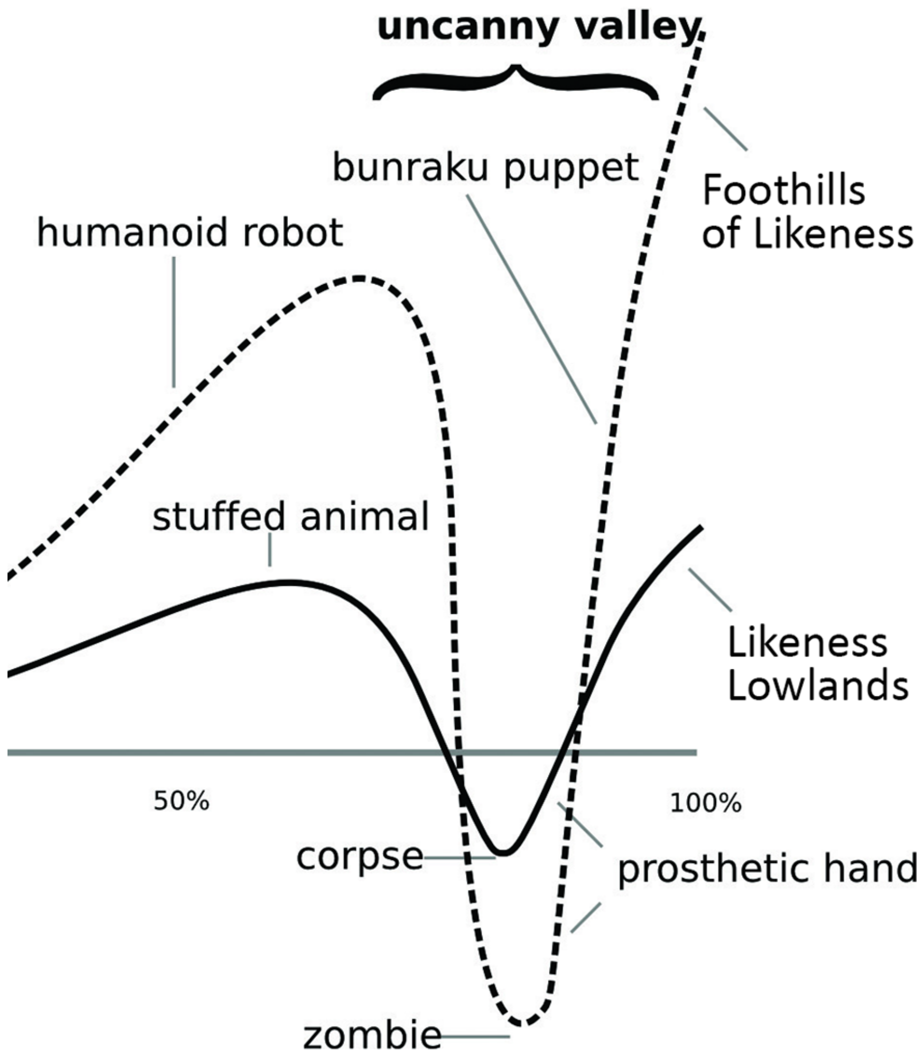


Figure 12. Giacoppo's "Likeness Lowlands" and "Foothills of Likeness" along the uncanny valley.

face. Fordham notes that model supervisor Paul Giacoppo banned his team from referring to the uncanny valley in the production of the film because he believed it was unhelpful. Instead, once the 3D models of the digital counterparts for Tarkin and Princess Leia looked believably human, they referred to that degree of quality as the "Likeness Lowlands". From there, any improvements to realism, whether through modelling, animation, or the simulation of physical movements, approached what they called the "Foothills of Likeness" (Figure 12) (2017, pp. 18-19).

Bode writes extensively about the phenomenon of digital posthumous performance, whereby the likeness of a deceased actor is reconstituted via digital technologies to produce a new mediated performance. A variety of techniques have been used to achieve posthumous performances, including creating CG replicas of actors modelled from film references or 3D scans, as well as digitally compositing a dead actor's face onto the body of a living double (2010, p. 48). However, puppeteering stars from beyond the grave is perceived as deeply uncanny, and is rooted in an intellectual uncertainty about reconciling death with (re-) animation. Our contemporary understanding of acting, which is so often linked to notions of an intentional and coherent subject, is challenged by posthumous performances in which screen actors have no agency in the creation of the acting that bears their likenesses (2010, pp. 59-60). Bode writes that

> [p]osthumous performances remind us of our uncertainty about the degree to which we are organic or artificial, and raise questions about the nature of personhood. Humanistic conceptions of acting and the popular understandings of the Method perhaps work to comfort and, to an extent, ameliorate these uncertainties, allowing us to infer or project a sense of authenticity and intentionality upon the figures we see on the screen. (2010, p. 60).

When producing digitally mediated posthumous performances, filmmakers are presented with a double uncanniness: not only must they achieve believable human physiologies and movements, but in order to overcome reservations about misappropriating a deceased's likeness, they must also be seen to pay respect in terms of both physical form and performance choices. Any deficiency in the physical believability of the character amplifies a ghoulish sense of wrongness in watching the dead perform.

4.2 The Truth Is Written All Over Your Face

Digitally representing human faces is complicated not just by the difficulties in producing accuracy, but also in overcoming a sense of "perfection" that sits at the heart of many standard practices for producing digital characters. For instance, when modelling a character (without scan data), it is common practice to model the character with symmetry enabled – that is, the modeller adjusts only one side of the body while the same changes are mirrored over to the other side of the body. This saves time and ensures consistency in form, but the symmetry causes problems if it is not later attenuated. Ed Catmull advises that “if we make a perfectly symmetrical face, we see it as being wrong. So we want things to be not quite perfect, have a lot of subtlety, but if they’re too imperfect, then we think that they’re strange.” (Tyler, 2000, para. 5). The process of animating a believable face is as much about capturing an overall embodied sense of emotion as it is in defining expressive details. This section looks at three concepts – emotions, expressions, and empathy – frequently used in acting for both live-action and animation, and which are physically related to the construction and articulation of the human face. As in previous sections, these terms are often misunderstood or mischaracterised, leading to a confusion in the language surrounding performance in animation and motion capture. Here I present suggestions for coherent meanings of these terms within the context of acting in animation and motion capture.

4.2.1 Emotions

Emotions are so intrinsic to our experience as humans that the term *emotion* is difficult to unequivocally define. Certain fields are more precise with their definitions of emotion, while others tend toward a pragmatism in the way that emotions are used within the discipline. For instance, within acting, Hagen & Frankel describe an emotion as something that “occurs when something happens to us which momentarily suspends our reasoning control and we are unable to cope with this event logically.” (1973, p. 47). Ed Hooks offers an alternative actor’s perspective and describes an emotion as “[a]n automatic value response.” (2011, p. 13). That is, when a person is presented with a stimulus, an emotion is the feeling elicited in the person based on his or her values about the stimulus. For instance, if you watch another person take a step off the side of a cliff, your reaction may be one of fear or joy depending on your values regarding that person. However, there are shortcomings in both these examples: Hagen & Frankel’s description of an emotion lacks a clear mechanism, while Hooks’ definition is self-referential: the value response is itself directly tied to an emotional state. Although such understandings of emotion may be useful to an actor or animator, they are teleological and lack critical rigour.

Bloch indicates that there is a tendency within the study of emotions for researchers to create definitions that only attend to the external signs of emotional reactions and their physiological symptoms, while ignoring what constitutes the phenomenological experience of "feeling" an emotion. In order to produce a more holistic understanding from a cognitive science perspective, Bloch defines emotions as "distinct and dynamic functional states of the entire organism, comprising particular groups of effector systems (visceral, endocrine, muscular) and particular corresponding subjective states (feelings)" (1993, p. 123). Blair provides clarity to Bloch's definition: that "emotions are physiological states of arousal, not cognitively defined perceptions of feelings", and an "emotional state" is not just a feeling but rather a "carefully articulated physiological response" linked to specific facial expressions, muscular tension, and postures (2008, p. 46). Power describes the process of an emotion as a rapid and unconscious appraisal of a person's immediate environment that results in a physiological response that affects attention, motivation, and action. This is "accompanied by a slower cognitive appraisal that assesses the appropriateness of the quick-and-dirty affective appraisal and that monitors, labels and feeds back into the overall emotional process" (2009, p. 112). The physiological response of an emotion results in a feeling: the conscious registration of the new body-state (Blair, 2008, p. 62). Blair describes a feeling as "emotion made conscious" and indicates that emotions play out through the body while feelings play out through the mind (2008, p. 68). This does not suggest a mind-body split, but rather that emotions and feelings occupy two sides of the same coin, the former being an embodied manifestation while the latter is a psychological interpretation.

Once a person is aware of a particular feeling, *behaviour* is the resulting action he or she makes based on that feeling. Blair indicates that in acting, as in life, behaviour generally moves a person (or character) toward pleasure and away from pain (2007, p. 61). In short, most of us (creatures included) like to feel good and dislike feeling bad, and therefore respond to our feelings with behaviour that maximises pleasure over hurt. Actors often work from a standpoint of engendering within themselves an emotion to elicit a feeling that justifies a particular behaviour required of the character at that moment in the story. However, as we've seen, this approach is primarily embedded in various techniques related to Method acting and does not represent the only technique for producing acting behaviour. In fact, Konijn argues that even when actors believe they are producing authentic emotions for a character, the very knowledge that the acted situation is one of artifice leads them away from producing "emotions in the true sense" – that is, the way the emotions would manifest if the actor believed he or she were in a real situation rather than a factitious one (2000, p. 79). Even at the best of times, actors must heavily rely on their imaginations to perpetuate a belief in the reality of their artificial surroundings. A MoCap environment that lacks most, if not all, of the normal set dressings that facilitate an actor's imagination places an extra strain on producing authentic emotions. However, this is mitigated so long as the actor is able to produce sincere emotional behaviour – that is, an imitation of the physical signifiers associated with a given emotion that the audience interprets as a genuine experience of the emotion. Konijn refers to these as "character-emotions" and states that "[c]haracter-emotions are not real emotions. They are representations of behavior suggesting emotion, used to create the illusion of the real emotion we know in daily life in a convincing and believable way." (2000, p. 80). However, as previously mentioned, it is arguable that sincere emotions may become more nuanced or convincing if they are accompanied by authentic emotions – that is, if the actor experiences connected emotions.

4.2.1.1 Basic Emotions

While emotions span a wide range of psychophysiological responses, it is widely accepted that any emotional experience can be linked to one or more "basic emotions". Basic emotions are generally understood as pan-cultural[39] discrete emotions from which all other emotions are "subcategories or mixtures (patterns, blends, combinations) of the basic emotions" (Russell & Fernández-Dols, 1997, p. 12). Bloch et al. believe that a defining characteristic of basic emotions is their presence in infants, "either as innate behaviours or

[39] Since the development of Ekman's theory of basic emotions in the 1960s, it has been widely accepted that basic emotions are universally recognised across all cultures. However, more recent research challenges such universal recognition and indicates that the interpretation of certain facial expressions may be culturally based. (This does not undermine the existence of basic emotions, *per sé*, but rather challenges the idea that basic emotions are identically expressed across all cultures.) This is supported by research into indigenous cultures that are mostly unfamiliar with cultures outside their own. (Crivelli, Jarillo, Russell, & Fernández-Dols, 2016).

apparent at very early stage of post-natal development." (1987, p. 4). However, the exact number and type of basic emotions has been the subject of debate since at least 1930. Some researchers propose as few as two basic emotions while others include up to 18 emotions (Ortony & Turner, 1990, p. 315). Paul Ekman, whose research in nonverbal behaviour and universal (pan-cultural) emotions helped to revolutionise psychology in the late 20th Century (PaulEkmanGroup, n.d.), attributes this confusion to a lack of proper organisation of emotions into families as well as a failure to distinguish emotions from moods, attitudes, disorders, and other affective phenomena (Ekman, 1992, p. 174). Ekman identifies a set of seven basic emotions recognised by all cultures, including anger, fear, sadness, enjoyment, disgust, surprise, and contempt (1986, p. 160). He promotes that all other emotions are "subcategories or mixtures (patterns, blends, combinations) of the basic emotions. For example, anger includes fury and annoyance as subcategories [...]. Anxiety is a mixture of fear, sadness, anger, shame, and interest" (Russell & Fernández-Dols, 1997, p. 12). While a certain emotional category may show an affinity to a given facial expression (e.g.: sadness corresponding to crying), this relationship is no more than an affinity. For example, crying is not distinct to sadness, and may occur during moments of happiness, while it is possible that during sadness, no expression may be present (Frijda & Tcherkassof, 1997, pp. 80-81).

Bloch et al. provide a list of six basic emotions from their performance-based research. These include happiness, sadness, fear, anger, eroticism, and tenderness. While these emotions are similar to those recognised by Ekman, they are not based on universal recognisability but rather on their utility for actors. They also provide examples of emotions derived from each of these basic emotions: "*happiness* (laughter, pleasure, joy); *sadness* (crying, sorrow, grief, depression); *fear* (anxiety, panic); *anger* (aggression, attack, hate); *eroticism* (sex, sensuality, lust); *tenderness* (filial love, maternal/paternal love, friendship)." (1987, p. 4). Bloch et al. plot these basic emotions along two axes representing continuums of relaxation-tension and avoidance-approach (Figure 13). Relaxation-tension refers to the degree to which muscles are activated throughout the body and face of a person experiencing a given emotion, and avoidance-approach refers to the inclination a person for moving toward or away from the stimulus of a given emotion (1987, p. 7).

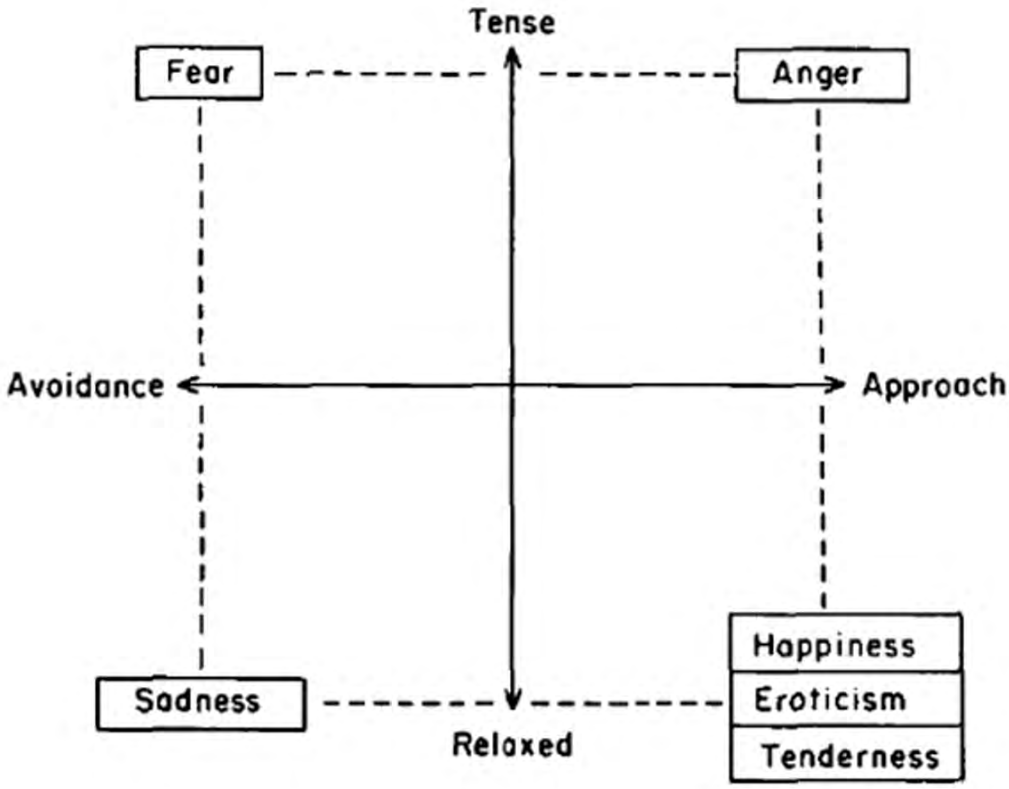


Figure 13. Representation of the six basic emotions of Bloch et al. in terms of muscle tension/relaxation and approach/avoidance parameters. Reprinted from "Effector patterns of basic emotions: a psychophysiological method for training actors", by S. Bloch, P. Orthous, & G. Santibañez-H, 1987, *Journal of Social and Biological Structures*, 10, pp. 1-19.

Bloch et al. associate an emotional reaction of fear with postural tension and avoidance of the stimulus, and sadness with relaxed muscles while still maintaining avoidance. By comparison, they associate anger with postural tension and an approach toward the stimulus, and the trifecta of happiness/eroticism/tenderness with relaxed muscles in concert with approach. While this chart may suggest an equivalence between happiness/eroticism/ tenderness, each of these emotions is characterised by a specific manner of breathing (Figure 14). For instance, happiness involves a saccadic expiration (sporadic exhalation) with the mouth open. By comparison, both eroticism and tenderness involve slow respiration that is low in force (as measured by a pneumatograph). However, while eroticism entails breathing through an open mouth, tenderness is performed through a closed mouth in a relaxed smile.

In contrast to Bloch, my acting training with the Kacie Stetson Studio focuses on only four basic emotions: love, hurt, anger, and fear. Hurt (sadness), anger, and fear are common

Emotion	Posture	Direction	Manner of Breathing
Happiness	Relaxed	Approach	Saccadic expiration (mouth open)
Sadness	Relaxed	Avoidance	Saccadic inspiration (mouth open)
Fear	Tense	Avoidance	Inspiratory apnea (mouth open)
Anger	Tense	Approach	Hyperventilation (mouth closed tight)
Eroticism	Relaxed	Approach	Small amplitude, low frequency (mouth open)
Tenderness	Relaxed	Approach	Small amplitude, low frequency (mouth closed in a relaxed smile)

Figure 14. Schematic representation of posture for the basic emotions based as described by Bloch et al. Reprinted from "Effector patterns of basic emotions: a psychophysiological method for training actors", by S. Bloch, P. Orthous, & G. Santibañez-H, 1987, *Journal of Social and Biological Structures*, 10, pp. 1-19.

between these two systems, and Stetson's understanding of love encompasses the full emotional range of happiness/eroticism/tenderness. However, while Stetson advises that active breathing is essential to good emotional activation, her teaching does not emphasise the conscious articulation of specific breathing patterns. I draw from Stetson's four basic emotions when I act in the practical components of this research – which means that I did not consciously employ specific breathing patterns.

Bloch identifies basic emotions as "phasic" reactions, which are short-lived responses related to the psychosomatic state of a body in a given moment. Most of these responses results in some kind of action, such as laughter, crying, attacking, flirting, etc. (1993, p. 125). This corresponds to Ekman's observation that emotions typically last for only a matter of seconds rather than minutes or hours. His research suggests that when someone reports experiencing an emotion for an extended period of time (i.e. 15 or 20 minutes), the person has actually experienced repeated but discrete emotion episodes that he or she rationalised as a continuous experience. When such symptoms persist for hours or days, they are no longer emotions but rather moods: highly saturated experiences of a given emotion or set of emotions. (1992, p. 186). Moods are not necessarily tied to a particular stimulus and include anxiety (a chronic state of fear) and depression (a chronic state of sadness) (Bloch, 1993, p. 125).

In live-action as with animation, the audience expects a character to exhibit sincere signs that equate to recognisable emotional effector patterns. The patterns of respiration identified in Bloch et al.'s research will assist animators when adding subtle breathing details to realistic characters. This could be in terms of ensuring that the breathing details were accurately captured via PeCap (as in the case of Jackson's King Kong), or in terms of creating the breathing details from scratch (as in Vogt-Roberts' Kong). Likewise, it is important for actors and animators to understand the phasic nature of emotions and that an actor (or animated character) can't play a mood. If a character is to experience a mood, the animator (and the actor, in the case of PeCap) must engage the character with stimuli that continuously trigger emotional reactions to perpetuate the mood.

4.2.2 Expressions

Ekman indicates that "[facial expressions] are a sign that an emotion is occurring. [… A] hallmark of an emotion is that it has a signal, in face and/or voice and/or bodily movement." (Ekman, 2004, p. 45). Russell & Fernández-Dols elaborate that a facial expression is "a universally understood signal, a visible outcropping of an otherwise hidden event, the triggering of a discrete categorical 'basic' emotion." (1997, p. 4). However, while facial expressions are understood to be evolved and genetically influenced components of emotion, Izard cautions that "*observable expression is not viewed as a necessary component of emotion.*" (1997, p. 57, original emphasis). He argues that there is no single "full-face 'prototypical' expression for each of [the basic emotions]" (1997, p. 60). That is, a facial expression indicates that a basic emotion is triggered, but a given expression does not necessarily indicate which basic emotion is triggered (i.e. someone who smiles when nervous or who cries when happy). While certain emotions have an affinity toward particular expressions, this is not a dependable relationship. Russell & Fernández-Dols suggest that the one definitive thing a facial expression reveals is that an emotion is triggered. However, the Alba Technique suggests that an expression can be recognised as the sign of an emotion without the signaller actually experiencing the emotion. Likewise, it is possible for an emotion to be "unobservable" due to a lack of expression. This means that while generally useful as measures of a person's emotional state, expressions are equivocal barometers that require an understanding of a person's history, context, and culture in order to reliably appraise an emotion.

Most people will be familiar with the phenomenon of witnessing someone express a *false* smile. Such smiles are not necessarily malicious in intent and can occur when someone is nervous and feels the need to produce an expression that conforms to social expectations. The French physiologist Guillaume-Benjamin-Amand Duchenne was the first to describe the difference between genuine and false smiles. In both types of smiles, the *zygomaticus major* muscles (which extend from the corners of the mouth to the cheekbones on either side of the face) pull the corners of the mouth upward, resulting in a smile shape. However, the defining feature of a genuine (that is, *authentic*) smile is the additional activation of the *orbicularis oculi* muscle responsible for raising the cheek and narrowing the eye aperture, pulling skin inward around the eyes and often resulting in crow's feet wrinkles among adults (Tinwell, 2015, pp. 91-92). Ekman notes that the majority of people are unable to voluntarily contract the orbicularis oculi muscle, and its activation is generally triggered through a spontaneous reaction, rather than a conscious one. In honour of Duchenne, Ekman terms smiles that incorporate both the *zygomatic major* and *orbicularis oculi* muscles as *Duchenne smiles*, and only Duchenne smiles are correlated with the experience of happiness (1989, pp. 155-156).

Attention to expression detail plays a significant role in how an audience perceives realism and empathises with a digital character. "Galaxy Chauffeur" (Kleinman, 2013), an advertisement for Galaxy chocolate, features a digital posthumous performance by a synthespian double of Audrey Hepburn (circa 1961). The digital Hepburn takes the role of a woman onboard a broken-down bus in Italy, and is rescued by a luxury chauffeur while she savours a bar of Galaxy chocolate. However, as Tinwell points out, Hepburn's CG face has impoverished expressivity, including a lack of perceptible nonverbal communication in her upper face compared to her wide smile in her lower face (Figure 15). There is missing perceptible activation of the *orbicularis oculi* muscles, and therefore no sense of crow's feet

Figure 15. A CG Audrey Hepburn in an ad for Galaxy chocolate, eliciting an unintentionally insincere smile. From *Galaxy Chauffeur*, Kleinman, 2013.

or skin bulges around the eyes, which would be associated with an authentic experience of happiness. This mismatch in expression could be interpreted as Hepburn hiding negative emotions – that she may not be as comfortable around her debonair chauffeur as she would have him believe. Likewise, as a response to eating a Galaxy bar, a false smile may suggest that the chocolate is not truly enjoyable (2015, pp. 93-94). For an advertisement that is meant to sell the wonderful experience of its product, the lead actress' performance produces too many vagaries that muddle the message.

It is reasonable to speculate why Hepburn's performance turned out the way it did. PeCap was not utilised in this production; instead, an actress who closely resembles Hepburn was filmed from multiple angles in order to provide comprehensive reference performance. No digital scans of Hepburn exist, so VFX studio Framestore developed a CG Hepburn head model based on still images from the actresses' oeuvre. Hepburn's human double was digitally scanned to produce an additional 70 facial shapes corresponding to muscle movements, which were adapted to Hepburn's physiology and incorporated into the Hepburn rig. However, despite looking similar to Hepburn, the actress double's facial expressions resulted in different nuances, meaning that animators had to refer to footage of Hepburn in order to accurately convey her mannerisms (Bartlett, 2013, paras. 1-7). A team of four animators worked to recreate Hepburn's signature smile through keyframe animation. VFX supervisor Simon French explains: "It is amazing how unique and how recognisable a person's smile is. When you see it in this detail, it really needs to look perfect." (Bartlett, 2013, para. 8).

Hepburn's lower face smile is accurately conveyed through her digital counterpart, but it is the smile detail in the upper face that is missing. This can be due to a number of factors, both technical and aesthetic. Historically, CG faces have been limited in their expressive capabilities based on the number of polygons used to construct a face. This was due to limitations on computing resources created by both hardware and software bottlenecks, and the result was often an insufficient amount of detail around the eye and mouth areas to precisely simulate human expression, especially when the character transitioned between expressions. The low polygon resolutions prevented the representation of believable bags and wrinkles around the eyes, which exacerbates the sense of incongruence between the upper and lower face (Tinwell, 2015, p. 92). However, by the time the "Galaxy Chauffeur" ad was produced in 2013, improvements in computing power provided increases in the resolutions used for realistic 3D characters.[40] Tinwell speculates that a problem in the source performance could be the issue: if the actress failed to convey sincere happiness and instead produced a contrived emotion, then any animator working to preserve that performance was unwittingly imbuing the digital counterpart with false emotion. Even more likely is that a reverence for Hepburn's iconic beauty may have meant that the producers were

[40] Even if the production was limited to a lower polygon count, they still would have been able to convey fine-level details through a variety of image maps, including normal maps, which add extra levels of detail without increasing the mesh resolution.

hesitant to display her with any age-related imperfections. Tinwell writes that

> [i]f designers are modeling a female protagonist character who is supposed to represent a beautiful heroine in a game or animation, then facial cues such as bulges and crow's feet wrinkles may be regarded as unattractive, unnecessary flaws rather than important and crucial signals. The designer may be reluctant to include such details in case they represent a more ugly appearance and imperfections in her flawless skin. A designer may even intentionally erase these less attractive wrinkles and folds in the skin to retain the character's perfect appearance. (2015, pp. 94-95).

Thus, a persistent form of sexism is also played out through computer-generated representations of women. We may imagine how such *digital Botox* (Variety Staff, 2008, para. 16) could be applied to a range of female protagonists, from de-aging extant actresses to producing more "perfect" depictions of deceased starlets[41]. However, just as with the CG Hepburn, wrinkles and bulges in the face provide crucial details in conveying sincere expressions. When those details are missing, the resulting false expressions may be interpreted as deceitful or even sociopathic. Attempts to improve attractiveness in this manner result in less empathetic characters who sink deeper into the uncanny valley (Tinwell, 2015, p. 95). Pallant indicates that the moments of PeCap that feel most jarring are those when the technology results in "an inability to accurately convey a character's emotional response." (2012, p. 39). While it is generally advisable to avoid mismatches in expressive signals, these lessons can be strategically deployed when designing characters that are intended to be understood as evil or uncanny (Tinwell, 2015, p. 117) – such as the criminal Harlan Fontaine in the video game "L.A. Noire" (McNamara, 2011).

Other features that may be missing from the CG Hepburn's performance are micro expressions – rapid facial expressions that last between 1/25 and 1/15 of a second. By

---

[41] One may speculate on how the bodies of other beauty icons such as Marilyn Monroe would be altered to conform to contemporary beauty standards, whether through digitally "thinning" her figure, reducing visible "peach fuzz" on her face, or through a similar application of *digital Botox* to remove any wrinkles.

comparison, macro (normal) expressions generally last between ½ a second to 4 seconds, often repeat, and likely correspond to what is said (Ekman, 2016, para. 4). Ekman explains that micro expressions occur "when people have hidden their feelings from themselves (repression) or when they deliberately try to conceal their feelings from others", both of which are visually indistinguishable from each other (2016, para. 2). A micro expression conveys how the person actually feels in response to a stimulus, even if this reaction is below conscious awareness. When creating a realistic CG character, there are many ways in which micro expressions could be unwittingly erased. As discussed above, lower resolution facial geometry may not be able to capture the nuances of a micro expression, especially if those nuances subtly vary from the expressions immediately preceding and following it. Also, *digital Botox* will similarly eradicate micro expressions. Another possibility has to do with the frame rate of the reference performance and the final animation. The quickest micro expression (1/25 of a second) occurs faster than the standard frame rate of film (1/24 of a second). This means that in a performance recording at 24fps, a micro expression may appear on only a single frame (or at most two frames, if the micro expression lasts for as long as 1/15 of a second). Such single or double frame variance in either a video file or PeCap data could easily be dismissed as noise by an animator or eliminated via signal filters in the software. Therefore, a knowledge of the psychology and physiology of human expressions is essential for an animator working with realistic performances.

The brows play a significant role in understanding the nature of expression across the rest of the face. The brows are the primary conveyors of nonverbal signals throughout the face, which help to provide context to the primarily articulatory processes of the lower face region. For instance, lowering a brow is associated with emotions such as fear, sadness, and anger, whereas a raised brow is likely associated with happiness or surprise. The brows are also used as "batons" to emphasise particular words or phrases, such as lowering the brow to accentuate a word with negative intent, or raising the brow to accentuate a word with positive intent (Tinwell, Grimshaw, Nabi, & Williams, 2011, pp. 742-743). The degree to which the brows squeeze together directly corresponds to a person's intensity of thought (Osipa, 2010, p. 23), while a person squinting the muscles around the eyes amplifies the intensity of an expression (2010, p. 27). The upper face region provides essential information about the affective state of a character, but this region has been commonly overlooked in many digital characters, leading to a strong uncanny response (Tinwell et al., 2011, p. 743). Busso & Narayanan suggest that the upper face region is the primary region to detect visual prominences and as a result, this area should be given special attention when creating a digital character in order to produce more realistic emotional representations (2006, p. 555).

The part of Ekman's research that has had the greatest impact on computer graphics and animation is the development of the Facial Action Coding System (FACS) (Ekman, 1979). Psychology professors Gosselin, Kirouac, and Doré describe how FACS taxonomises each discrete muscle movement in the face and encodes an action unit (AU) to it, distinguishing a total of 44 AUs, which are "the minimum number of action units that are anatomically separate and visually distinguishable." (2005, p. 249.). Examples of AUs include the inner brow raiser, brow lowerer, cheek raiser, nasolabial deepener, and lip funneler (2005, p. 250). For instance, the activation of the *orbicularis oculi* muscle is encoded as AU6, and the activation of the *zygomaticus major* muscle is encoded as AU12. When AU6 is combined with AU12 (denoted AU6 + AU12), the resulting expression has a high affinity toward being interpreted as happiness. All emotional expressions can be configured in similar ways (Ekman, 2004, pp. 46-47), although it is important to restate that facial expressions do not provide a fully reliable basis to determine a person's experience of a particular emotion. Both FACS and the effector patterns of basic emotions (EPBE) identified by Bloch et al. follow a taxonomic structure based on visually identifiable signals of a human subject. However, while FACS focuses on the physical relationship of discrete sections of the face, it does so based on the observation of still images. EPBE consider breathing patterns, which are likely important to signalling an emotion in a temporal medium rather than in just a static image. I suggest that an enhanced application of the FACS system to animation should factor in the emotion-specific manners of breathing identified in EPBE.

FACS is useful for 3D modeling and animation because it dissects a complex system into its component parts and allows the system to be rebuilt one unit at a time. For 3D modelers, this means that each action unit can be modeled separately to be combined by animators later. FACS allows 3D animators to "dial up" or "dial down" a specific action unit, and thereby control the nuance of any given expression. Likewise, this system for separating expressions into component parts is helpful for animators to understand how to read emotions and clearly see how the face moves. For these reasons, FACS is employed throughout the animation industry as a basis for modeling and animating faces, including at top animation companies such as Pixar and DreamWorks (Hampton, 2014, para. 2).

Returning to Catmull's caveat regarding symmetry in 3D characters, Gosselin et al. acknowledge that facial expressions of simulated emotions manifest with greater asymmetry than connected emotions. Expressions of simulated emotions are characterised by more irregular muscle contractions and often with missing components that are normally associated with a particular emotion (e.g. the missing contraction of the *orbicularis oculi* muscle in a happy smile) (2005, p. 243). Joe Navarro, an ex-FBI agent who served on the Nation Security Division's Behavioral Analysis Program (Psychology Today, 2020, para. 1), describes how asymmetric expressions can often betray our true sentiments when the left side and right side of our face suggests a different emotion. An emotionally connected happy smile will produce what Navarro's refers to as "emotional symmetry", in which the expression signifies happiness in equal measure on each side of the face, regardless of any natural asymmetry in the face's physiology. However, when a person simultaneously experiences multiple emotions, this disrupts the face's expressive symmetry (Navarro, 2016, paras. 2-3). Navarro describes such emotional asymmetry as "facial chirality", where *chirality* denotes "two objects that may appear identical but when folded over onto themselves, they are not symmetrical." (2016, para. 4). When a person suppresses an emotion, whether through conscious guile or subconscious ignorance, the expression may display on only one side of the face, while the other side may express something else completely different. Chiral displays are more likely to be subconsciously perceived, signalling to the viewer that something is wrong without the conscious awareness registering the precise reason why (2016, paras. 5-6). Navarro provides a mugshot that Jesus Oliveira posted to his own Twitter account while pleading his innocence on a charge of abuse (Figure 16). Navarro notes that one side of Oliveira's face looks different from the other half, as if one side expresses the desire to be believed while the other side expresses fear of being found out. The chirality present in Oliveira's expression is even more obvious when the image is split along the midline (Figure 17) (2016, paras. 8-9)

Animators are often encouraged to introduce asymmetry into poses and expressions, but an asymmetric expression endangers a digital counterpart's perceived sincerity and emotional connection. While the push toward asymmetry may work better in a graphic medium (such as in more stylised approaches toward character animation), animators should be aware of how facial chirality affects an audience's empathic response to characters embedded in a realistic style, especially when their performances are produced via PeCap. However, if a scene calls for a vactor who is struggling with multiple emotions, a director may wish to have animators introduce some chirality into the facial performance if it is not present in the source-actor's performance.

Just as with casting the right actor for a role, we may consider how animators are "cast" to handle various performance types, from big action to soft timing and subtle expressions. PeCap animators may be likewise "cast", depending on their skills at detecting, transferring, and augmenting certain types of performance. Strasberg criticises the actor who comes to play a character as if it were him/herself, "without trying to find out the difference between himself and the character" (Hethmon, 1965, p. 104). By contrast, Strasberg notes that when an actor is asked to play an animal, the actor never assumes to play it as him/herself, and instead is "forced to use observation to single out the qualities of a particular animal, and to find how he, the actor, can accomplish them." (1965, p. 104). Both the live actor and the animator-as-actor must look for traits that contrast with their natural manner:

> [a]n actor who is tough and robust and earthy does not work to be tough and robust and earthy in a scene that calls for a person with those colors. He works for sensitivity, for things other than his natural roughness and toughness, which are already there. But if an actor playing that part is not the rough and tough type of person, that is what he works to get. To the actor who already has sensitivity, sensitivity is no problem (1965, pp. 146-147).

Figure 16. Mugshot of Jesus Oliveira posted to his Twitter account. Facial chirality suggests reasons to doubt his proclaimed innocence. Reprinted from "Chirality: a look at emotional asymmetry of the face", by J. Navarro, 2016, https://www.psychologytoday.com/nz/blog/spycatcher/201605/chirality-look-emotional-asymmetry-the-face

Figure 17. By splitting the face into two distinct halves, the distinct expression on each side is more obvious. The left side appears to plead for trust while the right side appears apprehensive. Reprinted from "Chirality: a look at emotional asymmetry of the face", by J. Navarro, 2016, https://www.psychologytoday.com/nz/blog/spycatcher/201605/chirality-look-emotional-asymmetry-the-face

We may ask if the same principle holds true for animators, especially when they either produce their own acting reference or when they "read" the performance of a source-actor for PeCap. While further research would be necessary to confirm this assumption, my experience of teaching animators suggests that certain animators are (either innately or through training) better able to read and understand varying degrees of expression. It is worth investigating whether animators who are more experienced with expressing particular emotions are better able to read and transfer expressions linked to those same emotions within animation and PeCap. Strasberg describes how some actors are naturally inexpressive, which he attributes to "dead" imaginations due to psychological conditioning or lack of expressive development (Hethmon, 1965, p. 210). I tie this observation to a story I learned during a presentation at the Society for Animation Studies Conference in 2016 from Hillary Yeo, an animation lecturer at LASALLE College of the Arts in Singapore. Yeo noted that the majority of his students are Southeast Asian and are often insular, without much outward displays of emotion. He specifically referred to a paper he taught that involved FACS-based blend shapes for animation. According to Yeo, at the end of the semester, a female student who was not in his paper came to thank him. She was the girlfriend of one of the male students in his class, and she said that his paper saved their relationship. Apparently, her boyfriend was initially poor at reading her expressions and understanding her emotions, but became markedly better at this process as a result of learning FACS-based animation. This anecdote suggests that with the right training, animators can better recognise expressions and emotions, and can presumably use this skill to more accurately convey similar expressions through their animated characters. This is consistent with research that suggests acting training is associated with social understanding, including "higher levels of theory of mind, empathy, and adaptive emotion regulation" (Goldstein & Levy, 2017, p. 156).

4.2.3 Empathy

With the ability to read and understand facial expressions comes the opportunity to empathise with perceived emotions. As previously discussed, it is possible that when feeling an emotion, a person may produce no change in expression or no expression at all. An observer watching someone produce a facial expression *interprets* the expression as a particular emotion, but that is not necessarily the emotion the producer feels.[42] However, there is generally a significant likelihood that a particular emotion is highly correlated with a given expression, and the process of interpreting an expression as an emotion is the basis for *empathy*.

Preston & de Waal identify the complex history of defining empathy, and note that even today "the field suffers from a lack of consensus regarding the nature of the phenomenon" (2002, p. 1) due to a lack of a proximate mechanism. Empathy is a relatively recent concept in human affairs, dating back to 1873 when Robert Vischer coined the German word *Einfühlung*, which literally means "feeling into" (or "in-feeling") to describe "the projection of human feeling onto the natural world" (Blair, 2009, p. 98 [citing Pigman, 1995]). This concept was later expanded by the philosopher Theodore Lipps, who used it to describe the "psychology of intersubjectivity in terms of '*inner imitation* of the perceived movements of others', as a step toward understanding that other people have selves." (2009, p. 98, *original emphasis* [citing Gallese, 2001]). In 1909, psychologist Edward Titchener translated *Einfühlung* into English as *empathy*, derived from the Greek "to make suffer" (2009, p. 98). Catrin Misselhorn, a philosopher of science and technology at the University of Göttingen, notes that empathy is commonly understood as the phenomenon of "putting oneself in someone else's shoes", where a person is able to feel another person's emotion within oneself. By contrast, Misselhorn describes *sympathy* as "an emotional reaction to the situation of others which does not involve experiencing the emotions they feel." (2009, p. 351). This is in line with Hooks' distinction between sympathy and empathy, which he notes are frequently and erroneously used interchangeably. For Hooks, sympathy is "feeling for", whereas empathy is "feeling into" (2011, p. 15). The former provokes pity, whereas the latter is a form of identification with another person (2011, pp. 15-16).

---

[42] Anyone who has ever been asked why they look angry when they actually feel no anger can relate to this. In fact, this particular situation is so common that it has entered the popular culture lexicon as "resting bitch face".

However, the actual mechanism by which empathy is performed is hotly contested. In providing a rigorous philosophical definition for empathy, Sober and Wilson propose:

> *S* empathizes with *O*'s experience of emotion *E* if and only if *O* feels *E*, *S* believes that *O* feels *E*, and this causes *S* to feel *E* for *O*. (Misselhorn, 2009, p. 351 [citing Sober & Wilson, 1998]).

However, as Misselhorn points out, this definition is restricted to situations where an object *O* (the primary individual experiencing an emotion) feels an emotion *E*, which excludes any situation where a subject *S* (an individual who secondarily experiences the emotional state of *O*) may feel empathy for an *O* that either has no feelings or experiences a feeling that is different from what is perceived (i.e. a masked emotion) (Misselhorn, 2009, p. 351; Preston & de Waal, 2002, p. 4). Likewise, this definition precludes any empathetic reactions toward animated characters, which convey emotions without actually experiencing them. Obviously, since people report feeling genuine empathetic connections with animated characters, a definition of empathy should not be dependent on whether the object in question authentically experiences an emotion. As an alternative to Sober & Wilson's definition, psychology professor Heidi Maibom proposes:

> *S* empathizes with *O*'s experience of emotion *E* if *S* perceives *O*'s *T*-ing and this perception causes *S* to feel *E* for *O*. (Maibom, 2007, p. 168).

Maibom explains that "*T* stands for some behavior or expression that is reliably correlated with an emotion *E*, e.g. crying (reliably correlated with sadness) or laughing (reliably correlated with joy)." (2007, p. 168). This definition relies on the subject's *perception* of an object's emotional experience, rather than on the object's *actual experience* of an emotion. However, as with inanimate objects, we know that animated characters don't actually experience emotions – nor do they experience anything at all. Misselhorn proposes a modification to Maibom's definition that is inclusive of our empathic experience with inanimate objects, including within fictional contexts such as with animated characters:

> *S* empathizes with an inanimate object's imagined experience of emotion *E* if *S* imaginatively perceives the inanimate object's *T*-ing and this imaginative perception causes *S* to feel *E* for the inanimate object. (2009, p. 352).

With this definition, we have a clear psychological process for how a viewer may empathise with an animated character: if an animated character provides sufficient visual and auditory signals to suggest the character is experiencing an emotional state, then the viewer's imaginative perception of the character's emotional state provokes a similar emotional state in the viewer.

Preston & de Waal provide a Perception-Action Model of empathy that provides some psychophysiological basis for the experience of empathy. This model states that "*attended perception of the object's state automatically activates the subject's representations of the state, situation, and object, and that activation of these representations automatically primes or generates the associated autonomic and somatic responses, unless inhibited*" (2002, p. 4, *original emphasis*). Effectively, viewing someone experiencing an emotion generates the same physiological and neurological patterns in the viewer, resulting in the viewer experiencing the same emotion. The recent discovery of mirror neurons has led to an understanding of a more concrete physical mechanism for empathy. Mirror neurons in the brain activate whenever a person takes action, but also whenever a person watches another person's actions. That is, when observing another's action, an identical physiological response occurs in the observer's brain as if the observer were taking the same action. Bishko explains that this is "the mechanism through which we understand what we see", and empathy is "the physiological mechanism for how we perceive others, learn through imitation, develop language, and communicate" (2014, p. 50). Mirror neurons provide the basis for understanding our own actions and, importantly, they provide the basis for understanding the intentions that underlie the actions of others (Power, 2008, p. 32). This leads to questioning how reliable is the correlation between emotions and expressions? If expressions consistently invoke psychophysiological reactions that result in false empathy, this would seem to be of little evolutionary value. The existence of mirror neurons may suggest that the ability to accurately determine a subject's emotions from their expressions (or their emotional effector patterns) is most often reliable.

Another form of empathy may be encountered during the observation of movement. Carol-

Lynne Moore describes such *empathic attunement* in terms of the body's response while watching an intense sports or thriller programme. Fans may find themselves moving like the people they observe on screen, albeit in more limited ways, based on imitation and movement memory (Bishko, 2014, p. 50 [citing Moore, 1988]). Empathic attunement may play out in a unique way for animators during the act of animating. When referring to a reference performance, an animator becomes aware of the forces present throughout the body as a whole, as well as those present in more localised regions, such as the thighs and calves. Animators often rely on a type of empathic attunement to *feel* those forces *through* their bodies in order to translate a similar sense of force into their animated poses. Drawing on my own animation experience, I often try to physically recreate (to the best of my ability) most of my reference poses and actions with my body just to gain a more embodied sense of force and motion, and I refer to this phenomenological process as *listening to your body*. Good animators will listen to the bodies of their reference performances (via empathic attunement) while also listening to their own bodies in re-performing those same actions. This process is especially helpful when deciding how the weight in a pose should be balanced, which part of the body leads a particular action, and where the character would experience uncomfortable stresses in a pose. For instance, in my experience it is common for student animators to keep a character's feet planted on the ground when performing twists throughout the rest of the body. This doesn't make physical sense because a person would normally adjust the position of his or her feet in order to accommodate the twisted pose. Whenever I encounter this problem, I ask the students to stand up and act out the pose transition so that they kinaesthetically experience the problems in their poses, and to encourage them to frequently employ this method while animating.

4.3 The Screen Presence of Digital Performers

So far, I've examined several terms common to the language of acting – including *authenticity*, *truthfulness*, *aura*, *emotion*, *expression*, and *empathy* – for which standardised definitions are difficult to explicate. I now add *presence* to this list, and specifically how the concept is considered in terms of *screen presence*. Presence is subject to several interpretations, including the fact or condition of "being, existing, or occurring at this time or now" (Merriam-Webster, 2020). An actor may be considered *present* either on stage or in front of a camera by merely existing at a specific location and time. However, the act of existing is not in itself sufficient to produce the profound charisma generally associated with captivating actors (Kennedy, 2019b, p. 191). The playwright Joseph Chaikin defines *presence* in terms of the actor's craft, specifying that

> [it's] a quality that makes you feel as though you're standing right next to the actor, no matter where you're sitting in the theater. […] There may be nothing of this quality off stage or in any other circumstance in the life of such an actor. It's a kind of deep libidinal surrender which the performer reserves for his anonymous audience. (1991, p. 20).

When an actor is able to produce this same kind of charisma, chemistry, and appeal in front of a camera, the result is *screen presence* (Kennedy, 2019b, p. 192). Presence is a type of recognition of a particular actor's artistry when we are affected by "the actor's particular way of *doing* his role." (States, 1985, p. 165). The actor is his or her own artistic instrument, and Strasberg notes that "[t]he actor is the only art material capable of being both the material and the reality so that you almost cannot tell the two things apart." (Hethmon, 1965, p. 81). In order to increase his or her presence, the actor must be able to "embody or perform features specific to the character. For instance, a character in a stage performance who is a gifted whistler must be played by someone who is adept at whistling"; the better the whistler, the more likely the audience will fully believe in that actor truly is the character, thus amplifying his or her stage presence (Kennedy, 2019b, p. 192).

PeCap provides another form of presence, specifically through the digital counterpart, whether the counterpart resembles the source-actor or not. In a PeCap context, the actor and the digital counterpart are not the same object, and therefore screen presence of the digital character is independent of the source-actor – although the source-actor is often responsible for imbuing the counterpart with qualities that enhance its screen presence. However, since multiple artists are responsible for the digital counterpart's finished performance, its screen presence is not the product of a single individual. This deviates from States' understanding of presence as the particular way in which an individual actor performs his or her role. Instead, I refer to this type of presence produced through a divisibility of performance as a *gestalt* screen presence. Despite the numerous contributions to a *gestalt* screen presence, it must still be acknowledged that the source-actor provides perhaps the greatest single contribution to the *gestalt* (Kennedy, 2019b, pp. 192-193). For instance, Serkis is recognisable in the portrayals of each of his synthespian counterparts (including, Gollum, Caesar, Supreme Leader Snoke, and Baloo) even though none of them are human nor share the same physiology: "Serkis employs different voices and mannerisms in each role, yet something recognisably 'Serkis' still transmits through each character's 'digital makeup'." (2019b, p. 193). However, Serkis' presence is no longer a product of physical embodiment but rather exists as a form of pure kinesis. Indeed, Serkis' corporeal body is literally removed, leaving behind no record of his physical presence in any traditional sense[43] (Balcerzak, 2013, p. 203). Balcerzak explains that "our response to the performer onscreen can transcend the bodily and move completely into the realm of the spectral." (2013, p. 198). Thus, a digital counterpart's screen presence becomes akin to the mercurial double that Artaud describes as a "spectral effigy" in "*Le Théâtre et son Double*" (Artaud, 1976, p. 261).

We can also consider screen presence in terms of how well a digital counterpart integrates into the same screen space as profilmic characters and environments. The exact nature of a digital counterpart's screen presence in this regard will depend on whether the counterpart is fully digital or a partial digital augmentation to a profilmic actor. This is due to the ontological disparity between a flesh-and-blood actor and a fully digital actor, for unlike the living body of an actor, the fully animated actor "lack[s] sentience, corporeal flesh and an autonomous life force." (Hosea, 2012, p. 80). However, as Steve Dixon explains, presence is not linked to physical embodiment but rather in how an actor engages the viewer: "It is content, not container that asserts presence." (2007, p. 134). The quality of a digital counterpart's visual integration into the profilmic screen space largely determines its

[43] While this is true of most popular forms of PeCap, mesh propagation techniques provide a more "solid" recording of the source actor, and it could be argued that such techniques preserve more of the actor's physical presence.

believability as an authentic element within a given environment. The quality of the visual integration depends on a number of factors, including lighting, materials, levels of detail, movement, and range of expressions. Some of these are elements we take for granted and rarely noticed until they are missing, such as a sense of moisture along the lips, including stickiness as the lips part, mucus, and a sense of wetness along the contact between the cornea and the lower eyelid. Likewise, the digital counterpart must also believably perform alongside its co-stars, including matching their eye-lines and timing, as well as producing screen chemistry. Furthermore, the digital character must perform as if it is affected by gravity – a feature that is omnipresent in our Earthly experience but which must be manually constructed in a virtual space. For instance, if a character's movement does not correspond to a normal sense of gravity, the character appears "weightless"[44] through unexpected "weight shifts, sliding feet, rapid change[s] of orientation, physical intersections, lack of integrated body connectivity, and lack of complete 'movement phrases'" (Bishko, 2014, p. 48). If any of these elements is compromised, it negatively impacts the suspension of belief that the digital counterpart is truly present with its co-stars (Kennedy, 2019b, p. 193).

Digital posthumous performances provide a fifth distinction of screen presence, whereby acting is performed through the (digital) body of a dead actor: "[u]ltimately, whose presence is detected through the [digital counterpart's] performance: the original actor whose body is seen, the new actor who contributes to the gestalt screen presence, or a combination of both? (Kennedy, 2019b, p. 194). The synthespian double of Paul Walker as the character Brian O'Connor in "Furious 7" (Wan, 2015) provides a good example of posthumous screen presence. When Walker unexpectedly died midway through production of the film, the filmmakers decided to digitally recreate Walker in all of his uncompleted shots. That is, Walker's synthespian counterpart was designed to fill Walker's screen space throughout parts of the film, rather than being a digital replacement for him throughout the entire film. In most shots featuring the digital Walker, the synthespian is (with a few exceptions)

---

[44] "Weightlessness" is a term animators commonly use to refer to a lack of representation of sufficient weight in a character's movements, rather than a lack of weight entirely.

indistinguishable from the flesh-and-blood performer. The shots with the digital Walker are interwoven with shots containing the flesh-and-blood actor, and it is conceivable that continuing to see the real actor serves to "recharge" the audience's belief in the consistency of Walker's performance. This suggests that "[i]f the performance of a [digital counterpart] closely resembles […] its flesh-and-blood counterpart […], then we accept its performance *as if it were created by the flesh-and-blood performer*." (Kennedy, 2019b, p. 195, original emphasis). And,

> [w]hen the [digital counterpart] is indistinguishable from its living counterpart, the screen presence of the [counterpart] effectively *becomes* the screen presence of the actor it replaces. In this sense, it is at least hypothetically possible to achieve Paul Walker's screen presence without Paul Walker being directly involved in the performance. (Kennedy, 2019b, p. 196, *original emphasis*).

To summarise, within the context of digital counterparts, the concept of presence assumes additional layers of meaning. These distinctions include:

1) being or existing at a particular time and location;
2) the quality of an actor's charisma, chemistry, and appeal that emerges from the actor's particular way of doing his or her role;
3) a gestalt formed through the divisibility of performance in the production of a digital counterpart;
4) how believably a digital counterpart integrates into the same screen space as profilmic characters and environments; and
5) the tendency for a digital counterpart's presence to supplant the presence of its human counterpart (whether alive or dead) when the digital counterpart closely resembles its human counterpart.

In the next section, I build upon the ideas of performance and acting presented so far to develop taxonomies of different performance types specific to digital characters.

4.4 *Defining and Categorising Digital Performance*

Virtual acting performances are a recent phenomenon within the history of screen acting, the rare first instances of which appeared during the mid-1980s. Since then, several terms have been proposed to describe these virtual performers, but such definitions are based more on the means of technical production rather than the characters' performance abilities (Kennedy, 2019a, p. 304). Moreover, these terms are poorly defined and frequently confused, which results in an imprecise language surrounding the performances of digital counterparts. Based on my experience as both an actor and an animator, I propose formalised definitions for the different types of performance common to CG characters in films. I specifically choose to focus on *vactors* in feature films in order to limit my research to a manageable scope; evaluating all television shows and video games, not to mention other formats that utilise virtual characters, would prove untenable due to the sheer volume of performances. However, I suggest that the definitions and categories of performance I identify are equally transferable and robust for understanding virtual acting performances across a range of media.

4.4.1 *Vactors, Synthespians, and Digital Doubles*

One of the first digital performers was Dozo, the star of the 1989 music video "Don't Touch Me" (Kleiser, 2009). Dozo was created by the Kleiser-Walczak Construction Company as one of several computer-generated characters designed by the studio. Dozo represented significant advances in realistic 3D modelling, and the fluidity of her movements approximated human motion more accurately than previous attempts at digital humans (O'Neill, 2016, p. 26). Studio founders Jeff Kleiser and Diana Walczak coined the term "synthespian" to describe these new virtual performers, creating "a portmanteau derived from [syn(thetic) + thespian] and meaning 'digital actor'" (Synthespian Studios, 2015, para. 1). The construction of the term *synthespian* is acting-centric, but a lack of clarity around what constitutes acting within the context of CG characters led to confusion about what kinds of digital characters should be considered synthespians versus other types (Kennedy, 2019a,

p. 297). For instance, Creed's seminal article about cyberstars includes multiple different virtual acting terms, suggesting an equivalence of meaning across all of them:

> Like the silvery, slippery, 'liquid metal' T-1000 robot in Terminator 2: Judgement Day […] or the crowd that fills the dock scene in the opening sequence of Titanic […], the computer-generated figures have no referent in the real world. These are not actors playing a part: rather they are what is known in the industry as 'synthespians', 'cyberstars', or 'vactors' (virtual actors) enacting the parts of extras historically played by real actors. (2000, p. 79).

When the term *synthespian* entered more common usage, it was defined in numerous ways, including "a computer-generated three-dimensional character, either in a wholly animated film or in one that is a mixture of live action and computer animation" (Oxford Living Dictionaries, 2018), and "a computer-generated image of a film actor, esp [sic] used in place of the real actor when shooting special effects or stunts" (Collins English Dictionary, 2018). The Oxford definition is concerned with the technical apparatus that underpins the character, while the Collins definition assumes a synthespian is a digital manifestation of an existing actor that is used primarily as an effects/stunt double (Kennedy, 2019a, p. 297).

I propose unique and meaningful definitions for virtual performance terms that are frequently (and mistakenly) used interchangeably (Figure 18). Historically, there have been no standardised definitions for the terms "vactor", "synthespian", and "digital double", and I suggest that a standardised usage should be performance-based rather than technologically-based, as the latter is more likely subject to revision as technology advances. Also, performance-based definitions align these terms with performance rather than technology, which is important for advancing the appreciation of digital performers as valid actors. It is important to acknowledge "that not all virtual performances involve acting; that is, there are virtual characters that produce performances on screen that do not involve facets of acting such as empathy, simulation, impersonation, pretence, or even pseudo-deflection of reference." (Kennedy, 2019a, p. 300). With this in mind, I propose that "vactor" (itself a portmanteau of [v(irtual) + actor]) is an overarching category of digital counterpart performances, and includes both synthespians and digital doubles.

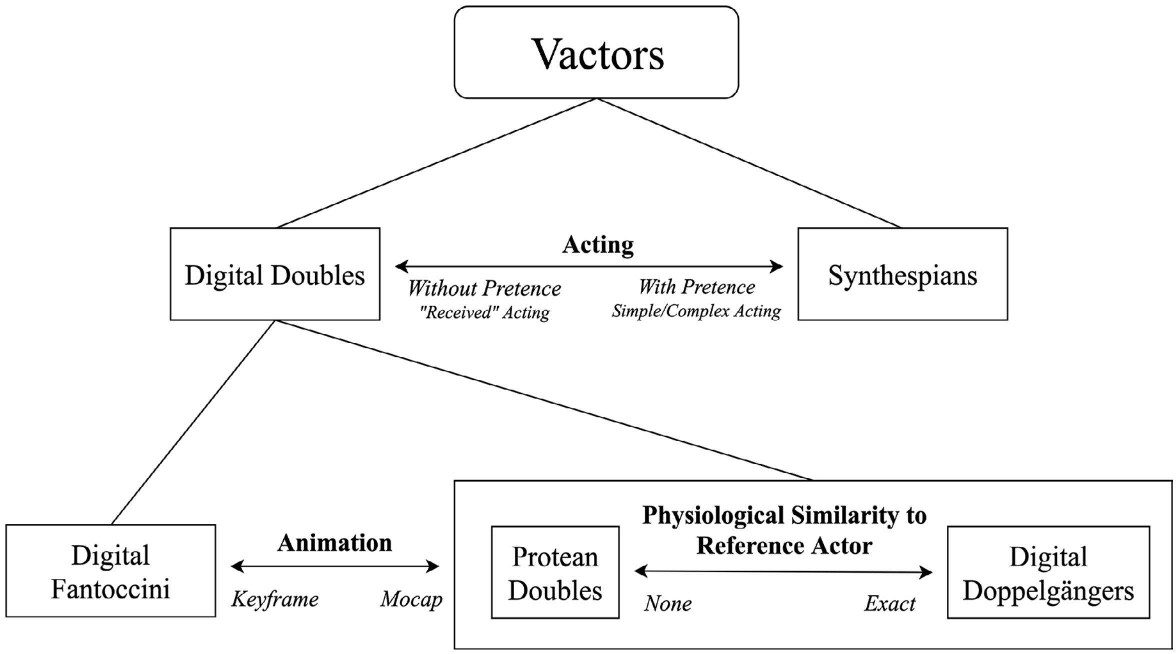


Figure 18. Diagram of virtual performance categories and relationships. Reprinted from "Acting-Centred Definitions of Vactors, Synthespians, and Digital Doubles", by J. Kennedy, 2019, *Proceedings of the 25th International Symposium on Electronic Art* (ISEA), 25, p. 301.

> Vactor: a CG or digitally-augmented character capable of producing a screen-based performance. (2019a, p. 300).

Vactors are defined simply by their capacity to produce a performance and not by the manner of performance they produce. This contrasts with synthespians and digital doubles, which are differentiated by the manners of their performances.

> Synthespian: a CG or digitally-augmented character that advances a screen-based story while producing acting that engages in at least one area of pretence.
>
> Digital Double: a CG character whose participation in a screen-based story is implied more by its physiology, costume, and surroundings than by its actions. (2019a, p. 300).

Based on Kirby's continuum of acting, synthespians are capable of producing simple and complex forms of acting, which makes them more akin to film actors. By contrast, digital doubles produce performances more akin to *received acting* rather than acting based on pretence. Digital doubles are often used as stunt doubles and in scenes where the character must embody a feature or ability the actor does not possess. For instance, when Spider-Man swings between the buildings of New York, it is likely that his performance will be created using a digital double whenever "the camera is located behind or far away from the character, thereby reducing visible sites of pretence." (Kennedy, 2019a, p. 300). Digital extras are often used to fill in large crowd scenes and are placed at such a distance from the camera that sites of individual performances are difficult or impossible to identify. As a result, the extras become more a single performance "mass", without any character visibly engaging in pretence (2019a, p. 300). Such scenes include the digital armies of orcs, goblins, dwarves, and elves in "The Hobbit: The Battle of the Five Armies" (Jackson, 2014), and the thousands of baseball fans in the Oakland Coliseum in "Moneyball" (Miller, 2011).

Digital doubles encompass several different performance roles and it is useful to provide subclassifications for each. Several academic articles published between 2000-2010 suggest the adoption of "digital doppelgänger" in place of "digital double" (Chimielewski, 2005; Bode, 2005), but I argue that a digital doppelgänger should be regarded as a subclassification to refer to "CG characters that are exact digital recreations of the actors on which they are based." (Kennedy, 2019a, p. 301). By contrast, "protean doubles" are CG characters that "do not resemble the actors on which they are based" (2019a, p. 301). In general, the performances of digital and protean doubles are at least partially derived from MoCap. I suggest the term "digital fantoccini" to refer to CG characters whose performances are solely derived through keyframe animation, and irrespective of whether the character resembles a human actor. *Fantoccini* is a type of puppet show that features jointed puppets manipulated by rods, string, and mechanical devices (World Encyclopedia of Puppetry Arts, 2018) – an apt metaphor for how 3D characters are attached to and animated by virtual joints. Digital doppelgängers, protean doubles, and digital fantoccini are all forms of digital doubles and therefore subject to the constraints of received acting (Kennedy, 2019a, p. 301).

It is important to point out a caveat about how I define vactors. Historically, the term "synthespian" was used to describe the first CG human characters that approximated a high degree of realism in films, including "Final Fantasy: The Spirits Within", "The Polar Express", and "Beowulf". The term was also adopted to include realistic CG characters in live-action films, including Jar Jar Binks, Benjamin Button, and the titular aliens from "Avatar". However, I argue that it is flawed to base the definition of "synthespian" in part on an assessment of visual realism because

> [w]hile performance ability remains relatively consistent after repeated viewings, it is common for a CG character that embodies the state-of-the-art visual accomplishment in one year to look far less convincing upon a repeated viewing several years later. This is due to the rapid acceleration of visual improvements regarding modelling, texturing, and rendering CG characters. With each of these improvements, viewer expectations about what is considered "realistic" shifts to the most recent visual marvel. It is therefore conceivable that after a certain amount of time, what was once considered realistic enough to qualify for synthespian status may suddenly fall below the same criteria. (Kennedy, 2019a, pp. 301-302).

Thus, relying on a qualitative assessment of visual accomplishment as a basis for determining synthespian (or vactor) status is specious at best. I suggest that a more reliable marker for attributing vactor status is based on the nature and degree of a CG character's performance.

If degree of visual accomplishment is excluded from our assessment, do we still include characters from fully 3D animated films, such as "The Polar Express" or "The Adventures of Tintin"? If we exclude an assessment based on visual realism, then on what other basis should we judge that a character in a fully animated CG film is either a vactor or "just" an animated character? While the characters in "Final Fantasy: The Spirits Within" or "The Polar Express" may have seemed like impressive examples of digital realism when they were released, it could be argued that their visual details pale by comparison to digital characters in more recent animated films such as "Kingsglaive: Final Fantasy XV" (Nozue, 2016). Do we continue to let certain characters retain the mantle of vactor status just because they have historically been labeled as such? Or can we set a different basis to underpin vactor status? I suggest that the category of vactors applies only to live-action productions – that is, screen-based performances set within a profilmic context. For the purpose of this thesis, I define a *profilmic context* as "a moving image sequence that is constructed, at least in part, by elements that were directly recorded in front of a real-world camera". This excludes from the category of vactors any characters based in fully animated sequences (a condition I call *ex-filmic*[45]), even if those sequences are part of an otherwise live-action film. This means that a CG character in a scene that includes at least one profilmic element would be a vactor, while the same character in an ex-filmic scene would cease to be a vactor. Therefore, vactor status is not a permanent attribute of a CG character but rather a context-dependent description of a particular type of digital performance. This becomes especially notable in films such as "The Smurfs", which include CG characters that we would normally consider animated but not necessarily vactors. Throughout "The Smurfs" (Gosnell, 2011), the titular characters exist in both a fully CG Smurf Village, as well as interact with flesh-and-blood actors in live-action settings. For as long as the Smurfs are fully ensconced by CG surroundings, they are merely animated characters. However, once they interact with live-action elements, the Smurfs become vactors in just those scenes. As a result, a profilmic basis is added to the definitions of vactors, synthespians, and digital doubles in order to provide clarity:

> Vactor: a CG or digitally-augmented character within a profilmic context that is capable of producing a screen-based performance.
>
> Synthespian: a CG or digitally-augmented character within a profilmic context that advances a screen-based story while producing acting that engages in at least one area of pretence.
>
> Digital Double: a CG character within a profilmic context whose participation in a screen-based story is implied more by its physiology, costume, and surroundings than by its actions.

---

[45] In previous research, I tentatively define "ex-filmic" as "a reality or situation solely recorded through software-based cameras" (Kennedy, 2019a, p. 298).

By excluding CG characters in ex-filmic scenes from vactor status, is there a different term we can use to recognise highly realistic ex-filmic performances? Again, the qualitative assessment of realism becomes problematic. For instance, the characters in "Hoodwinked Too! Hood vs. Evil" (Disa, 2011) in no way compare to the visual realism of the animated characters in "Kingsglaive", but where exactly should we set the bar for realism within a fully animated context? Are the characters in "How to Train Your Dragon: The Hidden World" (DeBlois, 2019) or "The Incredibles 2" (Bird, 2018) "real enough", and if not, why not? What does *realism* even mean in a context that is so technologically derived? This is a ripe subject for future research.

4.4.2 Categories of Vactor Performance

With meaningful definitions for vactors, synthespians, and digital doubles established, it is possible to create a taxonomy of vactor performance styles. I identify at least eight vactor acting categories common to films produced between 2010 and 2019[46] (and extendable to films prior to 2010). Just as with the definitions for vactors, these categories are performance-based rather than dependent on visual aesthetics. It is important to note that these categories are not mutually exclusive, meaning that a vactor within a film may participate in more than one category at different times.

4.4.2.1 Verisimilar Realistic Acting

Verisimilar realistic acting is a form of vactor performance that highly approximates real-world acting within the tradition of realism. These performances are most similar to our experience of the physical and emotional believability of flesh-and-blood performers within Stanislavski-influenced traditions of acting. As a result, this style of vactor performance is limited to synthespian performers. Verisimilar realistic acting simulates reality and is often

[46] This research originally emerged from an inquiry into how many vactors exist in feature films since "Avatar", which is why the research starts from 2010. Maintaining a range from 2010-2019 also helps to limit the scope of the work, as each year yields a progressively greater number of vactors.

performed via synthespians that are heavily influenced by or indistinguishable from their source actor counterparts (Kennedy, 2019a, p. 302). Examples of verisimilar realistic acting include Clu from "Tron: Legacy" (Kosinski, 2010), Gollum from "The Hobbit: An Unexpected Journey", Durotan from "Warcraft", Brian O'Connor from "Furious 7", Thanos from "Avengers: Infinity War", and Alita from "Alita: Battle Angel" (Rodriguez, 2019).

4.4.2.2 Stylised Realistic Acting

Stylised realistic acting is a form of vactor performance that transcends our experience of physical and emotional contexts. Stylised realistic acting encompasses a wide range of performance and visual styles, and is especially common among characters that possess superpowers or cartoon designs. Stylised realistic acting can be similar to verisimilar realistic acting in terms of a high standard for visual realism, but stylised realistic acting embraces imaginative performances that surpass human experience. Likewise, stylised realistic acting may embrace character acting and designs that are highly stylised, resulting in performances of cartoon characters that believably interact within a profilmic context. Due to its inclusivity, this is the most common vactor acting style category in feature films (Kennedy, 2019a, pp. 302-303). Examples of stylised realistic acting include Tony Stark/Iron Man from "Iron Man 2" (Favreau, 2010), Bumblebee from "Transformers: Dark of the Moon" (Bay, 2011), Papa Smurf from "The Smurfs", Malekith from "Thor: The Dark World" (Taylor, 2013), and BFG from "The BFG".

4.4.2.3 Non-Realistic Acting

Non-realistic acting is a form of vactor performance in which the styles of acting are not inherited from the Stanislavski tradition. "Non-realistic" in this case is not intended as pejorative but is rather used to distinguish between common approaches to Western and non-Western forms of storytelling in film. The non-realistic acting category "encompass[es] a variety of performance modes outside of the Western tradition" and is most common among Bollywood, Chinese, and Korean cinema (Kennedy, 2019a, p. 303). Examples of non-realistic acting include Chitti from "Enthiran" (Shankar, 2010), Almighty Foot from "Journey to the West: Conquering the Demons" (Chow & Kwok, 2013), and Zhong Kui from "Zhong Kui: Snow Girl and the Dark Crystal" (Pau & Tianyu, 2015).

4.4.2.4 Anthropomorphised Acting

Anthropomorphised acting is a form of vactor performance that involves digital creatures acting with human-like intentions. Animation has a long history of anthropomorphised creatures and this popular form of vactor performance continues the tradition. Digital creatures produce anthropomorphised acting through a human-like range of empathy, emotion, and intention, despite whether they bear any human resemblance (Kennedy, 2019a, p. 303). Examples of anthropomorphised acting include Reepicheep from "The Chronicles of Narnia: Voyage of the Dawn Treader" (Apted, 2010), the White Rabbit from "Alice in Wonderland" (Burton, 2010), Caesar from "Rise of the Planet of the Apes", and Baloo from "The Jungle Book" (Favreau, 2016).

4.4.2.5 Sentient Acting

Sentient acting is a form of vactor performance that involves sentient creatures performing according to their innate behaviour. This form of vactor performance is common among creatures in film who are unable to directly communicate like humans, but who are still intelligent and able to express clear intentions, emotions, and thinking. These creatures are often fantastical or alien and they communicate through their actions alone (Kennedy, 2019a, p. 303). Examples of sentient acting include the Homunculi from "Don't Be Afraid of the Dark" (Nixey, 2010), the Dragon Spirit from "The Last Airbender" (Shyamalan, 2010), the Troll from "Snow White and the Huntsman" (Sanders, 2012), the Kronan Marauder from "Thor: The Lost World", and the Cloak of Levitation from "Doctor Strange" (Derrickson, 2016).

4.4.2.6 Complex Animal Performance

Complex animal performance is a form of vactor performance that involves non-sentient creatures performing directed actions. By contrast to sentient creatures, non-sentient creatures lack clear reasoning abilities. The types of performances non-sentient creatures produce are either basic or complex. Complex animal performance refers to actions the creature would not likely produce of its own accord, and which would likely require extensive animal training. This can be a bit arbitrary given that many of the animals included in this category are imaginary, and therefore it is debatable what their innate behaviours would be (Kennedy, 2019a, p. 303). Examples of complex animal performances include the Scorpiochs from “Clash of the Titans” (Leterrier, 2010), the Bandersnatch from “Alice in Wonderland”, the Wargs from “The Hobbit: An Unexpected Journey”, the Gorgosaurus from “Walking with Dinosaurs” (Nightingale & Cook, 2013), and the Rathgars from “The Force Awakens”.

4.4.2.7 Basic Animal Performance

Basic animal performance is a form of vactor performance that involves non-sentient creatures performing according to their innate behaviours. These are the kinds of actions and behaviours we would expect of an animal without training it to perform in a particular manner. An important distinction here is that the basic animal performance must advance the story in some way – that is, the creature does not just exist in the background as a form of set-dressing (Kennedy, 2019a, p. 303). Examples of basic animal performance include the Adult Parasaurolophus in “The Tree of Life” (Malick, 2011), the White Elk in “Snow White and the Huntsman”, the Tracker Jackers in “The Hunger Games” (Ross, 2012), and the Buffalo in “The Lone Ranger” (Verbinski, 2013).

4.4.2.8 Physical Acting

Physical acting is a form of vactor performance that involves characters primarily performing a single type of action, such as fighting. This is a common form of vactor performance that often appears in the form of digital stunt replacements. However, the performances in this category must participate in at least one form of pretence, which means that digital doubles are by definition excluded (Kennedy, 2019a, p. 3030). Examples of physical acting include Tony Stark/Iron Man from "Iron Man 2" (as an example of how a character may participate in more than one vactor performance category), the Frost Giants from "Thor" (Branagh, 2011), Peter Parker/Spider-Man from "The Amazing Spider-Man" (Webb, 2012), Ryan Stone from "Gravity" (Cuarón, 2013), and Krrish from "Krrish 3" (Roshan, 2013).

4.5 Conclusion

While directors, actors, and animators frequently discuss the use and impact of emotions in their work, the actual means by which emotions are interpreted are complex and not fully reliable. For instance, while actors frequently discuss the importance of being emotionally connected (experiencing an emotion while expressing the same emotion), this chapter examined how actors may also experience an emotion without effectively expressing it (a masked emotion). Likewise, an actor may use particular facial muscle configurations and breathing patterns to convey an emotion without actually experiencing the emotion (a simulated emotion). An actor may also express an emotion other than what he or she actually experiences, which leads to concerns over misinterpretation of expressions. Understanding these fundamental relationships between expressions and emotions is critical for both actors and animators when improving the quality of their character performances.

While animators do not fit the common conception of actors, animators perform most (and sometimes all) of the same functions as actors. Director Brad Bird notes that an animated character's performance is the product of animators, and if that character provokes an emotional response from the audience, it is largely due to the little-known animators who invest themselves in the character. For Bird, animated characters are actors who produce strong emotions, but the emotions originate within the animators and therefore the animators create the performances, making them the true performers. (Crafton, 2013, pp. 15-16). While live-action actors are the stars who perform characters, in animation the characters themselves become the stars rather than the animators responsible for creating their performances. For instance, many will be familiar with Donald Duck, Homer Simpson, and SpongeBob SquarePants without knowing any of the animators responsible for their performances. Audiences are more likely to identify an animated character's performance with the voice-artist responsible for the character, even though the voice-artist may have little, if anything, to do with the character's movements (Hosea, 2009, p. 14). Within MoCap for films, the source-actor(s) and the animators share responsibility for a digital counterpart's finished performance, but it is only the source-actor(s) who receive clear performance credit. Again, individual animators' contributions remain largely unrecognised.

Animator Barry Purves is more cautious about labelling animators as actors. He believes that the concept of acting may be too entangled with connotations of theatre and film, and therefore it is safer "to say that all animators need to have the sensibilities of a performer, as that widens the field and brings in dance, mime, singing and a million related skills." (2008, p. xvii). Indeed, animators need to be able to understand concepts of performance at a wider level than most actors will ever encounter. An animator may never be a great dancer, but one needs to be able to understand body movement, force, intention, and timing in such a way that one could animate a tango master. The animator is just as likely to encounter scenes requiring physical dexterity, such as martial arts and acrobatics, to a diverse range of acting performances from subtle to bombastic – from Mark Rylance to Jim Carrey. The animator's great skill is in conveying myriad believable performances through a single animated body, and in so doing, surpassing the range and dexterity of most actors. Hayes & Webster summarise it well:

> [h]ow many "real" actors are required to undertake roles that cross gender or race, or cover age ranges from 1 to 101, let alone creating a performance across *species*. Animators are also asked to create character-based performances using inorganic objects and even abstract shapes that not only imbues them with a personality but

makes them completely believable. (2013, p. xvii, original emphasis).

Hosea discusses how during her postgraduate teaching at Central Saint Martins, at the University of the Arts in London, students studied a range of performance methods, including Laban Movement Analysis, Lecoq, Stanislavski, and the Method.[47] She notes that Method acting techniques work well for complex animated characters but can become overcomplicated in more stylised animated characters. Hence, an education that focuses on a broader and more exterior style of expression is appropriate to the production of cartoon animation (2012, p. 56), whereas a focus on more internalised techniques is relevant to more realistic styles of animation. The animator as an actor must be familiar with this broad range of approaches and draw upon the relevant techniques appropriate to the graphical and storytelling needs of a given production. Based on their individual skills, some animators may be better suited for specific types of animated performances. Animators with a good sense of action timing and economy of movement may be at home creating animations for games, whereas someone with a solid sense of character motivation, intention, and acting subtlety may be adept at animated drama (Hayes & Webster, 2013, p. 69).

Production constraints play key roles in the realities of producing animated acting. For instance, on small, independent productions, an individual animator may be responsible for the performance of a single character (or even a cast of characters) throughout the entire film. It is much easier to conceive of an animator as an actor in this type of situation. However, when working on larger-scale productions, animators work as part of teams that often include tens if not hundreds of other animators. A chief concern for animation producers is how to maintain performance continuity when so many individual animators contribute to the final result. Hayes & Webster note that

> [w]ithin large production crews, the anonymity of the animator is essential; the animator must become invisible, leaving only the personality of the characters on

---

[47] To this I would add a study of Feldenkrais, who can be used as a basis for movement awareness that improves the animators' sense of feeling an animation through their bodies ("listening to their bodies"). This is a further area for future research and I believe it would meaningfully complement a study of Laban Movement Analysis within an animation curriculum.

> show to the audience. […] Continuity throughout a film may be difficult to maintain particularly if it is being animated by numerous hands. This is where the virtuosity of animators will enable them to create quality performances regardless of the character, though virtuosity for its own sake, if this descends into showboating, is likely to be at the cost of both performance and of continuity (2013, p. 68).

Likewise, PeCap animators must be able to preserve the sense of the source-actor's performance even when editing the motion or creating new performance elements from scratch. The animator's role is to be a seamless performer who never upstages his or her animated character. The character's performance must be created through a single, unified vision across all members of the animation team. This is facilitated by character style sheets and production bibles that outline consistency of characterisation and visual appearance. With such heavily prescribed performance restrictions, it may seem at times that the animated character is responsible for animating the actor, and not the other way around. For instance, when an animator mouths words into a mirror to capture different facial configurations for lip-syncing dialogue, to what extent do the model sheets and the pre-recorded dialogue affect how the animator contorts his or her face? (Hosea, 2012, p. 70).

The key difference between animators and live-action actors is that animators never produce their work in the present moment. Likewise, animators do not require the kinds of relaxation techniques, concentration exercises, or vocal training common among actors. Animators produce the illusion of a present moment and use many of the same acting principles to get there (Hooks, 2011, p. 7), but the animators' method is a performance by proxy rather than through a directed performance of their own bodies. Walt Stanchfield noted to the early Disney animators that being an actor by virtue of being an animator doesn't mean that you necessarily are a good stage actor, but that you are "sensitive to poses and gestures that portray the various moods and emotions that storytelling demands" (2013, p. 81). This also means that an animator should be cautious of relying on a repetitive use of similar poses to convey similar meanings. One such example is the rubbing of the back of the neck when a character feels nervous or uncertain. This was considered a brilliant and unexpected example of psychological gesture when first used by Baloo in "The Jungle Book", but it quickly became a cliché used in countless other character animations (Osborne, 2015, p. 38). Animators must also be cautious of relying too heavily on live-action reference. Instead, they must select specific elements from the reference that they believe will serve the requirements of the scene, while avoiding copying the reference as a whole (Wells, 2006, p. 134). In this sense, animators are not mere slaves to live-action reference but are instead connoisseurs of motion.

5
SAPER VEDERE

5.1 Methodology

This research occurs at the intersection of the fields of acting, animation, and motion capture – areas that have limited consideration from an academic, pragmatic perspective beyond the texts discussed in Section 1.3.[48] More than a single methodology is needed to attend to the challenge of reconciling these interdisciplinary spaces. David Gray indicates that multiple methods/methodologies are often needed when reconciling a number of different research questions, as an approach used to address one question may not be appropriate for another. Multiple methodologies may also result in triangulation, whereby any perceived weakness of one approach is balanced out by the combined strength of a range of approaches (2004, p. 37). This is especially valuable in the production and evaluation of creative projects where the validity of the researcher's subjectivity may come into question. This chapter provides the basis and theoretical grounding for these methodology choices and how they interact with and reinforce each other.

This research is broadly framed within the philosophies of interpretivism and pragmatism. In this research, I explore the nature of how acting is produced through the motion capture process. As such, there is no "objective reality" that I can observe or study. Instead, my observations are interpreted through my own understanding of the separate crafts of actors and animators, as well as through my theoretical grounding as an academic in the field of animation. Interpretivism is well-suited for studying the actions of individuals, especially in terms of their unique, individual, and qualitative behaviours (Gray, 2004, p. 23). Thus, a

[48] That is, it is rare for academic research about acting in animation to be intended for use by actors or animators, themselves.

interpretivist epistemology grants me as the researcher the ability to qualitatively interrogate and assess a subject's behaviour, even when I am my own subject. My combined experience as an animator, actor, and academic professional is rare and provides me with an important nexus between these fields. Therefore, I require several interpretivist approaches that can accommodate this combination of skills.

I pursue this research through practice first and invoke a practice-led methodology to attend to what is revealed through my praxis. Due to the difficulty in distinguishing the roles of actors and animators in performance capture, in this research I produce PeCap experiments in which I am the sole actor and animator. As such, for the purpose of this research I am both the subject and the observer. A phenomenological research model is used to gain insight about how the roles of an actor and animator are constructed and reinforce each other in relation to the production of performance via PeCap. This research is pragmatically considered in that the ideas developed through this research need to be shown to work (D. Gray, 2004, p. 28), and that the results of this research are intended to offer animators and actors new and useful ways to consider performance. To achieve this, I employ a series of methods for self-reporting my insights, findings, and technical processes. I reflect on these reports to create recommendations for best practices, and in later stages of the project I implement and test these suggestions to determine whether they have meaningful practical implications.

5.1.1 Practice-Led Research

As this research heavily involves practice, a practice-aligned methodology is needed. Linda Candy notes that methodologies such as practice-based research and practice-led research lack consistent definitions and are often used interchangeably (2006, p. 3). For the purpose of this thesis, I rely on Candy's distinction between practice-*based* and practice-*led* research:

1. If a creative artefact is the basis of the contribution to knowledge, the research is practice-*based*.

2. If the research leads primarily to new understandings about practice, it is practice-*led*. (2006, p. 1).

The significant differences between these two methodologies is that for practice-based research, a full comprehension of the research can only be articulated in direct reference to the creative outcomes, whereas practice-led research "is concerned with the nature of practice [… which] may be fully described in text form without the inclusion of a creative work." (2006, p. 1). Practice-led research is used to either "advance knowledge about practice, or to advance knowledge within practice", and often occurs within the larger area of action research (2006, p. 1).

Candy's definition of practice-led research as *research that leads primarily to new understandings about practice* is an apt approach to this thesis. As Hosea acknowledges in her PhD thesis, the academic literature about animation is primarily theoretical from the perspective of the viewer (2012, p. 5). Importantly, the research in this thesis engages with practice as a means for gaining evidence and contributes a practice-informed perspective to animation and acting research. Through a series of performance capture experiments, I continually engage with practice through a process of reflection and reflexion. Redhead describes practice-led research projects as "frequently exploratory, open-ended, and often seen as specialised." (2012). This thesis should be understood as a series of exploratory acting/PeCap experiments that influence the design and execution of each subsequent experiment, as well as the acting and animation techniques used in each. The creative output of each experiment is an animation of live-acting conveyed through the performance of a digital counterpart. Each animation is used to evaluate the effectiveness of the acting and animation techniques utilised in its creation.

5.1.2 Phenomenology

The practical experiments within this research are conducted based on the following five stages (Figure 19):

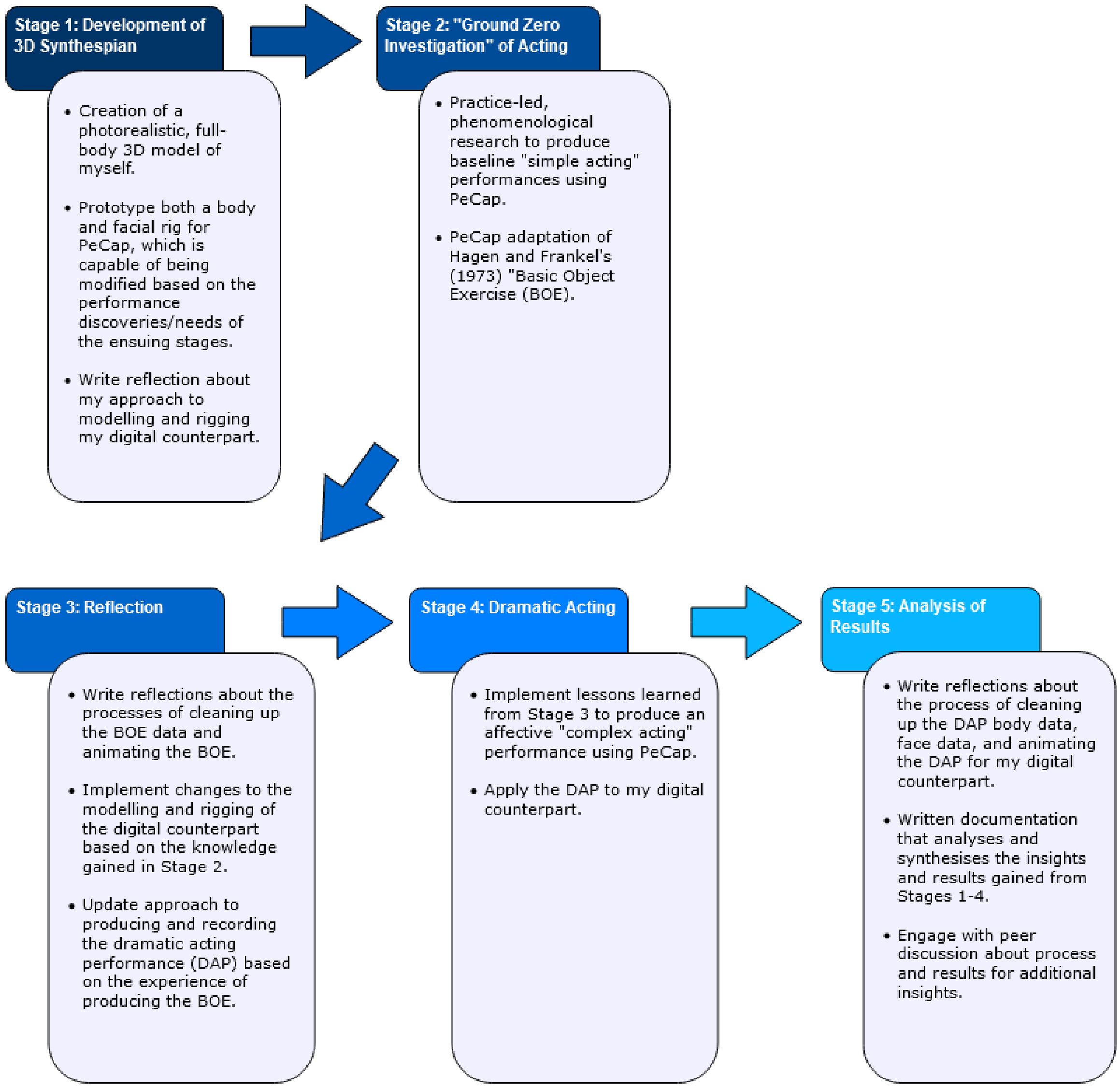


Figure 19. Overview of the research process in this thesis.

1) develop a photorealistic 3D character model and rig based on my physiology;
2) produce a "simple acting" (Kirby, 1972, p. 8) experiment using PeCap and apply it to my digital counterpart;
3) reflect on the results of Stage 2 and implement updated modelling, rigging, and performance strategies based on these results;
4) produce a "complex acting" (1972, p. 8) experiment using PeCap and apply it to my digital counterpart;
5) analyse and synthesise my articulations of the acting/animation results conducted during Stages 1-4.

5.2 Post-Merleau-Ponty Phenomenological Approach

During the mid-twentieth century, Maurice Merleau-Ponty challenged the prevailing characterisation of the human body as a mere material object, subject to general scientific laws, and no different from any other physical object. Drew Leder describes Merleau-Ponty's basic premise: that the body "is never just an object in the world but that very medium whereby our world comes into being." (1990, p. 5). While the English language only has one word for "body", the concept of "body-as-experiencer" is better articulated within German by the word *Leib* (living body), which sits in contrast to *Körper* (physical body) (1990, p. 5). Merleau-Ponty's contributions helped to reshape Cartesian mind-body dualism through an existentialist understanding of the lived body as both the first-person perspective of the body-lived-from-within, and the "object body" as externally perceived. In his later work, Merleau-Ponty emphasises that lived embodiment must "be *both* subject and an object available to external gaze." (Leder, 1990, p. 6). While the Cartesian split between mind and body suggests that the mind can achieve metaphysical privacy, Merleau-Ponty helps us to understand how when people struggle to withhold vulnerable feelings, their bodies often betray their moods and intentions through outward signs within public experience (1990, p. 6).

The experience of a phantom limb forms a critical discussion point for Merleau-Ponty. Phantom Limb Syndrome (PLS) occurs when a person loses a limb but still experiences sensations associated with the missing part, including itching, pain, touch, pressure, and numbness (Winchester Hospital, 2019). PLS brings into question our assumptions about our experience of our bodies from within (interoceptivity) and without (exteroceptivity). The phantom limb essentially becomes an extension of the body that is experienced as if it is still real. For example, Merleau-Ponty describes that having a phantom arm allows one to "to remain open to all of the actions of which the arm alone is capable and to stay within the practical field that one had prior to the mutilation." (2014, p. 84). The experience of this false reality where one's body is still complete provides the patient with "certainty of his [bodily] integrity." (2014, p. 84). According to Merleau-Ponty, our beliefs and experiences of the world are intrinsically based on our habits and interests that form our relationship to our environment and how our bodies are capable of interacting within that environment. A limb is not merely a sensory device capable of touching and grasping but also a means for a person to relate with the world through actions that can specifically be carried out through the possession of the limb. For instance, when we possess hands, we perceive certain objects in the world based on our ability to grasp them, such as door knobs. However, if we were to lose our hands, our relationship to door knobs would suddenly change as we are no longer capable of the act of grasping, and we would need to form a new relationship to such objects:

> At the same moment that my usual world gives rise to habitual intentions in me, I can no longer actually unite with it if I have lost a limb. Manipulable objects, precisely insofar as they appear as manipulable, appeal to a hand that I no longer have. (2014, p. 84).

In a MoCap context, either the limb or the object may be phantom-like. The source-actor may reach out toward a door knob that exists only in the virtual world and may witness the *grasp action* play out solely through a screen displaying a digital representation of the performance. The actor is still engaging with the experience of grasping a door knob, although the door knob itself is now the phantom, and the idea of "grasping" is limited to imagining the door knob's existence *vis-à-vis* its digital representation on screen. The actor never receives tactile confirmation of the door knob, but rather relies on vision to confirm its existence and his or her physical relationship to it.

Merleau-Ponty's conception of the lived body presupposes an awareness of the body, but as Leder explains, we are frequently unaware of our body and experience it through perceptive absence:

> While in one sense the body is the most abiding and inescapable presence in our lives, it is also essentially characterized by absence. That is, one's own body is rarely the thematic object of experience. When reading a book or lost in thought, my own bodily state may be the farthest thing from my awareness. I experientially dwell in a world of ideas, paying little heed to my physical sensations or posture. (1990, p. 1).

In this research I engage with a post-Merleau-Ponty phenomenological approach to catalogue and analyse my experience as the actor and the animator creating these 3D animation experiments. I find that in my own acting experience, it is necessary to maintain an awareness of what my body is doing. This is especially true in terms of hitting my marks, ensuring that the scale of my movement is appropriate to the framing of the camera, and preventing any subconscious signs of unease, such as unwanted blinking, over-gesticulation (especially in the head and neck), or irregular breathing. It is a challenge to make and respond to these observations while not letting this awareness impact my conscious, emotional presence in a scene. While lived experience may often result in an absence of awareness of the body, my experience as an actor suggests that acting requires a concerted effort of maintaining an awareness of the physical and mental states of my body.

While Leder identifies the phenomenon of body absence as an inadequacy in Merleau-Ponty's account, he still upholds the necessity for a physical body to be experienced as both subject and object. However, it is worth considering how the concept of absence could be extended to motion capture, whereby the physical body is displaced into the virtual realm. As a MoCap performer viewing real-time feedback of my body actions mapped onto a virtual character, I experience my lived body as subject, object, and simulacrum. I refer to the phenomenological experience of my body as a virtual representation as the "actor's double". In this way, absence does not just refer to the lack of perception of one's physical body, but also to the lack of interior perception of the virtual body. In the spirit of *Leib* and *Körper*, we can invoke yet another German word, *Doppelgänger* (literally "double-goer", referring to "an apparition or double of a living person" [Lexico, 2020]), to adequately refer to this process of simulated doubling.

This research posits that *Leib*, *Körper*, and *Doppelgänger* exist at separate poles that circumscribe an actor's perception of his or her body performing via motion capture (Figure 20). I distinguish these poles as follows:

- *Leib* refers to the living body as perceived from within, which can be thought of as the tool (or the "actor's instrument" [Huston, 1992]) through which the actor perceives both mental and emotional states. Tisseron, Tordo, and Baddoura

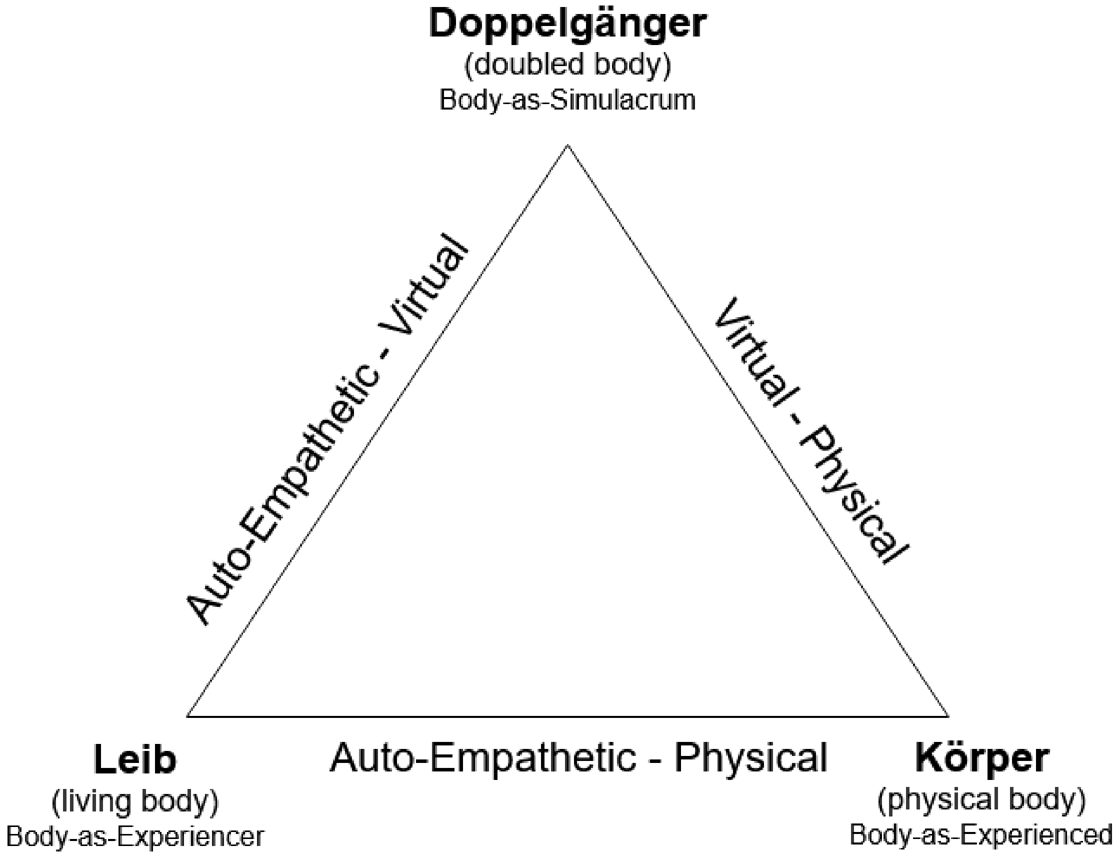


Figure 20. The actor's perception of the performing body in motion capture.

describe the phenomenon in which a person experiences representations of his or her own subjective states as "auto-empathy" (2015, p. 98).[49] In this thesis, I use auto-empathy to refer to the combined experience of a person's emotional state and mental state.

- *Körper* refers to the physical body as externally perceived, such that the corporeal body is perceived by both the actor and audience as an object. The audience experiences the *Körper* through the act of seeing, while the actor perceives his/her physical body as the tool[50] through which tactile and proprioceptive perception is transmitted.
- *Doppelgänger* refers to the virtual simulation of the body, which is always perceived

---

[49] This is as opposed to the more familiar notion of empathy as the ability to "put oneself in another's shoes" whereby one experiences the emotional and cognitive state of another person within oneself (Misselhorn, 2009, p. 351).

[50] Both *Leib* and *Körper* can be considered essential parts of the actor's instrument. Actors invest a great deal of time and energy into producing both mental and physiological awareness in order to provide greater emotional and physical flexibility to fulfil a role.

exteroceptively. The actor may respond auto-empathetically or physically based on perceptions of this virtual double, but the actor never experiences the double interoceptively; indeed, the spectral figure has no interiority to experience.

This model is in part used to reconcile the different perceptions of my body throughout the experience of various strategies of performance within a motion capture context.

The Leib-Körper continuum locates the degree to which an actor is simultaneously aware of his/her auto-empathetic and physical perceptions, and how the perception of one pole affects the perception of the other.[51] This "dual consciousness" is the relationship the actor has between remaining psychologically present in a scene while also maintaining an awareness of how his/her body is used (Zarrilli, 2004, p. 665). This latter awareness can occur in two ways: during the act of acting or when reviewing a recording of the performance. Phillip Zarrilli posits that a person's understanding of his or her body is actually a process of reconciling many different perceptions of bodies in everyday and extra-daily experience (2004, p. 655). Reflecting on my own practice, there is my body as I perceive it during the act of performance, and there is the body that I witness separate to myself when reviewing my performance as a recording either on film or via motion capture. When viewing the recorded performance, I evaluate what Zarrilli refers to as a set of "actions/tasks that constitute the aesthetic outer body offered for the abstractive gaze of the spectator" (2004, p. 664). It is during this review that I perceive my body as an external character rather than a medium through which I participate in direct experience of being-in-the-world.

The Leib-Doppelgänger continuum locates the degree to which the perception of an actor's auto-empathetic state affects the performance of the virtual body, and how the actor's perception of the performing virtual body affects his/her auto-empathetic state. For instance, during the filming of "Avengers: Age of Ultron" (Whedon, 2015), actor Mark

[51] It is also along this continuum that Leder's phenomenon of perceptive absence may be experienced.

Ruffalo was able to interact with a digital avatar of his character, the Hulk, before going on set (Stern, 2014, para. 26). Through perceiving how his expressions and physicality translated onto the virtual character, Ruffalo could be both mentally and emotionally affected by the virtual body's performance. This, in turn, creates a relationship between the auto-empathetic body and virtual body, helping to ground the actor in both the human and fantastic aspects of the character, as well as making him more comfortable in his performance (2014, para. 26).

Similarly, the Doppelgänger-Körper continuum locates the degree to which an actor's physical state affects the performance of the virtual body, and how the actor's perception of the performing virtual body affects his/her physical state. Geoff Gilson, a dancer and performance artist, describes his movements mapped onto a virtual character as quickly engendering a different "feel" to his body than his normal perception, whether it is through differences in gender, scale, proportions, or timing (Nikolai, Bennett, Marks, & Gilson, 2019, pp. 146-147):

> When my avatar went from a cuboid humanoid to a feathery, winged-arm humanoid, my instant desire was to lift off: my movements became lighter, and I rose up higher on my toes; movements felt like less effort; and I felt more stable, with a higher centre of gravity. In contrast, when my avatar was made stockier, with larger arms, my instant response was to see and feel these new powerful arms working and I began a series of more punctuated boxing movements. I literally felt more inclined to characterise this body. It is worth thinking about how the visual self-representation can influence movement decisions so readily and how, in conjunction with the sense of the three-dimensionally represented body, this can be a tool for further choreographic development. (2019, p. 147).

Andy Serkis describes that coming face-to-face with a virtual Gollum during the production of "The Two Towers" instantly affected his perception of the character: "I got into character as Gollum, hunching my back and crouching on my haunches, splaying my fingers, and in the goggles Gollum responded, simultaneously mirroring my every action, only in a more extreme way." (2003, p. 36). Based on his perception of the virtual Gollum's movement, Serkis modulated his own body performance to accommodate the physiological differences between his human body and his body-as-simulacrum: "I began to realize the potential for some pretty subtle, understated 'cyber-acting' as I lurked, loped, crawled, shook and slept as

Gollum." (2003, p. 36). The virtual/digital realm provides a range of new insights about the relationship between the auto-empathetic and virtual perceptions of self. This is reminiscent of the real-world practice of marionette puppetry, in which the puppeteer, while witnessing the puppet's movements, may be emotionally affected by the performance that he or she crafts through skillfull manipulation of the hand controls. This emotional response mechanism may in turn affect the nature and intensity of the ensuing dramatic choices within the puppet's performance – the artist's response to his or her own performance becomes part of the art itself. However, whereas a puppeteers primarily use their hands and arms to control the movement of marionette characters, MoCap performers must use full-body movements to inhabit the role of their digital counterparts. To further complicate this process, the virtual characters may dramatically differ by size and proportion to previous characters they've played, thus requiring the actors to adapt their performances through a wide variety of character designs.

This type of feedback works best when the digital counterpart of the actor's character is virtually available, but that is not the case in every production – for instance, sometimes MoCap is recorded before the counterpart's model and rig are finished. However, the motion data is usually able to be visualized in real-time in a more simplified format. In the case of the optical tracking used in this research, the tracking markers are represented as points in virtual space connected via straight lines that produce a rudimentary humanoid skeleton. This skeleton has the same proportions as the actor and provides an accurate representation of the actor's movement in real-time. While viewing the finished character model would be optimal, being able to perceive just a one-to-one correspondence between the actor and the virtual skeleton can still be useful. Here I draw on my own experience of chronic back pain, which often results in noticeable tension as well as a pronounced hunch in my neck and shoulders. I am not always aware of the physical effect this condition has on my body, but when I watch myself perform through a virtual skeleton, I can quickly identify and attenuate problematic movements.

5.3 Motion Capture Languages

While perception is a key to generating meaning for oneself, a thesis such as this requires that perception can be articulated and conveyed via a linguistic mode of argument in order to be understood by others. Michael Biggs identifies three principle types of experiential knowledge: explicit, tacit, and ineffable. Explicit knowledge is capable of being conveyed through words. Tacit knowledge is embedded within experience (such as kinaesthetic experience), which makes it difficult to efficiently convey through words. Ineffable knowledge defies linguistic expression altogether (2004, p. 6). An example of tacit knowledge occurs when identifying a friend's face from a crowd, but the same form of recognition could be difficult to describe to a police sketch artist. It becomes more efficient to point at the drawing of the face and say "the jaw is squarer" or "the lips are too wide". Biggs identifies experiential feelings as ineffable, and while we do out best to approximate our feelings with words, linguistic expression often falls short (2004, p. 7) – resulting in the common refrain that our feelings "sound stupid" once we describe them aloud.

In this research I am attempting to explicitly describe my experience of acting and animating, two creative disciplines which lend themselves more to tacit experience than linguistic expression. This is especially true of the animation process: while explicit descriptions suffice for technically oriented animation, moments that demand greater creativity and personal insight in animation are more often tacitly understood. Acting, whether through live performance or animation, can result in ineffable experiential feelings that we struggle to meaningfully describe at all. The construction of expressions throughout an animated scene can result in the perception of different emotions, the source of which is in tiny physiological cues throughout the face. An animator may be able to "feel" when the expressions are *just right* in order to achieve the desired emotional read, but he or she may struggle to express precisely *why* a particular physiological configuration produces a specific result. When actors and animators are able to describe their processes, they will generally do so within the inherited languages and vocabularies of their professional history (Delbridge, 2014, p. 61), such as the theories and techniques specifically used by their coaches and teachers. Such inherited terminology and experience may be opaque to anyone who is unfamiliar with them. When artists from different disciplines or stylistic traditions use similar language to describe their processes, their respective vocabularies may not have equivalent meanings.

Hosea cautions the would-be PeCap researcher to seriously consider the different languages and creative/technical traditions at play in a MoCap studio, including those of the actors, technicians, directors, and animators (2012, p. 61). The performance of a digital counterpart involves a highly mediated form of acting, which is prone to questions about the role of the actors versus animators in terms of authorship. Furthermore, 3D animated characters and especially motion-captured characters are usually the product of a team of artists, and that "it is extremely unlikely that there would be a one-to-one correspondence between one animator and one character's performance." (2012, p. 68). Due to the difficulty in distinguishing between the roles of actors and animators in PeCap, as well as pinpointing the exact contributions of any individual animator, I serve as the sole creative actor and animator in this research. Because of this, I can account for my thinking/creative process while engaging in acting and rest assured that the language I use to describe this process will be understood by the animator, who is also me. This minimises the number of variables I have to account for in terms of reflection and analysis. I am also able to relate the language and concerns of animation back into the acting process to investigate what effect the experience of creating 3D animation has on the process of acting in a PeCap context.

By using myself as the sole source of creative performance through the acting and animation process in this research, I am able to clearly locate and articulate the language and techniques of my professional history. My primary screen acting training has been through classes with the Kacie Stetson Studio in Auckland, New Zealand. Stetson uses Ivana Chubbuck's "The Power of the Actor" as a foundational text in her classes and supplements this with her own professional experience as an actor and director. Stretching further back, Chubbuck's techniques are modelled, in part, on the acting teachings of the renowned actress and acting coach, Uta Hagen, who emphasised truth and realism in acting through emotional "substitutions" (Hagen and Frankel, 1973; Hagen, 1991). Chubbuck's technique varies from Hagen's in that it is based more on psychology and behavioural science. While a more Stasbergian approach to Method acting could work within a performance capture context, his emphasis on emotional/affective recall focuses on the actor's past, whereas Stetson's technique involves a focus on emotionally charged relationships that are presently available to the actor. My acting experience is based on the latter, which is founded in the belief that active emotional relationships make for easier and better access to emotional authenticity.

My animation knowledge has primarily been self-trained but draws upon a variety of sources, including: the electronic art concentration within the Master of Fine Arts programme at the University of Cincinnati; the Mastering Maya: Advanced Digital Production online class through 3DBuzz.com; character and creature animation courses through AnimationMentor.com; as well as years of experience working directly with the Motion Capture Lab at Auckland University of Technology.[52]

Being both the subject and the researcher raises certain issues. For instance, I am unable to witness my own performance as it is created; instead, I have to play back the recorded footage through a camera or monitor. This prevents certain improvisations that a director may use to modulate an actor's performance. It also means that instead of maintaining my concentration on acting, I have to move back and forth between a creative "acting mode" and an analytical critiquing "critiquing mode". Due to my acting training, I am used to this type of modulation and have the experience of quick preparatory exercises to get me back "into the moment" of an acting scene.

[52] My language about animation and animated performance is also heavily influenced by a variety of practice-oriented books, including: "The Animator's Survival Kit" by Richard Williams; "Acting and Performance for Animation" by Derek Hayes; "Acting for Animators" (2011) and "Acting in Animation" (2005) by Ed Hooks; "Animated Performance" (2010) by Nancy Beiman; and "Thinking Animation" (2006) by Angie Jones and Jamie Oliff.

5.4 Autoethnography

Due to the difficulties in articulating meaningful tacit and ineffable knowledge about my acting and animation process, and because I locate myself as both the observer and sole subject of this research, I supplement my phenomenological approach with autoethnography. Brigg and Bleiker state that autoethnography "[reintroduces] the self as a methodological resource", overcoming the strict separation between objective researcher and subject-being-researched in order to situate "the researcher's experience at the centre of the phenomenon under investigation." (2010, pp. 788-789). Ngunjiri, Hernandez, and Chang state that autoethnography "allows researchers to dig deeply into their own experience, including the attendant emotions in ways that may not be possible [in other ways]" (2010, p. 8). While phenomenology is concerned with how the body perceives phenomena, autoethnography is specifically concerned with how the researcher perceives and is affected by those phenomena. This research engages with an autoethnographic account of my acting process, which enables me to articulate my experience of lived emotions and acting choices. Brigg and Bleiker indicate that social science often dismisses personal physical and emotional sources of knowledge as having no place in reliable analysis. However, by openly engaging with personal vulnerability, as well as my intuition and emotions, an autoethnographic approach helps to expand the range of data available within acting and animation research (Brigg & Bleiker, 2010, pp. 794-795).

In order to expose and retrace my thinking/feeling/intuition process while acting and animating, I require an autobiographical means to recall and record my choices, such as a think aloud (TA) method[53] and retrospective self-observation (RSO). The TA method involves a constant verbalisation of a participant's thoughts while engaging with a given task. Fonteyn, Kuipers, and Grobe indicate that there should not be a pause for longer than a few seconds between vocalisations (1993, p. 434). This method is useful for describing how subjects navigate their expertise through a given domain of knowledge, as well as

[53] Also called known as a "talk aloud method".

what information they focus on while problem-solving (1993, p. 430). However, TA can also interfere with the participant's normal train of thought, and requires the participant to constantly shift their attention between doing and vocalisation of the doing (Henderson, Henderson, Grant, and Huang, 2010, p. 7). This research requires constant self-reporting while engaging with MoCap cleanup and animation, during which time I invoke a TA method to continuously vocalise my creative decisions and thought process. I use Open Broadcast Studio (OBS), a software capable of recording the computer screen and audio through a microphone, to document my full process on the computer. I draw on my experience of producing hundreds of video tutorials about animation when producing these recordings. As with those tutorials, in order to maintain continuous vocalisation, I slow down my normal working pace to allow myself time to fully explain my thoughts. I later transcribe these perceptions and synthesise the major insights into a written reflection (see Sections 6.2 and 6.5). As Ngunjiri et al. note, it is unnecessary to articulate all of my experiences and choices made during this analysis, especially extraneous experiences such as contradictions, mishaps, and dead-ends (2010, p. 796). Instead, I emphasise my key choices so that a reader can retrace my steps of retargeting acting data onto a digital character. The TA method provides insights into how I encounter my acting through the lens of being an animator, and how I use both sets of experience/knowledge to produce the final acting product in my digital counterpart.

Prior to recording my dramatic acting performance (Section 6.5), I enlist the help of Dawn Glover, an acting coach who teaches as part of the Kacie Stetson Studio, to direct me through my chosen monologue for the scene. She coaches me through the monologue for one hour during the week prior to recording. On the day of recording, she is present to direct my performance and function as my "scene partner". This helps to provide the presence of another character who I address but who remains silent during this part of the script. I create as many takes as I feel are necessary to generate an *affective performance* – that is, a performance that clearly communicates my acted experience to a viewer's empathic experience. Following each take, Glover provides me with feedback and techniques for how to improve or modulate the performance. I create takes until I produce an affective performance, which I gauge by the degree to which both Glover and I feel an empathetic response to it. At this point I immediately engage in the RSO method described by Henderson et al. (2010). This type of stimulated recall involves participants reporting on what they had been "doing, thinking, strategizing, and/or feeling during a past activity." (2010, p. 8). This process relies on the participant's memory and an artefact reminder, such as a video of the activity, and its reliability is improved the sooner the recall is performed following the activity (2010, p. 8). While I engage with RSO, I watch a playback of the performance while I video record myself verbally articulating the acting intentions I used for the recorded take as well as what about my performance I feel works, what does not, and why.

In a project where I am a subject of my own research, it is important to engage with outside critique. Welby Ings notes that "[w]ithout some form of external feedback, purely self-referenced processes can result in designs that fail to explore a wealth of available options or fall short of their communicative potential." (2013, p. 680). It is for this reason that I seek Glover's coaching before and during the recording of my performance. However, this accounts for feedback specific to just the live-action component of this research. I further enlist Glover for feedback about how well the finished digital counterpart performance matches the emotional and performance cues from my live-action performance. Specifically, I ask her to perform a close read of the digital counterpart's performance to determine, in her opinion, what aspects of the original performance carry over and, importantly, which do not. While her feedback would no doubt be useful in further modifying the digital counterpart's performance, this research is concerned more with how we see and understand acting in performance capture rather than in producing a flawless PeCap result. Ings cautions that the research produced through self-inquiry can become confused or disconnected if the researcher feels pressure to adjust the work based on the input of different external perspectives. Instead, he advises that the researcher does not switch from a subjective search to an objective analysis of the creative product, but instead applies the critique to generating a fuller understanding of the available options and communicative potential of the research question (2013, p. 680).

Acting choices are a highly personal form of performance epistemology and the contents of the creative choices of both actors and animators are often underreported. When recalled after the fact, the moment-by-moment acting choices may become reconstructed, substituted, or forgotten. TA and RSO help to capture my clearest and fullest memories of my acting choices so as to mitigate the risk of them becoming altered with time. Through an autoethnographic approach, I engage with discovering whether/how my moment-by-moment acting choices inform my process as the animator responsible for translating the PeCap data to the digital counterpart model.

5.5 Ethical Considerations

Additionally, there are ethical considerations for using myself as the sole actor in this research. This is due to the nature of my acting process using the Chubbuck Technique combined with Stetson's acting model, which calls for emotional authenticity and transparency from the actor. For example, based on my training, if the scene I perform calls for the character to experience anger, then I need to find the person in my life who, at that moment, most strongly activates the emotion of anger within me. I employ various preparation exercises through imagining myself engaging with that person, with the goal of getting my mind and body into a state by which I genuinely feel angry. At that moment, I begin acting the scene. Naming the person who causes me to experience a particular emotional state requires vulnerability on my behalf (Brigg & Bleiker, 2010, p. 794), but it also makes that person an unknowing participant in my research. My preparation exercises are not intended to "tell the truth" about my relationship with that person so much as they are to engage with how I feel about my relationship with that person. As such, these exercises subjectively frame a view of that person (Ings, 2013, p. 682). While the person may never be named in the acted performance itself, this research involves documenting my entire acting process, including the emotional preparation exercises. As such, I edit out any reference to the names of the "unknowing participants" in my documentation. I also edit out any information that could potentially reveal their identities to anyone who is sufficiently familiar with them. By revealing this autobiographical material to the public domain, I also become vulnerable to an audience who may understand my acting process as a "true" account of my life, rather than a reality formed by selectively emphasising specific emotionally charged and highly subjective moments outside of a broader context (2013, p. 681).

5.6 Interviews

This research features two interviews with experts in their respective fields: the actor Andy Serkis and Weta Digital VFX Supervisor Dan Lemmon. In both of these instances, there were no requirements for seeking ethics approval from Auckland University of Technology to carry out these interviews. This is due to an exception to activities requiring Auckland University of Technology Ethics Committee (AUTEC) approval as outlined in Section 6.6 of the Ethics Approval Guidelines and Procedures:

> 6.6. A one-off interview of limited scope and depth with professional persons, authorities or public figures, (for example. politicians, scholars, prominent authors) in the area of their expertise. Such an interview must not comprise a major study. (Auckland University of Technology, 2020).

The information derived from these interviews supplements the thesis but does not comprise a vital part.

6
PERFORMANCE CAPTURE AND CONSTRUCTION

6.1 The Role of the Human Actor

The academic literature surrounding acting and performance in general, and for animation specifically, has limited meaningful accounts from actors and animators about their direct artistic experience.[54] Especially in regard to 3D animation and motion capture, the majority of the academic literature is written from an "outsider-in" perspective: animation theorists are rarely animation practitioners, and therefore do not have the direct experience of making animation to draw from in their analyses. Likewise, professional animators are rarely academic theorists, and there is little inclination or incentive for animators to discuss their process in detail at the level of academic analysis. The existing literature often provides a meaningful contextualisation of ideas, but it rarely articulates the pragmatic and creative experiences of the animators themselves[55]. Throughout this chapter, I contribute a thorough discussion of my process as both an actor and an animator in order to attend to this need.

Actors who perform through motion capture find that their acting processes largely remain the same as for stage or film. However, there are considerations specific to MoCap that actors frequently point out. For instance, when working with PeCap, actors are often requested by the PeCap technicians to participate in a preparatory phase before acting begins. This phase involves a preliminary capture of some basic facial shapes and expressions, including FACS shapes, phonemes, and possibly some of the actor's dialogue from the script and/or a range

---

[54] Among the examples cited in this thesis include doctoral theses by Hosea (2012) and Delbridge (2014). Also, the article "Behind the Scenes: A Study of Autodesk Maya" (2014) by Aylish Wood provides numerous accounts of 3D animators discussing their direct experience of (and frustration with) working with 3D animation software.

[55] To cite a few examples of non-practitioner film/acting/animation theorists who are immediate to my research, I include Vivian Sobchack, Scott Balcerzak, Barbara Flueckiger, Mihaela Mihailova, Lisa Bode, Barbara Creed, and Tanine Allison.

of extreme facial movements specific to an actor (Failes, 2017, para. 4). This preparatory phase provides both calibration data for the PeCap system, as well as reference data for the technicians and animators during the retargeting process.

Serkis places special emphasis on the importance of the breath in MoCap. Certainly, breathing is fundamental to the actor's connection with his or her emotions and body state. Serkis notes that he is able to create a clear sense of a character's internal rhythm through focused breathing exercises – a process he uses to embody his characters, especially non-human ones like Kong. He works through this type of physical manifestation of the character first, and once he finds that rhythm, he delves deeper to find an internal motivation for his character. PeCap roles frequently involve creatures from sentient aliens to terrestrial beasts, which means that PeCap actors are even more likely to encounter aspects of animal behaviouralism in their acting processes. Serkis frequently discusses how physicality factors into his roles. While a basic understanding of a character's physicality underpins any role, MoCap may more frequently require an actor to have a stronger physical connection to a role – to work as much with physical externalisation as with internal motivation (Balcerzak, 2013, pp. 199-200). This does not mean that the actor should over-gesticulate. Serkis explains that when he created his very first MoCap performances for Gollum in "The Two Towers", he felt compelled to produce a broader sense of movement consistent with his supposition that animated movement should be broad. However, Serkis quickly learned that more contained movements actually performed better through his motion-captured character. He notes that "[p]eople assume that performance capture is about a lot of movement […]. The thing I've learned is that it teaches you to use incredible stillness. You learn to trust yourself as if you were playing a conventional live-acting role." (Hiatt, 2014, para. 2).

A focus on the breath is important in MoCap in other ways. As the research of Bloch et al. indicates, an actor's pattern of breathing reflects the tension within his or her body (1987, p. 12). A common problem for actors is unregulated tension throughout the body as a whole or within specific regions. Strasberg noted that such tension inhibits an actor's ability to think and feel. A more relaxed body allows for natural impulses generated through thoughts and emotions to "pass without interruption into pure expression." (Hethmon, 1965, p. 88). The viewer is sensitive to an actor's tension, both in terms of the actor being physically rigid and emotionally stifled (insincere). MoCap technologies are especially sensitive to recording body tension.[56] Since motion-captured characters already run a high risk of uncanniness, any unintended body tension further increases the likelihood that a digital counterpart may appear uncomfortably strange.

Performing on a motion capture stage generally has more in common with theatre, especially blackbox theatre, than live-action filmmaking. This is particularly true of productions that utilise virtual cameras to produce the cinematography in postproduction, allowing actors to perform throughout the MoCap space without concern for the restriction of movement common to traditional filmmaking (Duncan, 2010, p. 86). Just as in stage acting, all of the actors in a motion-captured scene perform alongside each other throughout the entirety of the scene (Bestor, 2016, p. 177) – unlike in film where actors in the same scene may only be present while they are directly recorded, and in the case of some productions, may never meet together in the same scene at all. Reflecting on the making of "Beowulf", Anthony Hopkins, who plays King Hrothgar, was pleasantly surprised at how quickly the production proceeded to recording the actors. According to Hopkins, this was unlike his experience of standard film productions that require two weeks of rehearsal before filming commences (Roberts, 2007, para. 57). Director Robert Zemeckis also notes that the actors were surprised at how quickly the recording process was completed. Each take was shot continuously without concern for lighting or camera resets, allowing the actors to concentrate on pure performance. This allows the actors to "dictate the rhythm of the scenes […]. It was like theater, except that it was being captured in 3D." (Billington, 2007, paras. 2-3). Of course, this was a completely virtual production that resulted in ex-filmic performances; it therefore

[56] As someone who suffers from chronic back pain that leads to a great deal of body tension, I am constantly aware of how this carries over into the MoCap data I produce, and therefore into the digital characters onto which my performance is applied. Extra care must be taken in loosening the actor's body of unwanted tension or else it becomes embedded in the spectral trace of the actor's performance, and can be difficult for an animator to attenuate.

does not fully reflect the realities of working on films that utilise visual effects integration strategies.

While the use of virtual cinematography frees the actors from worrying about camera placements on set, this can be a challenge for actors who are used to modulating the direction and degree of containment of their acting within a given shot composition. In a preliminary study that investigated how experienced actors who were new to working with motion capture adapted to a MoCap environment, a significant discovery was that if any cameras were used on set to record the performance[57], the actors would default to playing toward those cameras. The actors were instructed to ignore the cameras but it was evident that their instincts were to be mindful of the cameras. This was in part because the actors were unfamiliar with how a performance could be recorded separately from cinematography. While this caused some initial issues with the actors understanding how to modulate their performances, the director, actors, and MoCap team came to the conclusion that when recording MoCap on a completely virtual set, the actors should default to treating the emotional component of their performance as if they were in close-up and the physicality component as if they were in a long shot (without over-gesticulating) (Kennedy, 2015a, p. 946).

These caveats don't necessarily apply for PeCap performances that exist within a profilmic context. It is especially important that the PeCap actor maintains a strong inner monologue in order to produce close-ups that read well for PeCap. An actor's inner monologue is a continuous dialogue the actor experiences within his or her head throughout a scene (Chubbuck, 2005, p. 172). It is akin to a person's normal thought process, but the actor uses a specific thought process that enables him or her to better connect with the immediate reality of a scene. Chubbuck explains that

[57] These cameras were placed at opposite ends of the MoCap studio and were used to record the actor's performances to provide reference for animators who would later clean up the MoCap data. The reference is necessary to assist MoCap artists and animators whenever there are errors in the data, or if the MoCap data is unclear for any reason.

> [g]reat actors always have a strong inner monologue playing behind their eyes at all times. […] The inner monologue plays out on their face, in their eyes, in the way they hold their bodies, their hands. Frequently, the actor who is using inner monologue displays more need and passion than the actor who is speaking. (2005, p. 194).

Inner monologue helps an actor to produce clearer external signs for the PeCap system to register, increasing the likelihood of producing an emotionally connected (or at least sincere) performance.

A further challenge with motion capture is in how actors interact with costumes, props, and their environment. A traditional actor will be dressed to suit his or her character, providing a set of tactile and visual details that many actors use to further ground themselves with the role. In my own experience, I find that the type of shoes I wear greatly impacts how I feel about my character, and in turn how I perform through that role. However, with MoCap in general, actors must substitute costumes for MoCap suits – generally form-fitting Lycra unitards or two-piece combinations. These suits either allow for tracking markers to be Velcroed to their surfaces, or have tracking sensors embedded within their fabric. While this provides optimisation for the technology that records the performance, it sacrifices the actors' personal connection with their costumes. The suits present actors who are new to MoCap with several challenges. The passive optical markers stick out from the surface of the suit by about 10mm, and it is common for them to become dislodged when actors make physical contact with each other (or during physically demanding actions). The markers may fall off the actor, switch location on the actor's body, or, with surprising frequency, *jump* from one actor to another when bodies contact. In the preliminary study mentioned above about actors adjusting to the MoCap space, the actors reported that they were quite aware of the physical limits of their MoCap suits and this led them to being less willing to engage with compassionate actions, such as sliding contacts, embraces, or maintaining general intimate proximities. For instance, if the actors needed to hug in a scene, their concern for disrupting the suits led them to pat each other instead, which communicated no real sense of intimacy or emotional sincerity. The actors reported that while they were still acclimatising to their suits, awareness of the novelty of the suits inhibited them from being fully present in their acting. The actors, none of whom were technologically proficient, felt pressure to perform within the bounds of the technology rather than trusting their performances would survive any technical mishap. They reported being overwhelmed by being surrounded by and performing through so much expensive and sophisticated technology. They conceived that any minor disruption to the default marker setup affixed to each of them would render a take unusable, much in the same way that an out-of-focus camera would ruin an otherwise excellent live-action take. However, once they were reassured that even a couple of dislodged tracking markers wouldn't necessarily harm a performance, their interactions became more natural. Likewise, when performing with a Vicon Cara PeCap helmet, the actors had to contend with cameras and LED lights mounted directly in front of their faces. The degree to which the camera arms stick out in front of the face also made it a challenge for the actors to get their heads physically close together. The study found that experienced actors working in MoCap for the first time generally require at least 30-60 minutes to engage with the MoCap space, suits, and equipment before they are as equally present in their acting as they would be for stage or camera (Kennedy, 2015a, p. 945). This is consistent with James Cameron's observation during the production of "Avatar" that the actors found the PeCap headsets "distracting for the first couple hours, but soon they started loving it and having fun." (Riley, 2010, para. 8). Delbridge rightly points out that the MoCap environment is an ecology of different creatives working alongside each other (animators, actors, technical operators, fight scene directors, etc.), but there is often "an absence of common knowing that participants are a collective part of." (2014, p. 121). All of the contributors in this environment need to ensure their input is heard in order to get what they need from the production. This research, in part, helps to articulate and record the experiences of different artistic contributors across the motion capture performance pipeline.

The use of props is a vital aspect of an actor's performance. Chubbuck refers to "the handling of props to produce behavior" as "doings", and indicates that "[d]oings are the physicalization of our intentions through the use of props." (2005, p. 142). Through doings, the actor engages with more natural behaviour that is guided by his or her emotions and intentions (2005, p. 143). However, the props used on MoCap sets are often substitutes for real items. For instance, a character who is meant to wield a sword may be provided with a mock sword created from a few pieces of wood glued together. Beyond just concerns for health and safety, the rationale for this is technological: the lights that reflect off a metal sword could be confused by the system for the retroreflective tracking markers. A wooden sword painted with a matte finish will not produce such sharp reflections and therefore not pose a tracking hazard. Additionally, props are often motion-captured in order to better integrate with the source-actors' motion. The design of some authentic props may be unsuitable for affixing tracking markers, and therefore approximate substitutes may function better for the sake of the technology. However, just like with costumes, an actor forms an immediate relationship with his or her props, and this is strengthened by a sense of a prop's genuineness. When acting on a bare MoCap stage, the actor must already engage his or her imagination to create the environment, and must further imagine inferior prop substitutes to be the real items.[58] Kade et al. advise that directors provide MoCap actors with as much visual reference about their environment, costumes, and props as possible in order to align the actors' imaginations with the directors' intentions. Additionally, Kade et al. note that among directors surveyed in their research, there are four attributes that directors commonly expect from a good MoCap actor:

1) Solid acting skills and the ability to express emotion throughout the whole body;
2) A good standard of physical fitness (such as an athlete, stunt performer, or dancer);
3) Creativity and possessing a good imagination; and
4) Familiarity with the language and systems of the MoCap environment (2013, p. 284).

Again, the emphasis on a good imagination mixed with a creative instinct is imperative

[58] Actors who are familiar with working in a blackbox theatre environment will likely be used to working in such an aseptic environment, and they will rely on similar tools of the imagination to fill out the MoCap space as they would the blackbox stage. Likewise, actors with experience of filling in the missing pieces on green screen or blue screen sets will likely be at home with the imaginative element of working with MoCap.

to feeling present in an environment that does not actually exist. Menache echoes the importance of an actor being familiar with the intricacies of a MoCap environment, especially in terms of acting within the boundaries of the capture volume, choreographing and repeating motions that are intended to be blended together (especially when recording MoCap for games), replicating ROMs, and starting and ending performances in a T-pose (2011, pp. 111-112). However, in the past decade, it has become more common for filmmakers to create approximations of set pieces to help actors better navigate and relate to their space. In order to minimise marker occlusion, these set pieces are generally see-through and designed from metal grids, simulating walls, doors, chairs, tables, and other objects that the actors may directly interact with. This also helps to contain the action within a specific environment should the environment be smaller than the capture volume – so as to ensure a character doesn't walk through a virtual wall or intersect a countertop, for instance. These set pieces are later replaced by virtual counterparts that correspond to similar measurements and placements within the finished digital scene (Pizzo, 2016, p. 51).

Finally, MoCap actors are sometimes equipped with prostheses to accommodate physiological differences between themselves and the characters or creatures they play. For instance, the actors playing the digital orcs in "Warcraft" were given custom dental appliances that pushed the actors' lips out in order to accommodate the effect that the orcs' tusks would have on speech and facial performance (Robertson, 2016, p. 39). Likewise, in the recent "Planet of the Apes" trilogy, the actors playing apes wore prosthetic fangs to distort their faces, as well as limb extensions that accommodated for the proportional difference between human arms and the arms of various ape species. These limb extensions ensured that the human actors produced movement akin to the loping, four-limbed gaits of their simian counterparts, rather than trying to restructure human movement into ape-like movement through animation alone. While these prosthetics are visible during the filming process, they simply become a part of the character once the motion data is retargeted onto the digital counterpart (Bode, 2015, p. 89).

In the next section, I describe the process for creating my first practical PeCap experiment in this research. Using language and anecdotes derived from this experience, I articulate the roles of the MoCap artists and animators in the ensuing sections, followed by a description of my process for creating a more advanced PeCap experiment. The insights gained from these experiments are supplemented and supported by interviews with a leading MoCap actor (Andy Serkis) and an Oscar-winning visual effects supervisor (Dan Lemmon).

6.2 Simple Acting – Basic Object Exercise

As the starting point for my inquiry into the nature of acting in PeCap, I produced a simple and realistic acting performance to elucidate what an actor does within a PeCap environment. This performance also clarifies what PeCap artists and animators do with an actor's performance in order to create a similarly affective performance in a digital counterpart. By affect, here I refer to Massumi's translation of *l'affect* in Deleuzean terms:

> *L'affect* (Spinoza's *affectus*) is an ability to affect and be affected. It is a prepersonal intensity corresponding to the passage from one experiential state of the body to another and implying an augmentation or diminution in that body's capacity to act. (Deleuze and Guattari, 1987, p. xvi).

I hypothesise that the transference of affect is directly related to the production of a similar sense of screen presence between the source actor and his/her digital counterpart.

Drawing upon Kirby's continuum of acting, this research is concerned with dimensions of acting that involve pretence, including emoting, simulating an action, or impersonating a character. Within Kirby's continuum, only simple acting and complex acting are based on pretence. Simple acting involves the actor engaging in only one form of pretence, while complex acting involves multiple simultaneous forms of pretence (1972, pp. 8-9). For instance, an actor experiencing/performing an emotion *and doing nothing else* would be considered a form of simple acting by Kirby's definition. However, Hooks believes that in terms of acting, "[e]motion by itself has no theatrical value. […] It is not even acting. Acting is *doing* something. It is doing something in pursuit of an objective while overcoming an obstacle." (p. 20). While Kirby's theory is acting-style agnostic, Hooks' position is based on a realist style of acting. These acting experiments form the basis of inquiry into if and how a source-actor's sense of sincerity (or emotional connection) can be preserved and transmitted through a digital counterpart. That is, how is sincerity constructed – by the actor(s), PeCap artist(s), and animator(s) – in the performance of a source-actor's CG counterpart?

Due to the complex nature of working with the PeCap system, performing the acted product, dissecting the acted result, and interpreting it through the animation software, it is important to produce a type of "baseline" performance that can serve as a guide for refining my process when engaging in more complex acting later in the process. Animator Tom Sito recalls a quote from Russian playwright Anton Chekhov: "On stage, the hardest thing to do is nothing, in character." (2013, p. 210). He relates this quote to his experience of watching performance-captured characters who he sees as producing believable performances when they engage in action, but who struggle to produce sincere and affective performances when they act from a fixed position (2013, p. 210). The practical experiments I conduct in this research specifically address this concern and reflect on how a performance-captured character may elicit empathy from an audience similar to its flesh-and-blood counterpart.

6.2.1 Approach to "Simple Acting"

As a first step toward understanding how acting translates from source-actor to digital counterpart via PeCap, I undertake a "ground zero investigation" of acting. That is, based on Kirby's definitions of acting, I produce a "simple acting" exercise whereby there is only one dimension of acting to keep the initial experiment as simple as possible. The acting in this exercise is simple by virtue of focusing solely on me making breakfast, and does not attempt to engage with any other dimensions of acting, including emotions or pretending to be someone other than myself. By doing so, I investigate how small gestures and minute facial movements translate from actor to digital counterpart. Since nuanced movements are essential to conveying a sense of realism in the acted product of a photoreal CG character, it is important to be able to isolate and examine these actions without distraction or interference from broader movements or extreme facial expressions.

Hagen provides a compatible acting exercise for this initial experiment. As I discuss in Chapter 5, my personal acting training largely comes from the Kacie Stetson Studio, which uses Chubbuck as a basis for its instruction. Chubbuck's techniques inherit much from Hagen's teachings. Hagen and Frankel outline a preliminary acting exercise that is both multidimensional and uncomplex – and perhaps as close to "simple acting" as I can get while still considering Hooks' admonition that emotions by themselves do not constitute acting. This is known as the Basic Object Exercise (BOE) and involves recreating two consecutive minutes from the actor's life, not in crisis. This performance is rehearsed and recreated "as if for the first time", with the goal of "convincingly [creating] two minutes on stage in which you exist as if you were alone at home" (1973, pp. 82-89). This exercise also accommodates my prerogative to be the sole actor (and animator) for my practical experiments: "The only difference between the [BOE] and a scene is that you will be using yourself instead of a character, and your life experience instead of a play." (1973, p. 86). By foregoing a script, this exercise eliminates the acting dimension of interpretation (1973, p. 87) and ensures the performance is fully personal, which is an important factor when reflecting on the process. By definition, PeCap involves more than just body movement, and so it is important that my "ground zero investigation" of acting in PeCap also attends to facial performance.

During the PeCap session for the BOE, I recorded video and audio of my performance in addition to the PeCap data itself. I placed HD video cameras at each of the four corners of the motion capture volume in order to maintain a clear view of my physical performance at all times. I captured audio with an omnidirectional microphone, which provided clear audio recordings regardless of my body's orientation. While my BOE performance did not involve speaking, the audio helped to record any natural body sounds (grunts, coughs, scratches, etc.).

I endeavoured to perform as many takes as I felt were necessary to generate an *affective performance* – that is, a performance that would clearly communicate my acted experience to a viewer's empathic experience. After overcoming issues with the PeCap equipment colliding with parts of the set in the first take, I felt I achieved an affective performance in the second take. Immediately following the second take of the BOE, I engaged in the RSO method and verbally articulated my acting intentions and what I believe worked well in the take. I video recorded myself engaging with stimulated recall to identify what about my performance I felt worked, what did not work, and why.

6.2.2 Summary of the BOE Performance

The BOE performance I developed involves recreating my morning ritual of preparing breakfast: a bowl of Sultana Bran with added raisins, along with juice and milk. In the scene are various props and furniture arranged to mimic the layout of my kitchen at home (Figures 21 & 22). These include:

1. A table (approximating the height and location of a cupboard) on which a box of Sultana Bran and a box of raisins are placed.

2. A bar fridge on top of a rolling platform. The fridge door sits at a height that approximates the actual height of my full-size refrigerator at home. This height means that when I reach for the milk and juice bottles on the fridge door, I am reaching at approximately the same height as I do during my morning ritual at home. Additionally, I added a grey cloth sheet to cover the back of the bar fridge because the mocap cameras detected reflections from its surface. Such reflections can be mistaken for tracking markers by the MoCap system and result in errors while recording. The cloth sheet eliminated this problem.

3. A stool with a tall box that sits lengthwise across its seat. This setup is meant to approximate the height of my countertop at home. On top of the box was a ceramic cup, ceramic bowl, and wooden spoon. Each of these items were chosen because of their matte finish, thus reducing the potential for tracking errors.

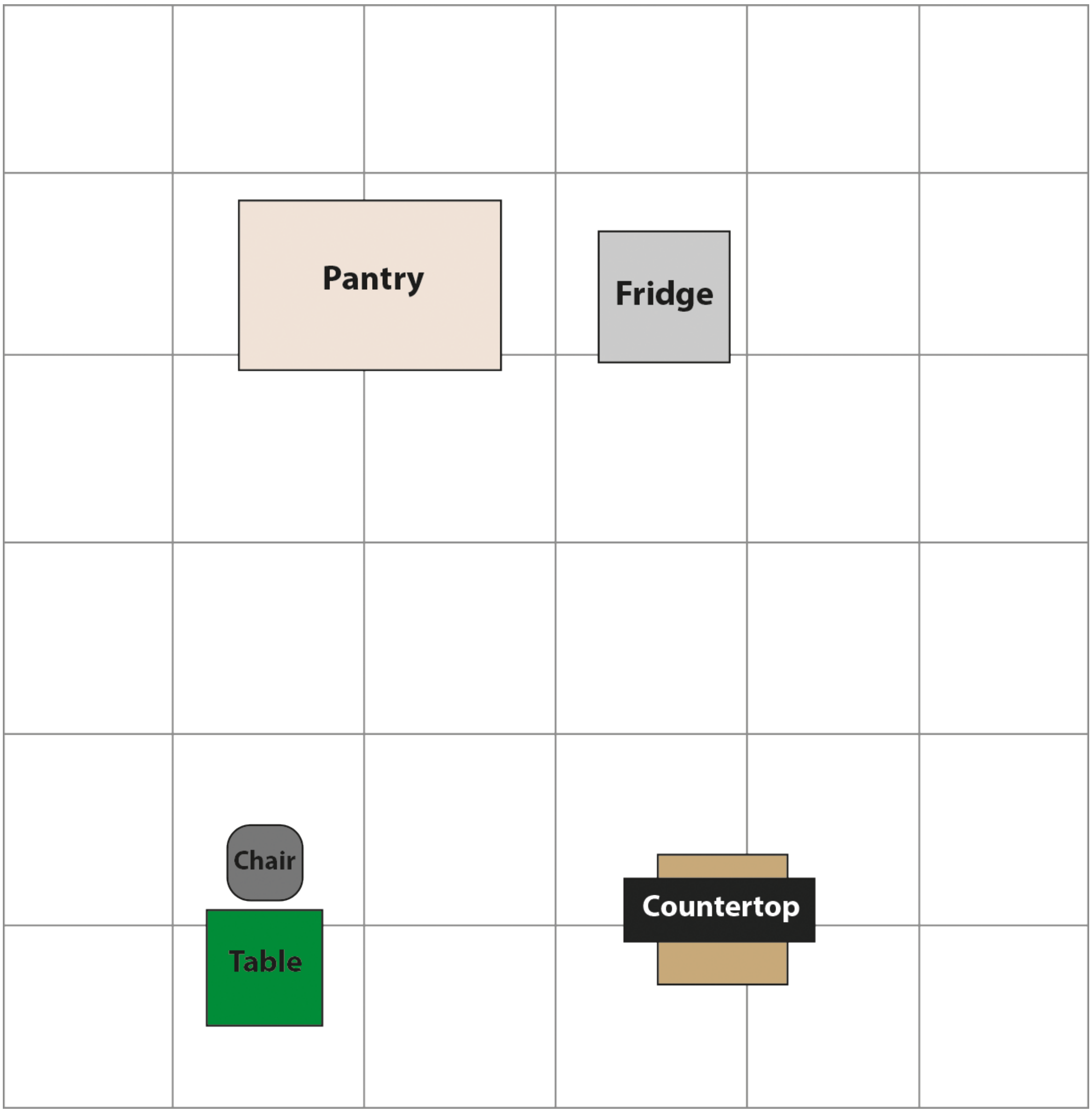


Figure 21. Diagram of the layout of furniture within the 6m×6m motion capture grid for the Basic Object Exercise (BOE) performance.

Unfortunately, it was not feasible to install a taller fixture to simulate the height of an actual cabinet above the countertop. As a result, I could only interact with the cup, bowl, and spoon on the countertop.

4. A table and a computer chair on casters. I removed the back support from the computer chair to minimise occlusion of the tracking markers on my back while sitting on it. The table is also relatively open, maximising space for the MoCap cameras to detect as many tracking markers as possible.

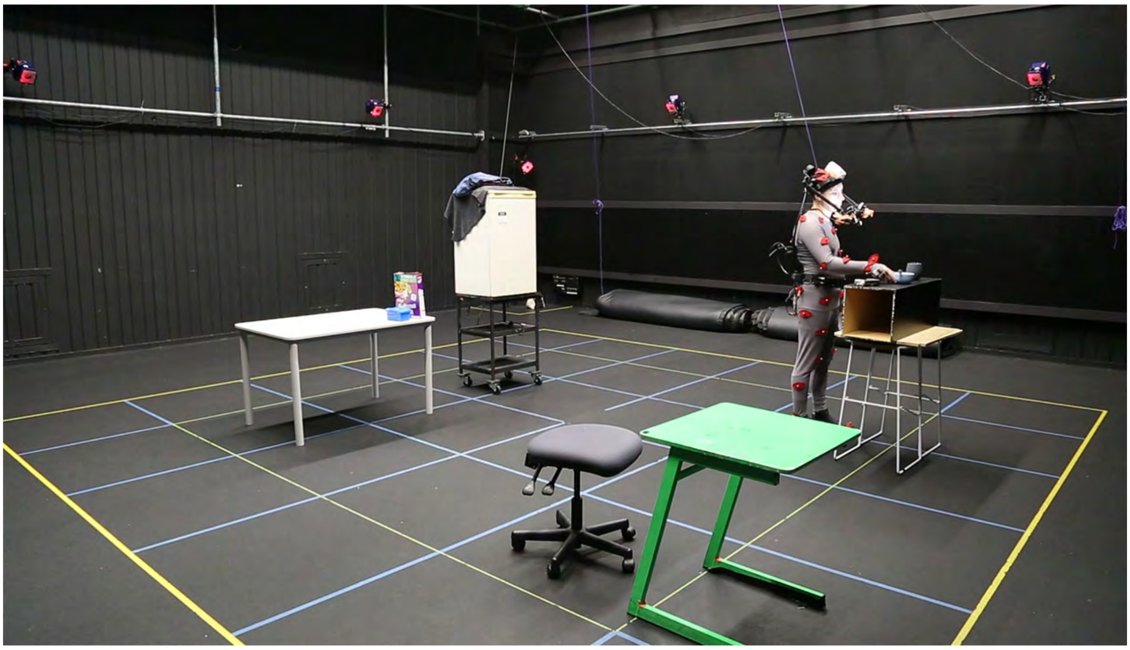


Figure 22. Screenshot of the setup for the BOE performance. Author's photo.

The major events of the scene are as follows:

1. Enter the scene near the pantry and cross to the countertop.
2. Move cup and bowl closer to me on the countertop (this is an attempt to simulate pulling these items out of a cabinet above my head at home).
3. Cross to the pantry to grab the boxes of cereal and raisins.
4. Set box of raisins on countertop and open the box of cereal I hold in my hands.
5. Pour cereal into bowl and close the box of cereal.
6. Return the box of cereal to the pantry.
7. Move back to the countertop, open box of raisins, and place several handfuls of raisins into the bowl of cereal.
8. Close box of raisins and return it to the pantry.
9. Cross to the fridge, open the fridge door, grab bottles of milk and juice, and close the fridge door while crossing back to the countertop.
10. Place milk bottle on the countertop and shake the bottle of juice in my hands before pouring juice into the cup. Close juice bottle and place it on the countertop.
11. Open milk bottle and pour milk into cereal

```
    bowl.
12. Close milk bottle, grab juice bottle, cross
    to the fridge, open fridge door, and place
    both bottles back inside it.
13. Close fridge door and cross to the countertop.
14. Grab the cup and bowl of cereal and cross to
    the table.
15. Place cup and bowl down on table while sitting
    on chair.
16. Scoop cereal onto spoon and prepare to bring
    it toward my mouth.
```

Prior to the first take, I reflected on my mind and body state. Earlier in the day I experienced anxiety due to some issues with getting into the MoCap studio, but I felt grounded and present by the time recording began. I noted that I was eager to complete the performance because it had been delayed by ten days past the day I expected to record. I also commented that my mouth movements felt more restricted than normal due to my awareness of the facial tracking markers around my mouth. I was concerned with whether the markers would fall off my face and I compensated by limiting my mouth movement in general. This was a psychological barrier that I recognised I needed to overcome, and I indicated that I would see how far I could push my expressions before any markers would fall off. In the end, I did not encounter any marker loss and I was able to relax into performing without being concerned about this issue.

One of the issues encountered during the first take was unexpected collisions between scene props and the Vicon Cara camera arms in front of my face. This also led to making it impossible for me to eat the cereal in a natural manner. During the first take I was unaware I would run into this issue, so part of the performance in the moment of eating became me trying to figure out how to move the spoon around the camera arms. I quickly realised the futility of this, but this disrupted the flow and intention of the BOE performance. I considered whether to simply mime the action, but such a solution would break the simple acting exercise by introducing an additional element of pretence. If for no other reason, this problem warranted recording an additional take.

The second take of the BOE performance was an improvement on the first take in regards to mitigating issues where various props collided with the bracer arms of the PeCap helmet. With those concerns removed, I noted that I felt like the BOE reflected my normal morning experience of preparing breakfast – the sole exception being the presence of the heavy and conspicuous PeCap helmet. By the end of my second take, the weight and compression from the helmet resulted in a headache. Despite the headache, I remained focused on the individual task actions in a similar manner to my standard morning routine.

The second take took two minutes and ten seconds to complete, at which point I immediately recorded a retrospective self-observation about my experience. In the RSO, I reported a few differences between my normal routine and the performance within the MoCap studio: "The one thing I noticed was… Of course, I don't have a cabinet up here anywhere to pull stuff down, so I'm just kind of pulling the bowl and the cup closer toward me." I then refer to the props I used to simulate a countertop, which consist of a tall box that sits lengthwise on top of a stool: "Because [the box] sits high right here, which is meant to sort of approximate a countertop, I was just trying to make sure I wasn't losing or placing the juice anywhere down lower than that, because there actually is a ledge down here." By "down here", I refer to an area on top of the stool that is not completely covered by the box sitting on top of it. It was possible to place the juice bottle on the exposed area, but this could have resulted in the bottle and the markers on my hands becoming occluded from some of the MoCap cameras. As a result, I was unable to fully recreate my experience of preparing my breakfast at home because I had to accommodate certain prop and technological limitations within the MoCap space. However, this discrepancy is relatively minor and did not seem to disrupt the flow of my experience with the BOE: "There were a few little actions that were probably somewhat more artificial than usual simply as a way of navigating around [the *ad hoc* countertop space]. But I didn't feel like that really had too much of an impact – but it is something to pay attention to for later on when I deal with more dramatic scenes. For the most part I was able to pour the juice and pour the milk and all that stuff the same way I normally would."

One other problematic prop was the fridge. During the first take when I tried to open the fridge door, the door collided with the camera arm on the Cara, which jolted the footage. This not only lead to a slightly disrupted performance but also may have caused issues with generating an accurate recording of the performance during those moments. I was conscious of this problem when I approached the door during the second take: "The second time when I went to the fridge door, I didn't knock anything, which was good. I kind of realised to give myself more room. And I feel like all of this has been done without the sacrifice of the performance itself." The additional amount of room I gave myself was minimal and sufficient just for the purpose of avoid a collision; therefore, the resulting performance does not look artificial or strange during this moment.

6.2.3 The BOE Performance Clean-Up Process

During the recording phase, the Faceware camera was placed along the bracer arm for the Vicon Cara rig (Figure 23). This was an attempt to simultaneously record both Faceware and Cara data with the goal of determining whether one recording system was better than the other at capturing and transferring subtle facial performance. In addition to retroreflective face markers glued to my face, I also affixed a series of white dots to my face to serve as tracking markers for the Cara. One advantage of having a large range of facial markers is that they provide a clear indication of when minor movements occur in regions that may otherwise be difficult to clearly see, especially if the Faceware video pixel detail in that region is low. The markers would not inhibit Faceware's tracking system, and I hypothesised that the markers might even provide extra detail to assist Faceware's tracking.

The Cara's bracer arm sits lower than the camera arm on the Faceware helmet, and therefore the Faceware footage was recorded from a lower angle than normal. For instance, the Faceware camera should be placed so that it is nearly in-line with the tip of the actor's nose; by comparison, the camera in this experiment was located in-line with the front of the actor's chin. This resulted in a Faceware recording that constantly looks up the actor's nose. This provides good detail for nostril flares and is generally fine for the mouth, but it becomes

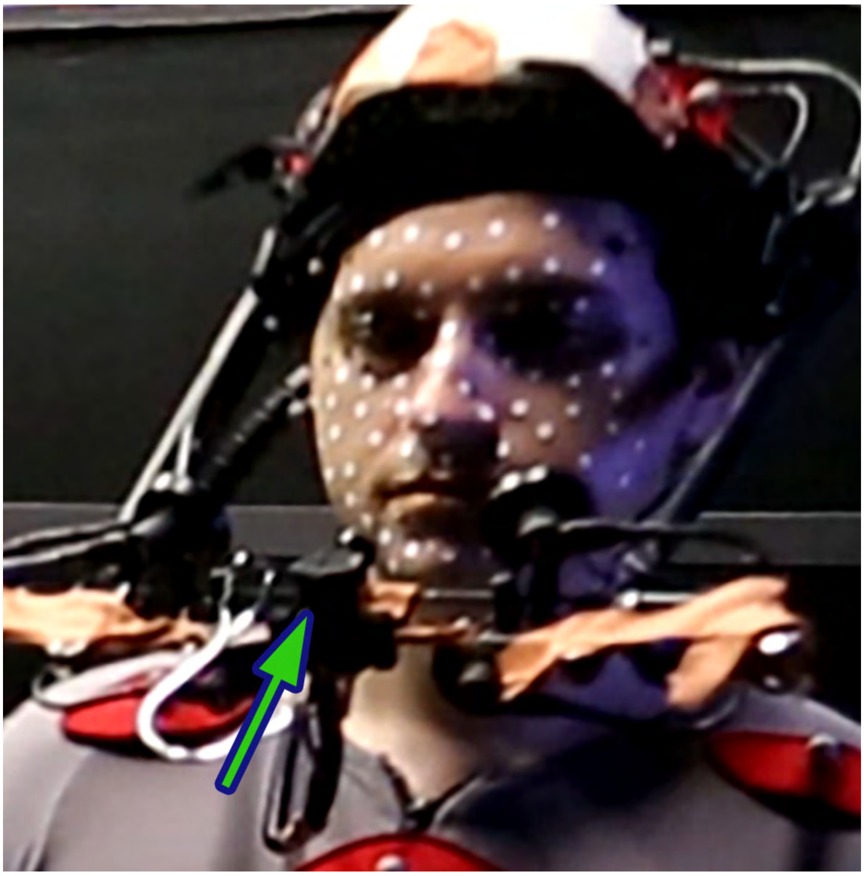


Figure 23. Screenshot of the combined Vicon Cara and Faceware camera systems. The Faceware camera (indicated by the green arrow) is attached along the centre of the bracer bar between the two cameras on each of the Cara's arms. Author's photo.

more difficult to track the performance details of parts of the face that are farther away, especially the eyes and the brows. The Faceware system is designed to interpret these regions from a more central position within the face; due to being attached to the Cara's bracer arm, the lower angle of the camera results in the actor's eyes and brows appearing smaller by proportion to the lips and nose. The lower eyelids also lack prominent curve detail, which complicates knowing precisely how open or closed the eyelids are.

Unfortunately, the data generated by the Cara rig proved unusable. Both the MoCap technician and I had the least amount of experience working with the Cara, and this resulted in some foundational issues with our setup. It became evident when attempting to resolve the positions of the tracking markers in the CaraPost software that the four Cara cameras had not been properly aligned to my face to ensure that each marker was visible to enough cameras to create reliable triangulation. The MoCap technician was later able to create a more reliable alignment of the Cara cameras, but this was long after the BOE performance. Fortunately, the Faceware system produced a usable result, but due to the lower than usual position of the Faceware camera, the performance required extra work to clean up. I was careful to accurately track the original performance using Faceware Analyzer 3.0 (hereon: Analyzer), as well as when I retargeted the Analyzer data onto the rig using the Faceware Retargeter 5.0 for Maya[59] plugin (hereon: Retargeter).

A true 3D tracking system such as the Vicon Cara would have provided more accurate automatic tracking, whereas Faceware relies on the skills of the PeCap artist and animator to accurately copy the actor's performance onto the 3D model. The Cara represents a truer sense of 3D performance capture, whereas Faceware doesn't capture performance in the same way. Instead, it records a 2D video and the performance in that video is tracked in a two-dimensional plane based on human input, which is then used to drive a one-to-one correspondence between 3D poses created by an animator. This means that the results from the Faceware system is only as good as the human input it receives at every step along this process.

6.2.3.1 BOE Analyzer Process

The Faceware system works by first tracking key points on the face using Analyzer. When tracking the source footage in Analyzer, the tracking is broken up into three regions (called "face groups" in Analyzer): the mouth, the eyes, and the brows. The workflow from here involves choosing a face group and creating a series of poses on the 3D character that resemble the facial shapes in the corresponding source video frames. I elected to work on the mouth group first. When I recorded my performance, I used a different colour of lipstick for my top and bottom lips. This was based on a hunch that the difference in colour might assist Analyzer (and the PeCap artist tracking the data) with detecting the separation between

[59] Autodesk Maya is the principle 3D software used in this research. Retargeter is a plugin that extends Maya's utility to interface with and retarget PeCap data from Analyzer.

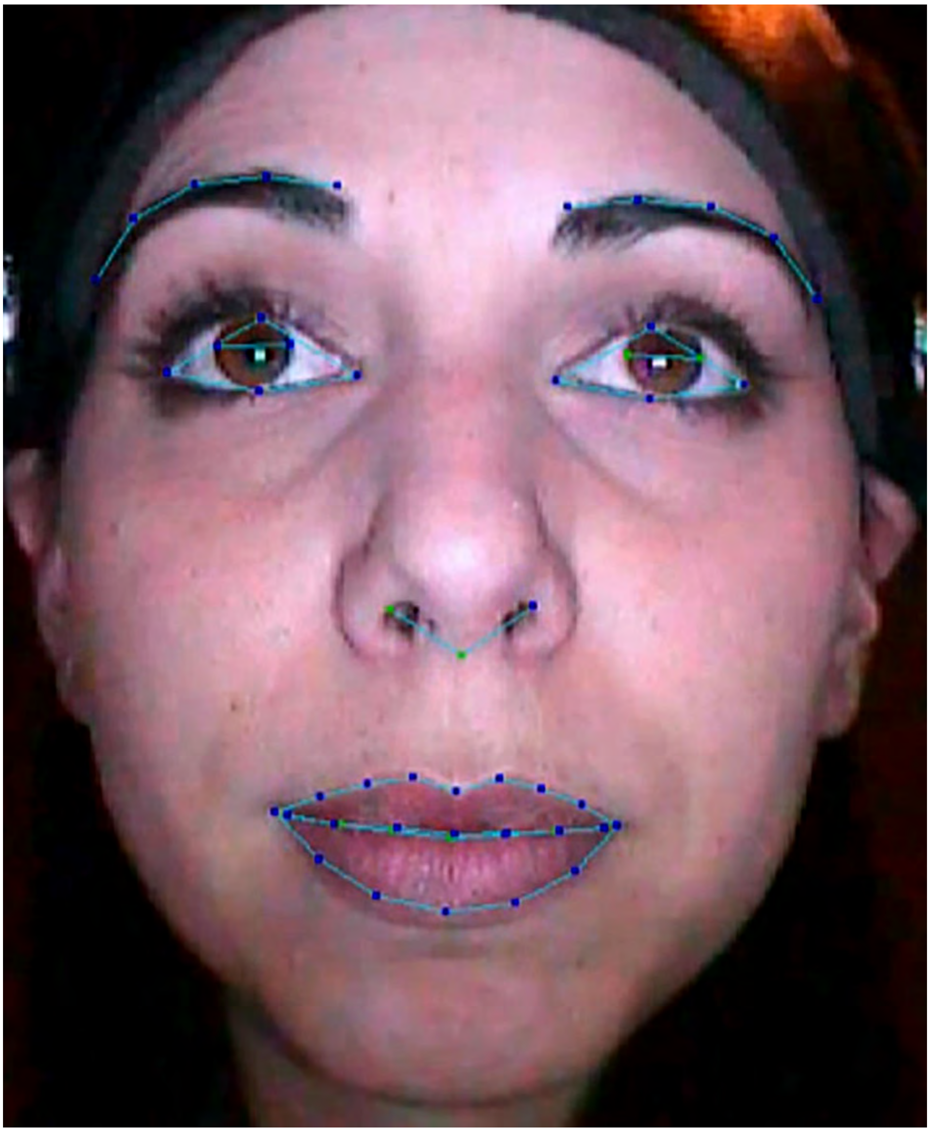


Figure 24. Screenshot of face groups in Analyzer (blue dots connected by green lines).
The outer ring of the mouth group traces the *vermilion border*.
From *How to Use 'Neutral Frames' in Faceware Analyzer*, Faceware Technologies, Inc., 2013.

the lips, especially during moments when motion blur obfuscates the lips' features and complicates individual pixel tracking. Even though the mouth possesses the most complex range of motion, I found it the easiest region to track in this exercise.

A ring of tracking markers encircles the outer edge of the lips (the *vermilion border* [Figure 24]), while a further ring of tracking markers encircles the inner edge. The outer edge remains constant, whereas the inner edge is subject to change due to the lips either pushing outward or pulling/sucking inward. This means that it is not possible to just track the

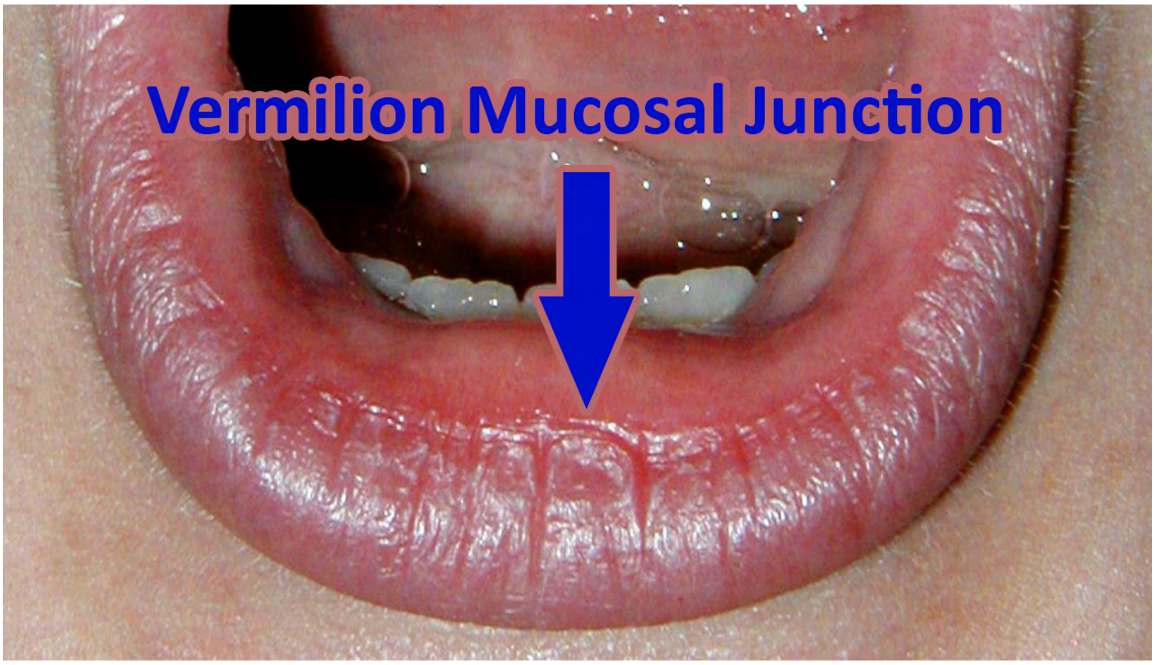


Figure 25. The *vermilion mucosal junction* identified by the zone of the inner lip where the fleshy outer part of the lips transforms into the wet, mucosal region inside the mouth.

visible inner edge of the lips because what constitutes the inner edge is not necessarily a fixed series of points on the lips. Due to the orientation of the camera to the lips, this is mostly a concern for the lower lip. In order to mitigate this problem, I use the *vermilion mucosal junction* (Figure 25) as the edge along which I track the inner mouth markers. It is critical that each marker corresponds to a consistent and conspicuous landmark on the face and should not vary from the position of the centre of that landmark by more than a few pixels over the course of the performance. Tracking the mouth becomes difficult when the mouth moves quickly enough to produce motion blur in the video, confusing the exact pixel location of a landmark. Occasionally, landmarks on the mouth change shape, especially around the *oral commissures* (the edges of the mouth where the top lip transitions into the bottom lip). This region in particular is subject to significant stretching and requires careful manual tracking to maintain an accurate fix on any chosen landmark in this region.

Each eye is tracked by four markers, one for the top eyelid, bottom eyelid, and each corner of the eye. I align the eyelid markers to the base of a conspicuous eyelash at the centre of each eyelid, and I align the corner markers to the outer edges of the medial and lateral angles of the eye (Figure 26). While the eyes have a more limited range of movement than

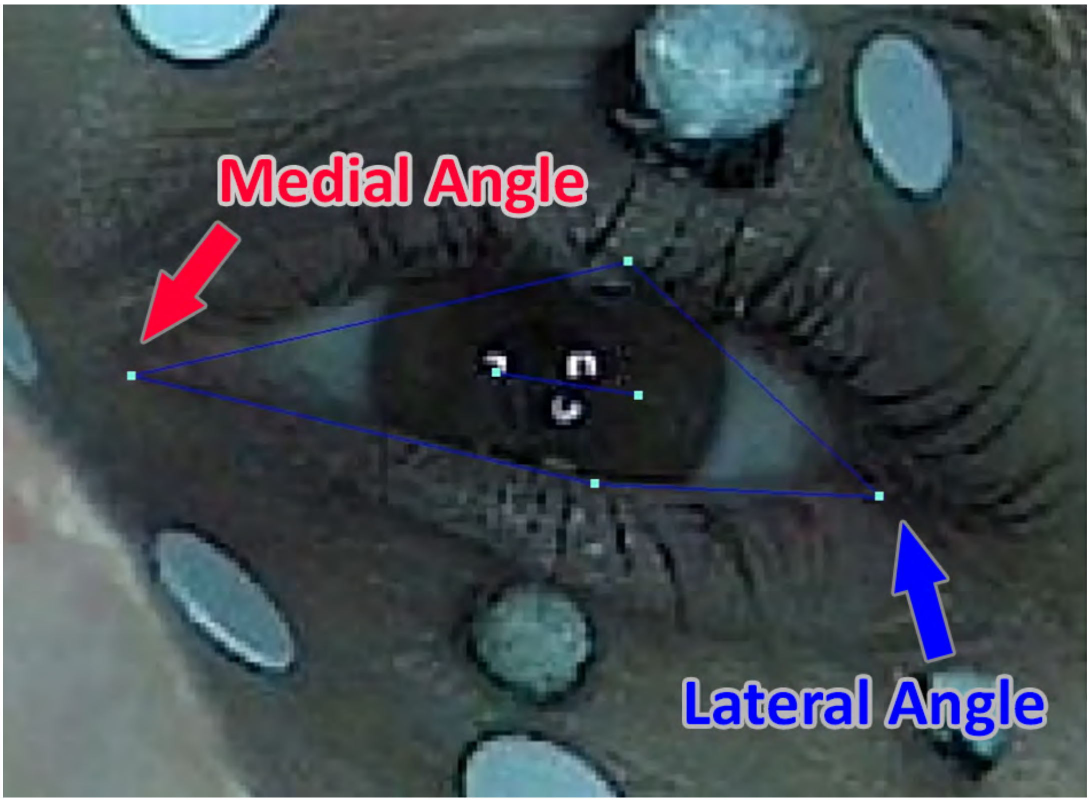


Figure 26. Medial and *lateral angles* of the eye with tracking markers.

the lips, the eyelashes often produce motion blur or occlude the landmarks associated with the eye markers, thus complicating tracking. Due to frequent opening and closing of the eyelids, the constant shape changing of the eyes provides a challenge for Analyzer to fully and consistently track the eyelids during these transitions. This requires a lot of manual correction, which is time-intensive. Analyzer also tracks the irises to account for eyeball direction, but at times this is also difficult to accurately track if the eyelids frequently close or are held shut for a period of time. These reasons all contribute to the eyelids being especially prone to data noise. In my experience, the eyeball tracking, at best, only ever produces approximations of the eyeballs' movement, and frequently needs to be replaced in its entirety by keyframe animation in Maya.

The brows are also challenging to track: while there are only five markers per brow, the outer edges of the brows can be difficult to track because they wrap around the temples of the skull, making it difficult to sight particular landmarks during quick movements. The brows don't have special features in themselves; instead, I place the markers along the top edge of the eyebrows and attempt to track conspicuous brow hairs.

6.2.3.2 BOE Retargeter Process

In order to produce an optimal alignment between the real-world recording and its digital representation, I created a virtual camera (which I named "Faceware Cam") in Maya and positioned it at the same focal length, angle, and distance to my face as the real-world Faceware camera. This method enabled me to gauge facial shapes on my 3D character from the same angle as the video of the facial performance (as opposed to looking at the reference video from one angle and trying to align the shapes to a different angle on the 3D character). As mentioned above, it is difficult to precisely see the shapes of the eyelids and eyebrows due to the low position of the Faceware camera in this experiment, and it is possible there are some small degrees of variation in these areas in the finished animation. I worked back and forth between Maya's default perspective view and the Faceware Cam view in order to judge whether a facial shape looked correct not just from the original performance view, but from other angles, as well.

The Faceware Cam is positioned so that I see the same features of the head as I do in the reference footage – that is, from the base of the chin to the top of the skull. However, when trying to create small-scale detail adjustments, on the lips and eyes especially, it is useful to have cameras that are positioned closer to those features. For this reason, I created a series of additional cameras that focus on the mouth, eyes, and brows, and named these:

- Mouth Cam
- Eyes Cam
- Left Eye Cam
- Right Eye Cam
- Left Brow Cam
- Right Brow Cam

With each of these cameras, I tried to preserve the same apparent focal length and angle to the camera as in the reference video. While working with these fixed-position cameras, I

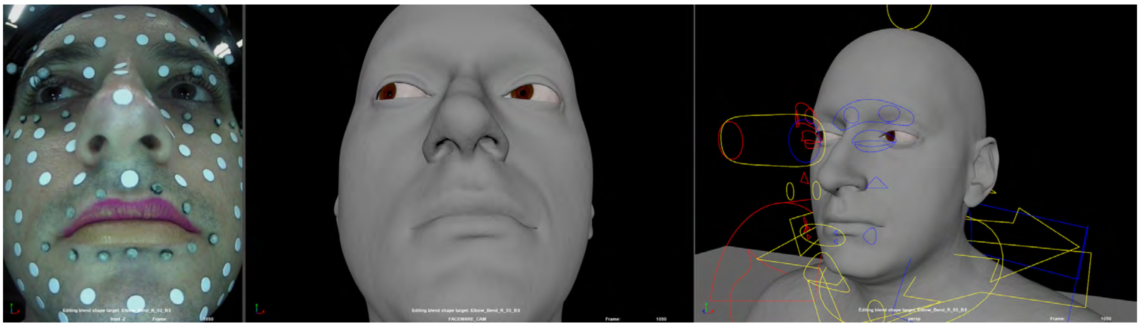


Figure 27. Side-by-side comparison of the real-world Faceware camera and my custom "Faceware Cam" view in Maya. An additional perspective view window is open to assist with facial shape adjustment from multiple angles.

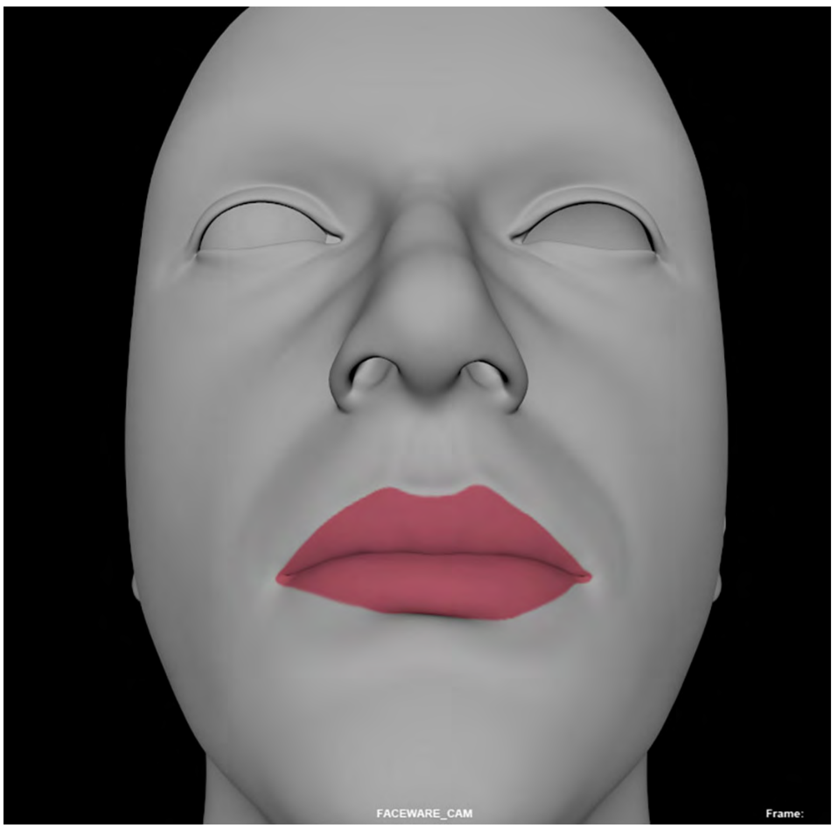


Figure 28. Lips colour map applied to model in order to better visualise the *vermilion border.*

also used a freely adjustable perspective view camera to orbit around the model to better see the deformations in three dimensions. Since I work with a dual-monitor setup, I set up the perspective view on the right monitor while on the left monitor I worked with both the source video and fixed-position camera view (Figure 27).

During my initial attempt at retargeting the performance onto the 3D character, the model displayed just a single grey colour material over the entire body. This uniform colour did not clearly establish visual landmarks associated with the skin. For instance, it was difficult to see the *vermilion border*, which in turn made it difficult to clearly define the lip shapes I wanted to create in any given mouth region. In order to provide a better colour separation between the lips and the rest of the face, I painted a colour map in Mari, a 3D texturing software by The Foundry (Figure 28). This colour map provided a basic red texture for the lips while retaining the default grey material for the rest of the model.

I started the retargeting process by identifying a neutral frame in the source video – that is, a frame in which every region of the face is as expressionless as possible. In such a frame, the lips, nose, cheeks, eyelids, and brows exhibit no tension or muscle pull. Ideally, if the 3D model is well-made, the neutral frame should match the base geometry. With the best neutral frame identified, I made small-scale adjustments to the facial controls and markers on the 3D character in order to closely match the neutral expression, including a few adjustments to the shape of the 3D character's face in order to establish more natural asymmetry consistent with the source video.

When creating each pose, I evaluate how much the lips stretch, compress, pout, suck inward, and contort by proportion to their positions in the previous pose. Likewise, I evaluate how other facets of the mouth region move with respect to the previous pose, such as nostrils, cheeks, the naso-labial fold, and the chin (see Appendix II.ii). The animation rig contains both large-scale and small-scale controls for the mouth. I first animate the large-scale controls to produce the majority of the movement, followed by the small-scale controls (face markers) to make specific, localised adjustments to the shape of the lips. It is important to clearly articulate how much (or rather, how little, in the case of the BOE) the movement of the lips in a given region change from one pose to the next. Frequently, a pose that occurs later in the animation may be similar to an earlier pose; due to the extremely subtle nature of the movement in this animation, if two poses that should be near-identical vary at all in their shape, this creates calculation issues when the performance is retargeted, resulting in data noise and performance errors. When dealing with subtle performance data like this, it is critical to maintain consistency across the entire performance. To reduce the chance for error, when I create a new pose, I first locate an existing pose that most closely matches the new pose. If no previous pose approximates it, I instead copy the shape of the previous pose (on the basis that the previous pose is likely to resemble at least some parts of the new pose). Whenever I create a new pose, I set a key on every value of each control associated with that particular facial group before I pose the new shape. Likewise, when I complete a pose, I key every value of each control associated with the facial group. This ensures that the pose is truly “locked-in” and that no control undergoes any unintended value interpolation, which would alter the shape of the pose.

When animating the lips, it is important to accurately represent the distance between the lips, or if they are closed, the degree to which the lips press together. Other important factors to consider include the angle of the *oral commissures*, whether the lower or upper lip presses out/in more, if either lip is curled in or rolled out, and whether the lips fully part or remain “sticky”. If there is any compression in the lips, this needs to be accurately represented by creating a sense of vertical squash with resultant horizontal stretch. Considerations such as these are equally applicable when animating other parts of the face. Any of these considerations may be prone to error due to the fact that my only close-up view of the face is from a front-on angle in the source video, but the facial shapes actually deform in three dimensions, not just two. As a result, it is especially easy to miss problems that arise along the depth axis. As I work, I keep a separate 3D perspective view open while matching my poses to the source video. This way I can orbit around my model to see how the poses hold up from every angle, rather than being limited to just a single point of view. This problem is specific to monocular facial tracking methods, such as Faceware and would have been less of a concern with the Vicon Cara system. While I use four static cameras to record reference of the full-body performance, those cameras do not have a high enough resolution to record meaningful facial performance to “fill in the gaps” left behind by the monocular PeCap system.

As I evaluated the initial retargeting tests, I took note of a lot of noisy data over a short series of frames in a specific time region. This was due to a particular action I performed during the shoot that involved shaking a bottle of juice. The motion from that action also transferred to my neck and head. The Faceware helmet was not fixed tightly enough to my head, which meant that the helmet moved independently from the motion of my head. This resulted in unwanted secondary motion, in which the Faceware camera moved in a direction different from my face. Although this was a minor amount of movement, such viewpoint variance created the impression in the source video that my lips physically moved up and down when they actually remained still. In situations like this, it was necessary to override the tracking/retargeting and resort to keyframe animation alone.

While creating the mouth poses, it became evident that the simple lip texture I created in Mari provided insufficient detail to serve as a meaningful guide to posing the finer detail of the mouth region. This was due to the fact that the rest of the face remained smooth with only a flat grey colour applied to it. A full-face texture better serves the task of animation by clearly identifying how both large-scale and small-scale details change from one pose to another. I used Mari to create a more detailed face texture by projecting photographic references of my head onto the 3D model (Figure 29). This step was more time-intensive to complete but ultimately resulted in a faster and more precise animation workflow due to improving my ability to register the deformation of the face (see Appendix II.iii).

The workflow I developed was to exclusively adjust the facial blend shape controls first (see Appendix II.iv). Once the blend shape controls were fully animated, I made minor adjustments with the skin-weight-based (SWB) controls. As previously mentioned, one of the distinct challenges in creating the BOE mouth poses was how subtly the lips varied in shape over time. In order to create accurate poses, the lips (and the surrounding parts of the face) needed to be checked against the source video to determine any slight differences in position, orientation, and tension (Figure 30) (see Appendix II.v).

Due to the additive nature of blend shape deformations, they sometimes cause unintended

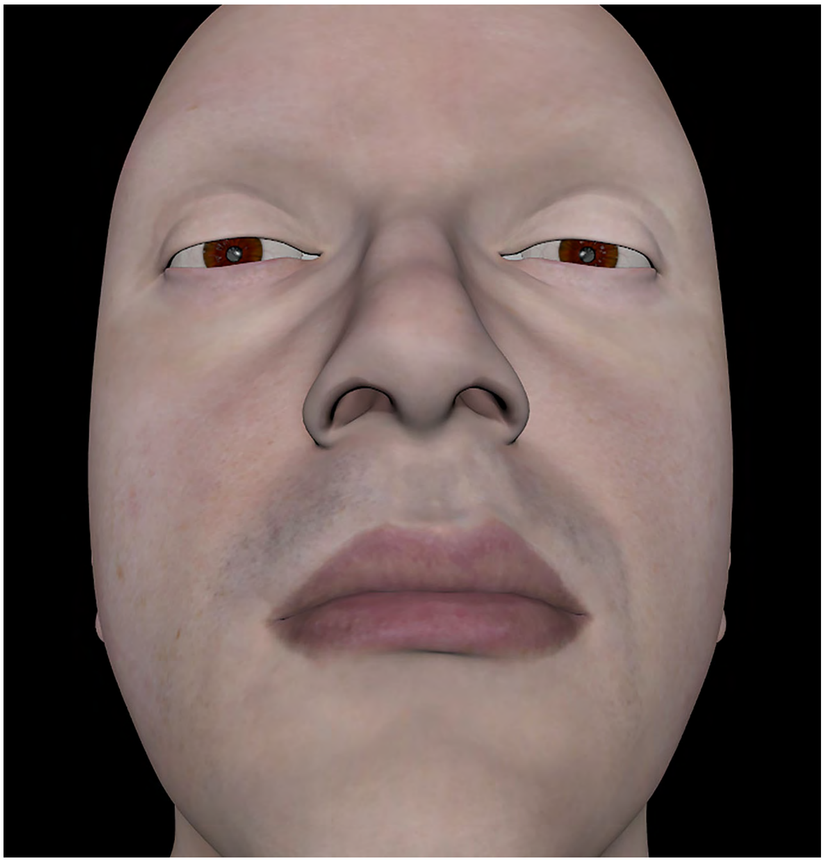


Figure 29. Full colour map with more prominent facial details applied to model.

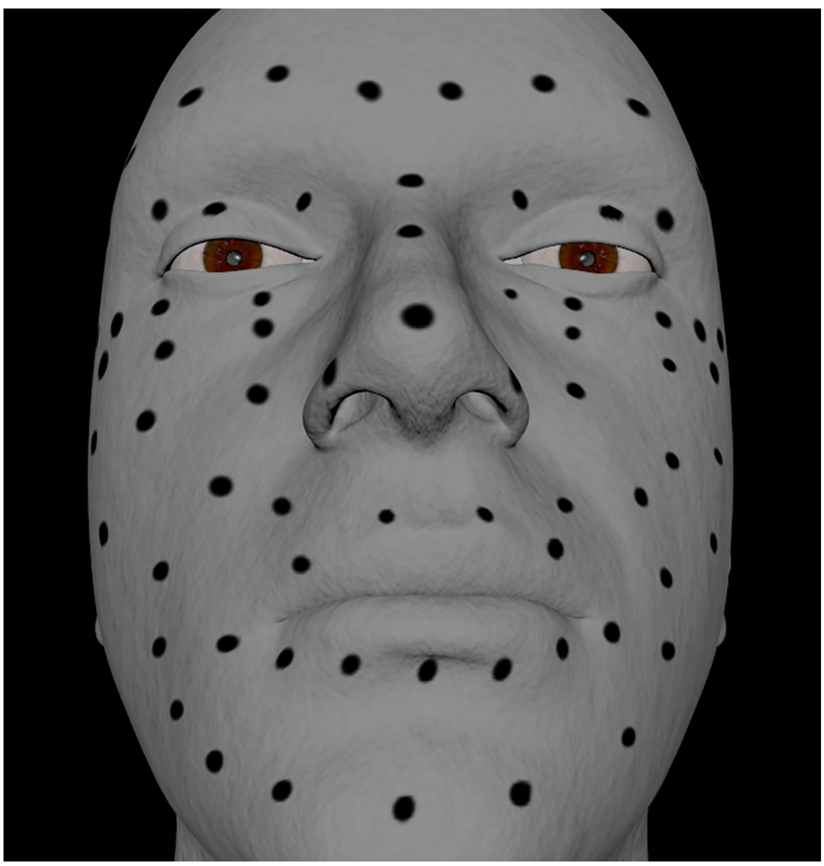


Figure 30. Fractal bump map and locations of tracking dots applied to model for greater visible detail.

physical changes to occur when an area of the face is deformed through the combination of multiple blend shapes. This is more likely to occur when the blend shapes are near their maximum influence, and can be a sign that a blend shape is being overused. The SWB controls are useful for making fine-tuned adjustments when trying to resolve such problematic deformations. The SWB controls are especially helpful when creating more accurate mouth and eyelid shapes for the Faceware poses. This is because the blend shapes can only deform within the limits of their defined shapes, which do not represent the full range of physical movement of the face, whereas the SWB controls build in minute localised shape adjustments on top of the blend shape deformations. Often this is accomplished simply by moving the SWB controls, but some shapes can only be achieved when rotating the controls. This is a potentially problematic solution in that it introduces twists within a localised region, but so long as a twist is not excessive and is attenuated or removed in subsequent poses, it usually does not produce noticeable issues. Such rotations were used extensively along the eyelids in the BOE, especially to prevent the upper and lower eyelid geometries from intersecting each other (see Appendix II.vi).

As indicated in the example of shaking a juice bottle, unwanted secondary motion from the PeCap helmet makes parts of the face appear to move at times. The helmet may slide along the scalp and become out of sync with the head movement, creating viewport variance. Ravikumar notes that this is a well-known issue in PeCap and can significantly affect the quality of the performance solve if the spatial offset is not accounted for (2017, p. 109). It is important that when evaluating any movement, the MoCap artist factors in these technical considerations, or else he/she may introduce false movement into the data. This is especially important to keep in mind when comparing poses between frames in order to maintain consistency. A cursory visual examination may suggest that a part of the face moves because it sits at a slightly different position between the two frames, but in reality, this is due to the movement of the helmet instead. Again, during the creation of the BOE performance, this was a particular concern because of the minor differences between the facial expressions and the need to produce consistency within that subtlety.

Even though I had applied a colour texture for the full face when creating the poses for the BOE performance, the degree of facial movement nuance was such that the texture still did not provide enough subtle detail to facilitate the kind of precise small-scale movements I needed to make with the SWB controls. In order to help with this, I replaced the face texture with a high-frequency fractal texture map applied through the face material's bump channel. This greatly helped with visualising every part of the skin that moved. I painted dots on the texture map to correspond with the tracking markers in the source video, so as to better visualise localised, nuanced movements (Figure 26). I also created lights on either side of the head that I positioned to clearly reveal the relief of the fractal noise. One light was red-tinted and the other blue-tinted so as to provide visual separation and detail. Using this method, I could swap back and forth between the colour texture and the fractal texture, depending on the needs of a pose. When creating poses for the eyebrow region, there was no eyebrow geometry to guide this process. In order to visually track the changes in the eyebrows, I added some basic eyebrow details to the colour texture.

The location of the eyelid relative to the iris plays a significant role in registering the expressiveness of the eyes, as a small difference can result in a more intense expression (too much of the sclera showing) or a sense of tiredness (too much of the iris occluded by the eyelid). I judged that the original irises were too small, so I remodelled the eyes to enlarge the irises and created a blend shape that allowed me to manually adjust the size of the pupil through animation. Human eyelids do not necessarily close the same amount at exactly the same time when blinking; while watching my performance, I noticed that my left eyelid was droopier[60] than the right and completed its blink 1-2 frames after the right eyelid. Therefore, it is important to not rely on simply copying the action of one eyelid across to the other because in a realistic context, they are not exact. This is a good example of why any kind of artistic intuition in the posing and animation of the face needs to be informed by a close analysis of the underlying source performance, rather than being based on assumptions about how the face functions.

---

[60] Related to a mild condition of ptosis.

While creating the Faceware poses, it became evident that a number of subtle facial shapes were not possible to achieve with the original set of blend shapes I developed. While aspects of these shapes could be achieved using the SWB controls, they often lacked important fine-details, such as wrinkles or skin bulges. As a result, I created a number of new blend shapes that were directly informed by the animation process itself, rather than being solely FACS-derived.

Once the initial retargeting of all three facial regions was complete, I compared the resulting performance to the source footage. I noticed that the upper eyelids were angular and flat in some regions where they should have been rounder. Eyeballs should generally lock into position when looking at something, rather than experiencing a constant series of micro-movements. However, noise in the retargeting data meant that the eyeballs continued to have a small degree of motion when they should have been still. By contrast, there was insufficient movement from the lower half of *orbicularis oculi* muscle whenever the lower eyelid closed or squinted. This movement is necessary to detect a squint as part of an expression, and without it, this region of the face lacks vitality. When the eyelids close, they do not just rotate but also have a small lateral translation toward the nose. This lateral translation was also missing after the first retargeting. These types of missing expressive components significantly contributed to a sense of uncanniness in a performance. I made notes of several inconsistencies between the retargeted animation and the source video to guide me through adjusting the poses before creating a second retargeting. Most of the poses that were problematic after the first retargeting were on in-betweens that Retargeter automatically interpolated rather than on keys that I manually set. In any problematic region, I looked for the one frame that I felt most exemplified the problem and made a note of what specifically needed to be adjusted on that frame. My rationale for doing this was that if I adjusted the most problematic frame, the other surrounding frames should proportionately inherit similar adjustments.

The initial retargeting also revealed some modelling issues around the eyes, especially in terms of interpenetration between the inner eyelids and the corneas. I also further smoothed out some of the geometry along the edges of the eyelids, particularly when they closed together. It is important to note how the animation process can reveal problematic issues with the modelling and blend shapes that are not apparent in an inanimate model alone. After I applied the modelling changes and fixed the issues identified in the initial retargeting, I completed a second retargeting pass, after which I cleaned up any extraneous motion and jitter. A key element was in determining when movement was due to jitter versus the nuances of subtle face motion. It also became apparent that while the motions in the face matched the source video, the beginning and endings of the motions frequently were either a few frames early or late. This required manually adjusting frames using Maya's Graph Editor[61] to ensure the timings of the motions were as precise as possible. Likewise, sometimes when the eyes or lips closed, they did not fully close to the same degree as in the source video and had to be further clamped together. Because Faceware doesn't provide a true 3D tracking solution, it relies on the quality and accuracy of the poses to match meaningful performance details in the various face regions. Therefore, its output is only as good as the quality of its input – but even with excellent quality input, this will always be more limited than a true 3D tracking solution that records the position of each marker on every frame.

Midway through the facial clean-up, I decided to attach the facial animation to the body animation. Until this point, I had animated the facial performance in isolation because the body movement was not yet relevant to that part of the process. However, it is advised when evaluating the facial performance that the body movement is factored in because the two are inextricably linked. I also recognised that there was much facial performance that did not need to be highly resolved if it wouldn't be seen in a shot. To this end, I created a digital version of my kitchen scene using freeware 3D assets. I placed the full body character into the scene and choreographed the beginning and ending of the character's motion so that it logically navigated the layout of the kitchen scene. At times, this required some creative adjustments to the location of the character, but any of these changes were masked

[61] Maya's Graph Editor provides a graphic representation of animation data over time, which is often a more direct method to edit values and the spacing of actions than through manual adjustments in the 3D viewport.

via the cinematography and editing process. I used Maya's Camera Sequencer to interactively edit the cinematography of my scene, producing shots via virtual cinematography that could continually be edited for length and composition. Several shots emerged where the character's face was hidden from camera. During these shots, I edited the data of only the parts of the face that could be seen, if any. Likewise, if a part of the body would not be seen, such as the legs and feet in a close-up of the head, then I avoided cleaning up the motion on those parts so long as they did not produce jitter or other errors that affected the parts of the body visible within the shot. This saved a great deal of time and meant that I was not working on elements that would never be seen.[62] This method proves effective at establishing the cinematography following a PeCap recording but prior to motion clean-up, and ensures that PeCap artists and animators efficiently spend their time on only relevant details. This is relevant for fully virtual (ex-filmic) captures, whereas profilmic captures may be determined by the cinematography of the live-action footage.

6.2.3.3 BOE Body Clean-Up Process

The body tracking provides its own challenges, particularly when the character needs to interact with digital counterparts to real-world objects. In the source video, I grab both a bowl and a mug from a makeshift ledge that is meant to represent a cabinet. However, the ledge is not high enough to believably function as a cabinet, and in fact serves dual purpose as both the purported height of the cabinet and the countertop. Therefore, when I reach for the bowl and the mug, I pick them up and place them back down at the same height (Figure 31). To overcome this discrepancy and in order to minimise the difference in height, in the digital version of this scene I place the bowl and mug on a dishrack[63] (Figure 32). It could be argued that in this moment, I have added a degree of pretence that changes the performance from simple to complex acting. However, I quickly return to a simple acting performance once the action is complete.

---

[62] Of course, if I were to alter the cinematography later and reveal the face in those sections, I would need to edit their motion.

[63] The fact that the dishrack is far away from the sink is an idiosyncrasy that I hope the audience will overlook.

Figure 31. Screenshot of the setup for the BOE performance as I adjust a bowl and mug on a makeshift countertop that also serves as an *ad-hoc* cabinet.

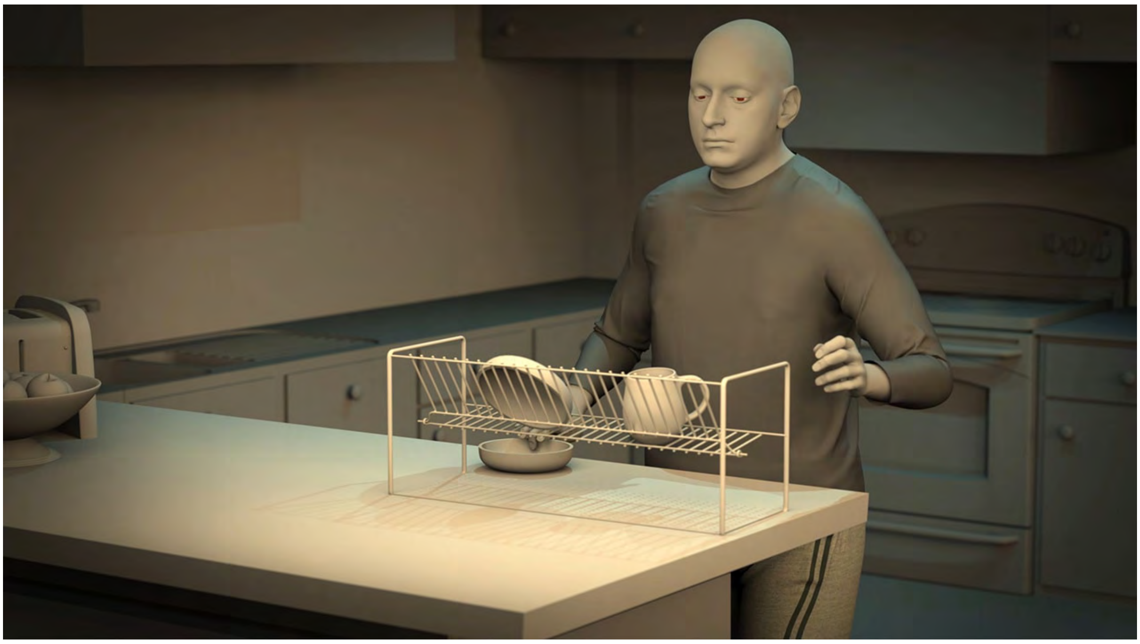


Figure 32. Screenshot of the digital counterpart performing the same action as in Figure 31.

The orientation of a bowl drying on a dish rack is different from the orientation of a bowl sitting in a cupboard. Therefore, not only did I need to adjust the height of my arm, but I also needed to adjust the orientation of my hand and fingers as it reached for the bowl, flipped it around, and placed it on the countertop.[64] Furthermore, in the original performance, I moved the bowl only fractionally closer to me, whereas the addition of the virtual dishrack required that the bowl move substantially closer to my digital counterpart. This required quite a bit of keyframe animation to resolve because the MotionAnalysis (full body MoCap) system doesn't provide accurate finger tracking. As a result, the fingers always require keyframe animation, but the adjustments to the performance of the hand and arm required custom animation on top of the existing MoCap data. Because this bowl action was not part of the original performance, I had to film separate reference of myself removing a bowl from a dishrack and flipping it around to rest upright on a countertop. I filmed this action from several angles to ensure that I captured all of the relevant performance details. Fortunately, the mug's position and orientation in the original performance was compatible with a mug on a dishrack; therefore, I was able to preserve the original performance of my hand picking up and placing down the mug while adjusting for the difference in height between the 3D dishrack and countertop.

There is a significant challenge when keyframe animating the motion of a realistic performance, especially when it needs to seamlessly match the movement of the rest of the body's PeCap data. While stylised 3D animated characters may possess believable movement through a more liberal application of animation principles, the keyframed movement for a realistic animation must be as similar as possible to the underlying performance data elsewhere in the body. For precise movements, especially when animating interactions with props, this often requires keys to be set on every frame over a given sequence (sometimes referred to as "brute force" animation) in order to take full control of the animation in a believable fashion. Brute force animation is also necessary in order to resolve problematic rotations when blending between the MoCap animation and the keyframe animation. This

[64] My original gaze remained a good match for this modified action and did not need to be adjusted.

was a particular problem in regard to the wrist when placing the bowl onto the countertop: Maya's automatic interpolation caused the wrist to constantly twist in an undesired orientation, which required that I carefully reorient and reposition the wrist on each frame over a 16-frame sequence. I performed these actions by transitioning the character's arms from being evaluated with forward kinematics (FK) to inverse kinematics (IK).[65] It is generally preferable for an FK-IK transition to occur over a shorter interval, such as 4-6 frames at most, but in some cases an extended interval is necessary to hide the differences between the keyframe-animated versus performance-captured versions. This was the case for the left hand reaching for the mug, which required 25 frames to produce a convincing blend between the keyframe animation and the MoCap data.

Furthermore, the geometry of the dish rack and the countertop presented some limitations for the motion-captured movement of the arms: with the character positioned in front of the dishrack, the right hand would at times intersect the dishrack while the left hand intersected both the dishrack and the edge of the countertop. I used Maya's animation layers system to subtly adjust the positions of the hands to avoid these collisions while preserving the basic trajectory of the Mocap data in the arms and hands. To minimise the amount of layered animation needed to solve this problem (in order to preserve the original performance as closely as possible), I adjusted the dishrack model by scaling its geometry to be significantly wider – thus reducing the degree to which the hands intersected its edges. Encountering this problem is good in that it is symptomatic of problems sometimes faced between the actions capture on a MoCap stage versus the digital environments into which those actions must be placed. In an ideal scenario, the 3D assets would match the dimensions of the real-world set pieces, but this is not always achievable and workarounds such as this must be enacted.

[65] FK calculates the position and orientation of each sequential joint of an appendage (such as an arm from the root of the appendage (the shoulder) to the position of the end effector (the wrist). IK works in an opposite fashion by first evaluating the position of the end effector and using this as a basis to calculate the joint angles between it and the root of the appendage. FK is more intuitive for animating most arm movements and is the type of animation directly produced through MoCap, but IK is useful when animating against resistance (such as a hand or foot pushing against a surface) or for exacting control when overriding existing MoCap data.

While the overall performance is preserved, it is important to note that the moment when the character picks up the bowl and the mug has been significantly changed in the animated version due to the differences in height and orientation of the objects within the scene. As a result, the hands, arms, and fingers are, to varying degrees, the animator's performance rather than the source-actor's performance, although the source-actor remains principally responsible for all other actions throughout the character's body during this time. The animator's job is to complete any actions that were not part of the captured performance while attempting to respect the pace and intent of the original performance. This includes providing "invisible seams" between the moments when the custom-animated performance joins the original performance. This can be challenging, especially if the custom animation significantly varies from the original performance and there are only a few frames available to create a smooth transition between the two performances. If there is a large spatial difference between the two performances and the transition from the custom performance to the original performance is too quick, it may result in motion that "snaps" or "pops" in an undesirable fashion, leading to an uncanny moment. One such moment that I could not overcome in this animation was at an FK-IK transition on the right arm after the character lifts his hand away from the bowl. Due to the differences in how FK and IK rotations are calculated, this resulted in a subtle but noticeable pop in the rotation of the forearm muscle near the crook of the arm. I tried several methods to fix this error, including the possibility of moving the mug in front of the problematic transition at the moment the error occurs in order to visually hide the pop. I planned on dressing the character anyway, so I elected for some loose clothing that could hide the underlying problem while also supplying some digital modesty.

When cleaning up the MoCap animation curves, there is a fine balance between removing noise from the data and preserving the natural variances within movement that makes it identifiably human, and preserves the index and presence of the source actor. In standard 3D animation, the animation curves tend to possess more graceful, consistently curved arcs throughout the data, which produces a sense of fluidity to a character's movement. However, natural human motion, while heavily influenced by arcs, is not so perfect and

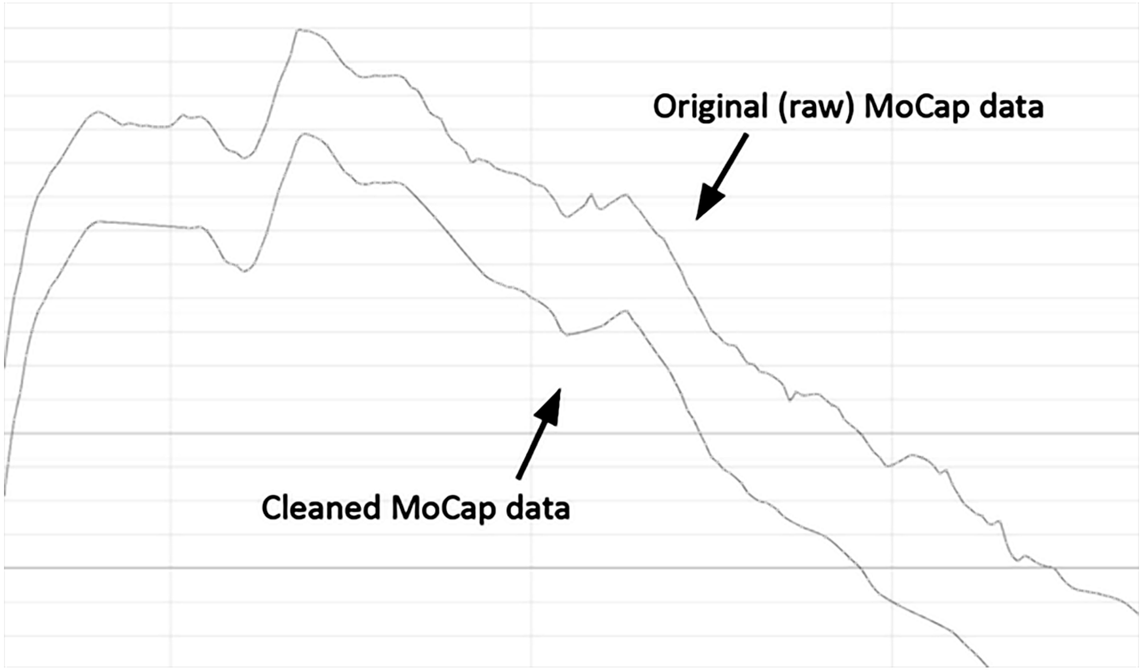


Figure 33. A comparison between the shape of the raw (uncleaned) MoCap data and the cleaned MoCap data for a single control over a 200-frame interval. Note that the positions of the curves in this image are for the sake of visual clarity and do not represent changes in values.

can seem overly animated if the MoCap data is cleaned from the perspective of smoothing all curves in the same way as in standard 3D animation. Essentially, in MoCap, any data that contributes a sense of jitter to the motion needs to be removed, but the data clean-up should stop there. Being able to make this type of judgment is a skill that is learned through hundreds of hours of experience with cleaning MoCap data and assessing how it impacts a digital character's performance. Ultimately, the decision should come down to the effect it has on whether a part of the body visually jitters, rather than how smooth the underlying animation curves are. This consideration is primarily for keys that show a minor degree of variance by comparison to the overall trajectory of the animation curve. For keys that greatly deviate from the curve's overall trajectory, their variance is most likely due to noise and they need to be removed (Figure 33). However, it is still important to confirm how this visually affects the character's animation.

In the original performance, after the right hand lets go of the bowl, the right hand's fingers remain in static positions while the arm swings next to the body and the body turns clockwise. This looks fine in the source video but I do not believe this action feels right when watching it as an animation. I attribute this to the phenomenon in 3D animation that if a part of the body becomes completely static, it is perceived as "dead". There is something particular to the graphic nature of 3D animation, even highly realistic animation, that demands there should always be even a slight amount of movement to keep a given part of the character (i.e: the fingers) "alive". Therefore, I added a minuscule amount of finger flex throughout this section. The result has little effect on the overall performance but is critical to overcoming any sense of uncanniness that would otherwise result. This ties in with the different modes with which an actor perceives his or her performing body via motion capture discussed in Section 5.2. From the perception of my body as Doppelgänger, the virtual construct of my body performs exteroceptively. I experience an uncanny reaction to the movements of my Doppelgänger when its movements do not match my expectation of how my own body moves, especially when the movements appear too stiff. In such instances, I am taken out of the moment with the character and my empathetic connection is disrupted. I add in minuscule movements to support my digital counterpart's underlying actions until I perceive my empthatic connection with the Doppelgänger restored.

While I attempted to accurately translate and preserve all of the details of the source performance onto the character's face, an unexpected issue arose due to the angle between the character's head and the virtual camera within the scene. When the character turns clockwise to screen left and looks in that direction, the original eyelines from the source footage make the eyes appear to look too far to the character's right – effectively, the irises disappear from view and only the whites of the eyes remain visible. This feels visually problematic and may suggest to the viewer that the character's sight line could be better composed. In this case, it seems justified to alter the sight lines on the basis of aesthetics; rigorously holding to the source performance doesn't produce a better result, and may in fact negatively affect the audience's perception of the character in that moment. This problem is specific to fixed-perspective cinematography; in the case of a video game, virtual reality, or an interactive film, the audience may be able to move around the character, thereby overcoming the limitations of a fixed composition.

6.3 The MoCap Artist and Analysing Data

The BOE performance provides some foundational insights into the nature of the roles of a MoCap artist and animator. The process is heavily technical and needs to be guided by an eye for close detail and a clear understanding of how performance is constructed. It is important to note that I elected to produce the BOE as an animation only for as long as it provided novel insights to me about the process. This is consistent with my practice-led methodology, specifically in relationship to participating in research that "leads primarily to new understandings about practice" (Candy, 2006, p. 1). The goal of the research is not to complete the animation but rather to articulate the process of creating the animation. After a certain period of time, there are diminishing returns in terms of new knowledge learned from the process, as much of the effort is either repetitive or based on previously discovered insights.

While the BOE performance is 117 seconds in total, the amount of finished animation encompasses just eight seconds. A great deal of work was required to complete even this short interval, which also illustrates the sheer scale of the work and the necessity for so many 3D artists to be involved in the process of creating PeCap performances. Consistent with my methodology, I serve as the sole performer and 3D artist in this research, and there is only so much work that can be meaningfully expected from a single contributor to what is normally a collaborative process.

In the appendices, I discuss in detail my approach to modelling (Appendix IV) and rigging the 3D character (Appendix V). As with any technical process that requires a long time to implement, my experience lent many ideas for how I would modify the process given a second opportunity. During the process, I came across techniques for creating a more efficient rig, which would have allowed for real-time animation playback. Unfortunately, one of the most limiting factors in the rig I created was that it was computationally demanding and I could not interact with it in real-time. I believe that real-time interaction facilitates an animator's intuitive understanding of how to modify a performance. When working with non-real-time feedback, the animator must instead rely on software to produce a *playblast*[66] of the animation. These calculations occur only at the scene's maximum available framerate and can sometimes take minutes to hours to complete, depending on the duration of the sequence and the complexity of the calculations. However, they are the fastest means by which to preview an animation when real-time playback is not available. The playblast creates a delay between an animator's modification to a character and the result of that modification, which inhibits intuitive feedback. While animated acting is never created in real-time, I hypothesise any process that facilitates a quicker response between the actions of the animator and the performance of the character produces a more natural sense of directly performing through the character. That is, in real-time feedback, there is a stronger link between action and affect for the animator-as-actor. When working with a delayed response between action and affect, instead of feeling like an actor who directly controls a virtual puppet, I feel like a mere adjuster of discrete controls whose overall effect will not be immediately apparent. I liken this to the experience of playing the piano – if you have to wait ten seconds to hear the result of the keys you just pressed, the loss of immediate feedback inhibits the discovery of the performance through meaningful experimentation and improvisation.

As mentioned in the previous section, the Vicon Cara PeCap system was originally intended to complement the Faceware ProHD monocular PeCap system used in this research. However, due to technical issues and a lack of experience of working with the Cara, the data we produced with it was unusable. This is unfortunate as the Cara data could have been used for comparative analysis in terms of the fidelity of performance capture between the two systems, and the Cara would have provided much more performance detail. Also, the Faceware system proved to be less reliable when tracking long, continuous sequences of 30 seconds or more. This is because the Faceware system is more often used in the context of recording PeCap for video games, where it is more common to track shorter performance

[66] A playblast refers to a series of viewport screen captures of each frame of the sequence, which Maya calculates and assembles into a separate video file that plays back in real-time.

sequences. The type of pixel tracking algorithms that Faceware utilises seem to be more accurate over shorter intervals, and they appear to become less reliable over longer sequences. I tracked the entire two-minute BOE performance, which required a great deal manual keyframe creation in order to produce an accurate tracking result within Analyzer. It is conceivable that if I broke up the performance into smaller intervals and tracked each interval separately, an accurate overall result may have been faster to achieve. If the virtual cinematography of a scene can be determined shortly after following the PeCap recording, then the cuts in the virtual cinematography could be used as the basis for the tracking intervals. It would have been meaningful to test how the Cara's tracking would have behaved under similar circumstances, as a true 3D-tracking system should theoretically remain more accurate over longer intervals.

6.3.1 Reflection on the BOE Performance

When I've solicited feedback for the BOE performance, I am consistently told there is a strong sense of naturalism to the body movement. This should be expected, given that the animation is based on a PeCap recording, but it is encouraging that viewers confirm the animated movement is believably human. I notice many traces of my own body in the finished performance – specifically, a pronounced stiffness in the shoulders and upper back related to my chronic pain condition. The shoulders are also hunched forward, producing my walking posture. Such idiosyncrasies preserve and convey an actor's aura in the finished animation. One person's feedback was that the actions "felt like a typical sleepy morning". There is a sense that while the character's actions are not quite aimless, they are also not urgent. This is appropriate to the overall feeling of an exercise meant to communicate a natural moment that doesn't involve crisis, and is consistent with my normal approach to preparing breakfast.

While there is a general sense of stiffness in the upper back region, there is a sense of looseness in the arm movements, facilitated through a natural degree of overlapping and secondary action, especially in the forearms. This is particularly noticeable during the first few steps before the character stops in front of the countertop. This is a nuanced form of natural movement that can often be missing from keyframe animation if it is difficult for the animator to clearly see all of the subtle details in a reference performance. These small-scale details help to add a greater sense of weight to the character as they suggest a series of natural and unconscious forces throughout the body.

In the original PeCap data, the character's toes frequently intersect the floor. This is due to a lack of capture detail in the foot region. In the cleaned performance, I animated the toes to remain on the floor and help push the foot off the ground when the character takes a step. I also added in a subtle degree of overlapping action to the toes by delaying the toes contacting the ground by a few frames after the foot comes into full contact with the ground on each step. This is common technique to break up the animation on different parts of the body in order to create a "looser" sense of animation. It is a barely noticeable detail, but it helped make the foot feel like it is composed of several anatomical structures rather than being just a single bone. After the scene switches to the second shot, the feet are hidden and I ceased cleaning their animation because such work would have no visible effect. However, I animated the knee controls (pole vectors) throughout the entire animation because misalignments in the knees in the original PeCap data produced numerous deformation errors around the hips.

The facial performance in the first shot is very minimal, in part because there was also minimal expression in the source performance. Also, the virtual camera is far enough away from the character's face so as to render most nuanced movements imperceptible. By contrast, the second shot features the character in a mid-shot with his gaze directed at a bowl on the dishrack. Due to the MotionAnalysis system tracking only the hand motion and not the individual fingers, the fingers in the original PeCap data are rigid; all finger movements throughout the finished animation are completely keyframe animated. In the finished first shot, I pose the fingers to match the slightly cupped shape of both hands, which appear relaxed. This also provides visual contrast to the following action when the character's right hand opens wider a few frames before his fingers reach the bowl. His fingers curl around the bowl before the he lifts it off the rack and his eyes follow the bowl. Here the fingers may feel somewhat rigid as a result of fitting the character's finger geometry around the lid of the bowl. The character's hand geometry is not a perfect match for my own hands, nor is the bowl a perfect match for the actual shape of the real bowl used in the source performance. As a result, the hand-bowl interaction is at best an approximation, but hopefully one that transpires quickly enough to not draw much attention. The middle finger holding the bowl also has a conspicuous crease along one of the knuckles due to being bent further than normal.

After the bowl is placed on the countertop, the character's right arm remains manually controlled through keyframe animation until just after the aforementioned pop in the forearm, at which point full control transitions back to the PeCap data. Throughout the manual control, there is a sense of the arm being slightly more mechanical in its motion when compared to the fluidity elsewhere in the body. The pop in the right forearm is only noticeable if one specifically looks for it. This could be further refined, but since the other arm becomes the focus of the animation after the bowl is placed down, I resolved the motion to the point where any defects would not be noticeable upon a single viewing (as opposed to the scrutiny of a frame-by-frame analysis conducted through repeated viewings). However, this serves as a meaningful illustration of the importance and challenge of matching keyframe animation to MoCap data. Likewise, any natural stiffness in the original performance could be interpreted as an animation error for the 3D character, such as the stiffness of the character's right forearm as he begins to turn clockwise. It is tempting to add a small amount of movement to this motion in order to overcome the "deadening" effect that stiff poses produce in 3D animation. As a result, the nature of 3D animation is such that some subtle performance modifications need to be made to attend to the idiosyncrasies of the medium.

The keyframe animation of the left hand is perhaps more successful than the right hand in terms of achieving consistency with the MoCap movement elsewhere in the body. The motion of the arm as the hand reaches for the mug is both quick and purposeful, and the fingers convincingly wrap around the handle of the mug. They also release in a convincing manner once the mug is placed on the countertop, and the hand naturally relaxes as the body turns. The character's eyes watch this action as the hand reaches for the mug, but the character changes his focus to the bowl while he lifts the mug off the dishrack. This suggests that the character is aware of the movement of the mug via his peripheral vision and proprioception, but he doesn't need to give the action his direct focus in order to complete it. As previously mentioned, the direction of the eyeballs was altered at the end of this sequence to keep the irises visible.

Less prominent elements in the animated version include subtle adjustments to the lower *orbicularis oculi* muscle, especially when the character looks to his right before turning clockwise. Without this, the upper cheek region feels rigid, and the bags underneath the eyes do not feel fleshy by comparison to the cheek muscles just below them. Overall, the character's eyes saccade accurately and provide sufficient random micro-motions to generate a naturally complex sense of living movement. However, the challenge is in differentiating meaningful micro-motions from noise in the eyeball data, even when manually keyframing their movements (Fordham, 2017a, p. 19). There are a few interpenetration issues when the eyelids close, as well as a sudden downward shift in the inner left brow's position. The BOE animation was originally completed with an older (e.g. less accurate) version of my face model; when I later added a more accurate 3D scan of my head as a full-face blend shape replacement, this resulted in some inconsistencies in terms of existing blend shapes and produced unintended geometry intersections. Most of these issues are only apparent upon an extreme close-up perspective of the face viewed through the virtual camera that mimics the source video camera (the Faceware Cam view). Therefore, there is a trade-off between how close the camera is to the character and how much detail needs to be cleaned up to exactly match the source video. At a certain distance, some fine-scale adjustments are simply not detectable. Also, the Faceware Cam view doesn't account for the angle at which the final composition "sees" the character, and therefore some of these issues are purposefully ignored because they are turned away from the camera.

Beyond the eyes, the rest of the face has little visible movement. This is not surprising as the BOE is likely to generate performances of limited intensity. Such performance subtlety reads well on a human actor but is somewhat problematic on a virtual character due to the deadening effect (previously mentioned and discussed further below). Returning to Sito's observation that performance-captured characters produce more believable performances when they engage in action rather than when they perform from a fixed position (2013, p. 210), the BOE performance illustrates how a believable performance is also more straightforward to animate when the character is dramatic rather than reserved. When examining the BOE performance, there are no clear expressions to evaluate whether the performance is "on" or "off" in regards to communicating a given emotion.

The concept of *deadening* is commonly cited in discourses surrounding PeCap production. Following the release of "The Curious Case of Benjamin Button", director David Fincher is famously quoted saying that the image analysis PeCap software used on the film "tend[ed] to sandblast the edges off a performance", which required a great deal of work from "artists" (e.g. *animators*) to restore the nuance of the original performance and overcome the deadening effect[67] (Duncan, 2009, pp. 83-84). Fincher describes the image analysis software as throwing out details that the "computer didn't think were important" (Variety, 2008, para. 17). While this is partly true, it is only part of the story: the software is less likely to discard details on the basis of importance and more likely to evaluate a detail incorrectly[68] where the degree or intensity of expression is compromised, or the timing is less than perfect. Fleming suggests that despite the rhetoric of flesh-and-blood actors providing a performance that the animators must honour, perhaps animators are the ones truly responsible for the finished digital performance, with source-actors "merely providing a screed of digital information for [animators] to work upon." (2012, p. 208). He bases this

---

[67] In this case, Fincher usually refers to the deadening effect in regard to performances that feel overly smooth, dull, or desensitised. This is certainly one connotation of deadening, and the phenomenon where a 3D character that lacks motion feels dead is another.
[68] It may be inaccurate to suggest that the software is "aware" of such details at all. It is more likely that mathematical filter operators evaluate and modify the data solely on the basis of a "signal space", rather than some kind of "performance detail space".

supposition on Fincher's description of the PeCap process as reliant on the manual override of animators to correct errors introduced through the image analysis process. In "Benjamin Button", the PeCap recordings are discussed as "raw information upon which animators then work." (2012, p. 208). Animation supervisor Steve Preeg discusses this process in detail:

> Image analysis figured out the timings of when the brows went up or when the mouth moved, and all of those correct timings would be in the performance of the head. But that was just the starting point. We still had to determine if the face was doing what it was supposed to in terms of expression and emotion. Quite often, coming out of the image analysis, it would look like it was moving more or less the same as Brad Pitt in the performance capture, but the emotion would be a bit off. He might look angry or sad, when that wasn't the emotion we were going for at all. Some shots were great right out of the image analysis, but most of them required some hand-manipulation. (Duncan, 2009, p.83)

Of course, Preeg refers to much older software and numerous improvements have been made since 2008. However, there still remains a sense that animators are asked to decipher the equivalent of clay tablets or ancient papyri with 40% of the letters either missing or illegible. In a separate interview, Preeg notes how even 1mm shifts in the position of an eyelid cause dramatic changes in expression, and how the image analysis software was never precise enough to replicate the source performance without manual intervention from the animators (Flueckiger, 2011, p. 11).

There is no clear quantification of how much of the final motion in a performance-captured character is directly derived from the PeCap data versus how much is augmented or created from scratch by animators. Speaking about "Avatar", director James Cameron is quick to estimate that at most ten percent of the finished performance is keyframe animated, but in this estimate he accounts only for features on the Na'vi alien characters that differ from their human counterparts, such as the ears and tails. As Flueckiger points out, this estimate does not refer to the inner mouth region, such as the tongue and inner lips, which is a critical performance zone that is also complex and difficult to convincingly animate. Likewise, any time human action is exceeded, such as with many of the Na'vi characters' leaps and dives, keyframe animation is required (2011, p. 23). It is likely safe to assume that many other manual adjustments were made by animators to match the human source performance to the Na'vi physiology, especially in terms of the massive differences in the eyes and noses between them. Cameron's estimate is at best a low-ball figure that serves his purpose in promoting performance capture as a tool that completely translates a source actor's performance onto a digital character.

In the context of the BOE performance, there is a clearer sense of how much the original data drives the finished performance versus how much is augmented by manual keyframing. When comparing the raw retargeting data (before clean-up) to the finished retargeting after clean-up, there is minimal accurate eye tracking in the raw data despite great pains taken to carefully track the eyes in Analyzer. All eyeball movements were manually keyframed in the cleaned version. In the raw data, there is only one moment where the mouth region moves at all – a minor widening of the lips. There is likewise little movement in the mouth region of the cleaned version, but the movement that exists is more purposeful: there is a sense of the lips pressing together rather than just widening, which feels like a more natural movement rather than just an arbitrary blend shape interpolation. However, there is still some jitter present around the lip corners toward the start of the sequence, which could be removed. This is jitter that is surprisingly not present in the raw data, which suggests it was accidentally introduced through the clean-up process. This is a distinct challenge when matching such nuanced facial movements: it is easy to accidentally exaggerate a movement by creating a fractional adjustment on an animation control. Another animation pass would help to tone down any such mistakes, but fortunately these errors occur during a section of the scene where such minor movements are not visible to the camera. Much of the clean-up process involves not just removing jitter but also repositioning many distinct keyframes to better match the shape of the source video. There are no brow movements in either the raw or clean versions. This makes me wonder whether it would be wise to add some micro-movements to the brows when the eyes look down, for instance, so as to overcome any deadening effect in that region of the face. The source video doesn't clearly reflect such a movement, but with movements as minimal as this, they may be difficult to detect. Also, there is a lot of noticeable camera jitter based on the location of the Faceware camera on the Cara bracer arm, and it is possible that any small-scale brow movements are confused by the large scene movements of the camera. However, even if I added in a minor brow movement, the likelihood, as with the mouth, is that they would only be visible from a close-up camera. In the end, I estimate that about 40% of the BOE facial performance is constructed via manual keyframing, which accounts for all of the eyeball movements and many subtle shape adjustments to the underlying expressions within the raw data.

Providing an estimate for the body performance is complicated by the fact that the animation required a great deal of adjustment to overcome the placement of the bowl and the mug. Excepting this adjustment, the major manual interventions include ensuring the toes do not intersect the floor and fully animating the fingers, along with some basic jitter reduction throughout the body. Including the adjustments for the bowl and mug, I estimate that about 30% of the BOE body performance is constructed via manual keyframing. This value reduces to 15% if I exclude the adjustments for the bowl and mug. There is no clear method for quantifying how much the face versus the body contribute to a performance, so these values are, at best, understood within their individual contexts. However, if I had to provide an estimate as to the amount that manual keyframe intervention played in the overall BOE performance (including the adjustments for the bowl and mug), I suggest it is between 30-40%. That is, due to the nature of the performance equipment and attendant software, only 60-70% of the source-actor's performance is directly reflected in the raw data.[69] It is the job of the MoCap artists and animators to fix the data to restore the original performance as closely as possible and, in certain cases, create an alternative performance to match the intention of the scene. It is worth noting that this estimation is based solely on the product of one particular experiment and should not necessarily be seen to suggest a value of what should be expected from either the Faceware or MotionAnalysis systems. Instead, this is the result achieved based on the given circumstances and participants in this research.

---

[69] This value is consistent with industry estimations of the amount of raw MoCap data versus keyframe animation in the final performances of digital characters (Sito, 2013, p. 212).

6.3.2 *The Role of the MoCap Artist*

The term "MoCap artist" is more of a catch-all designation than a specific role. The responsibilities of a MoCap artist may vary for each production and depend upon the size and skills of the creative team. It is not always clear what, if anything, qualifies a MoCap artist from a MoCap technician, since the two roles often overlap with each other during production and postproduction. MoCap technicians are primarily responsible for the maintenance, preparation, calibration, and recording of the MoCap equipment, props, and sets. The MoCap technician needs to understand and communicate how a MoCap actor will interact with other actors and props, and how to provide accurate motion tracking of these elements. The MoCap technician must make accurate measurements of the performance volume and understand not only the layout of the scene but also the topography of the floor – for instance, if it is uneven or sloped – as this affects how the MoCap data will be interpreted during the clean-up and animation phases. As Menache points out, these are details vital to the MoCap process that extend well beyond the information typically represented in storyboards alone (2011, pp. 84-85). The MoCap technician also serves as an authority about MoCap best practices for producers and directors who may not be familiar with the process or the specific equipment used in a particular studio, including how to match a performer to the role in question – such as football players for a football video game, or martial artists for a fighting game, and stunt performers for more complicated physical stunts. Likewise, the MoCap technician should advise that the performer match the intended proportions of the virtual character as closely as possible (2011, p. 110). MoCap supervisor Remington Scott recommends that video cameras record each take from the periphery of the capture volume, providing long-shots, mid-shots, and close-ups of the actors' performances (if sufficient cameras are available) (Serkis, 2003, p. 86). These are obvious concerns from a MoCap artist's perspective but which a person unfamiliar with the process may not consider.

MoCap artists also have roles downstream from production, including analysing and tracking data (as I demonstrate in this research). MoCap artists are also responsible for cleaning MoCap data, which involves a keen sensitivity to knowing when minor variations in movement are intentional rather than the result of noise. MoCap artists also use software such as Cortex, which is used to clean the optical MotionAnalysis data in this research, to reconstruct marker trajectories that are either corrupted or missing in the original data. On larger budget productions, MoCap artists may also acquire the preparation and calibration data captured by the technicians (basic facial shapes and expressions, phonemes, and sample dialogue) in order to train (often proprietary) facial solvers. As Failes points out, "[t]his is the key solution [Weta Digital] has R&D'd over several projects […] to take human facial performance and re-target it to a cg [sic] character." (2017, paras. 4-5). A MoCap artist must possess a fine eye for detail when tracking video-based performances, and understand how a minor variation in an actor's expression can have a significant impact on how an audience interprets the performance.

6.4 The MoCap Animator and the Mimetic Imperative

The common refrain that the role of the MoCap animator is to "honour" the source-actor's performance is typically attributed to Andy Serkis, who, as I discussed in Chapter 2, characterises PeCap as a type of "digital makeup"[70] that preserves the source-acting (Hiatt, 2014, para. 6). While discussing "The Two Towers", creature facial lead animator Bay Raitt acknowledges the importance of the source-actor's contribution, but also advocates for the significant role animators played in the final performance of Gollum:

> They took the data as a starting point, embellished it, sweetened it, and made sure the intent of the performance was really clear. Sometimes the footage was so blurry or bad, or the direction had changed, that it was entirely keyframing. It is important to note that there is no facial motion-capture data, at all, on Gollum. The only motion-capture data is for his torso, legs and arms. (Singer, 2003, para. 8).

[70] In my interview with Serkis, he is quick to distance himself from some of the controversy generated by his frequent use of the term "digital makeup", pointing out that this is a phrase originally developed and promoted by Weta Digital (personal communication, June 11, 2015).

Serkis frequently compares the concept of digital makeup to the full-body prosthetics worn by John Hurt in "The Elephant Man". According to Serkis: "[t]he makeup was incredible, but no one disputed the performance was Hurt's." (Hiatt, 2014, para. 6). However, Gollum in "The Lord of the Rings" differs from the nature of Hurt's performance in a crucial way: while Hurt can claim full ownership over the finished performance, Hurt's performance could not be post-processed in the same way as Serkis' performance of Gollum. This is an important distinction because Hurt's performance was essentially "locked in" at the time of filming, whereas the final performance of Gollum could easily be exaggerated or altered beyond Serkis' source-acting.

However, the techniques for filming PeCap performances have significantly changed since the early experiments with Gollum. Today, PeCap actors are often filmed on set or location interacting with actors producing profilmic performances. This provides a more integrated template to work from than the old system where PeCap was recorded separately from other actors' performances. As Serkis (personal communication, June 11, 2015) explains of his experience working with director Matt Reeves on "Dawn of the Planet of the Apes", the PeCap actors are filmed on location and their performances are then cut and pieced together through standard film editing techniques. This results in a version of the film that tells the whole story as it will eventually appear on screen, but with the source-actors not yet replaced by their digital counterparts. This version of the film determines the full arc of each character, and it is screened before many people in order to get final story approval before the visual effects team comes on board. In this production scenario, the acting choices are fully determined during the filming and editing stages, and it becomes the role of the visual effects artists to match those performances as closely as possible through the digital characters who eventually replace the source-actors in a scene. Within this type of production strategy, it is rare for any new character performance to be determined during postproduction, whereas this was a more common scenario in earlier films featuring motion-captured characters. It is in these more recent examples of PeCap production that the concept of "digital makeup" becomes more appropriate, but this still does not fully attend to the decisions that animators must make when translating performances from human source-actors onto characters of significantly different physiologies. Also, the question remains: are ex-filmic character performances more likely than vactor performances to be modified from the source-actor's original intent?

The role of the MoCap animator is partly based on the quality of the performance data. For instance, if the data quality is accurate – that is, it has minimal noise and produces a one-to-one correspondence between the movement/expressions of the source-actor and digital counterpart – then the MoCap animator's role is primarily to make minor adjustments to create the "final 10%" of the performance, especially in terms of animating features of the digital character that vary from the source-actor. However, if the performance data quality is compromised or the director requests a performance adjustment, then the MoCap animator may also need to spend time adjusting the automated retargeting of expressions and poses to more closely match the source performance. This is also a consideration if a digital counterpart's physiology or proportions do not closely match those of the source-actor. These scenarios are time-intensive processes that require significant amounts of artistic judgment on behalf of MoCap animators, including the ability to perceive, interpret, and translate both large-scale and small-scale performance details. For instance, the bodies and faces of the Na'vi characters of "Avatar" have significantly different proportions to their source-actor counterparts. The Na'vi have broad, feline noses and eyes, and stand approximately 2.5 metres in height – approximately 1.5 times taller than an average human. Flueckinger notes that such proportional differences between a digital character and a source-actor can pose "significant problems in the scaling of the movement data because the pattern of movement changes in relation to age, height and, most significantly, mass." (2011, p. 19). The Na'vi also lack prominent eyebrows, which complicates facial animation because the eyebrows play a significant role in the recognition of many expressions (2011, p. 21). In a detailed analysis of Zoe Saldana's character Neytiri, Thompson and Bordwell note numerous differences between Saldana's source performance and Neytiri's finished performance (Figure 34). For instance, they attribute the effectiveness of Neytiri's snarl to exaggeratedly long canine teeth and suggest that her mouth is proportionately smaller to her head by comparison to Saldana, with a less prominent tongue. The fact that both Neytiri's tongue and lips are blue further

Figure 34. A comparison between Zoe Saldana's input performance and the finished performance translated onto her digital counterpart, Neytiri, in the movie "Avatar".
Retrieved from http://www.davidbordwell.net/blog/wp-content/uploads/Avatar-mo-cap-21.jpg

contributes to an increased contrast with her teeth; in turn, her snarl is more prominent and well-defined than Saldana's input (2010, para. 8). Throughout the rest of her face, there are significant differences to the wrinkles on her nose:

> [t]he wrinkles seem to be derived from canine or feline faces as well, extending from the inner end of the eye and arcing down toward the tip of the nose. The human frown lines at the lower center of Saldana's forehead are transformed into larger, longer, curved wrinkles at either side; these start between the eyebrows and move up and to the sides. There they get extended by the curved areas of darker blue that radiate across the upper forehead, so that the lines of anger seem to cover more of the face. […] [E]ven the shape of Saldana's naso-labial folds has been slightly altered (2010, para. 10).

The degree of realism within a character's design strongly affects how closely the digital character's performance must follow the source-actor's performance, with more stylised characters (such as Tintin) allowing more room for performance interpretation on behalf of the animators[71] (Figure 35). In the Blu-ray special feature "Animating Tintin", VFX supervisor Joe Letteri describes this process as looking for the "caricatured elements of the design" to determine what it is about Tintin, for example, that makes him unique and

---

[71] By comparison, the Na'vi possess both realistic and stylised design elements, at least in the sense that the aliens have no real-world counterparts to compare against. Therefore, Neytiri is more open to performance interpretation than a fully realistic character such as Tarkin in "Rogue One", but less so than a stylised character such as Tintin.

Figure 35. A comparison between Jamie Bell's input performance and the finished performance translated onto his digital counterpart, Tintin, in the movie "The Adventures of Tintin". Retrieved from https://d13ezvd6yrslxm.cloudfront.net/wp/wp-content/images/tintinmotioncapture.jpg

exaggerating those features to "make the character come alive on the screen." (Spielberg, 2011). Previs supervisor Jamie Beard indicates that the actor's performance is the first step, but "from there, in a very animated way, we have to go through and sculpt it, and just make it feel right on the character itself." (Spielberg, 2011).

If part of a performance needs to be completely redone, such as with the dishrack actions in the BOE performance, then the MoCap animator is also responsible for crafting these actions, ensuring that they match the rhythm of the performance-captured animation, and that the keyframed actions seamlessly blend between the PeCap data. In the case where a character possesses features wholly different from the source-actor, these features must be animated as an extension of the character's range of expression. For instance, Neytiri's long and pointed ears resemble those of an enraged animal and significantly contribute to the impact of her snarl expression, further pushing it beyond Saldana's source performance (Thompson & Bordwell, 2010, para. 9). Cameron describes these animation contributions, combined with Neytiri's tail swishing back and forth, as providing "100 percent of what [Saldana] did, with a turbocharger on top of it." (Duncan, 2010, p. 138). However, despite Cameron's public promotion that the source-actors' performances fully translated onto the Na'vi characters, much of the eye and brow data were replaced with keyframe animation. Additionally, Sito notes that the movie's crucial scene in which Jake confesses his mission to Neytiri was heavily reworked through keyframe animation (2013, p. 213).

“Warcraft” director Duncan Jones explains that with increased fidelity of PeCap, the finished performances are less reliant on keyframe animation. For him, this is a benefit because he believes keyframe animation often leads to uncanny results when paired with performance capture (Robertson, 2016, p. 40). Freeman suggests that the Uncanny Valley phenomenon in motion capture “may be a direct result of the common practice of treating or ‘cleaning’ the raw data with ancillary animation” (2012, p. 46) – that is, additional animation that is necessary to improve upon deficiencies in the underlying raw data. While some highly skilled animators have demonstrated their abilities to produce sufficiently nuanced keyframed performances, this is a rare skill. Such nuanced keyframe animation also requires a significant time expenditure, as is demonstrated in the experiments within this PhD research.

MoCap animators must also create interactions with props and sets, both in the real and virtual worlds. When digital characters must interact with real-world sets, the ideal scenario is where the PeCap is also directly recorded on the real-world set. However, on-set lighting is rarely optimal for passive optical tracking (which is a core reason why active optical tracking was later developed). When source-actors vary by proportion to their digital counterparts, scale differences are distinctly noticeable in the ways they interact with a scene. For instance, in “Avatar”, Sam Worthington is significantly shorter than his Na’vi counterpart; this required specialised setups and multiple takes when filming a scene in which his avatar is confused and disoriented inside a medical bay (Flueckinger, 2011, pp. 24-25). MoCap animators are responsible for ensuring that the digital characters don’t interpenetrate walls, floors, or other props and set pieces. This is important in situations where the floors are sloping or irregular, especially if the digital character’s proportions do not match those of the source-actor.

With increased fidelity of performance capture, MoCap animators are more frequently instructed to make no adjustments to the performance data without explicit approval from the film director(s) (Robertson, 2016, p. 40). For instance, on “Warcraft”, animation supervisor Hal Hickel notes that the raw PeCap data was often in little need of refinement:

“[t]he great thing about our facial capture is that animators can concentrate on the 10 per cent [sic]; the sweet stuff.” (Robertson, 2016, p. 39). “Avatar” animation supervisor Andy Jones describes the last ten percent in terms of adding minute details missing from the facial capture, such as “a quiver of a lip, a slight movement of one eyebrow. […] We weren’t changing what was there to begin with; we were just trying to get every last nuance out of that original performance.” (Duncan, 2010, p. 135). Additionally, proprietary software innovations such as ILM’s Snap Solve proved especially effective at enhancing the realism of the orc performances in the film. Snap Solve works by matching tracking dots on a source-actor’s face to rigid solves of the jaws and eyes, automatically moving the performance frame-by-frame to fit the character’s physiology while conveying a sense of external forces acting on the body. Visual effects supervisor Jason Smith describes a scene in “Warcraft” where the character Draka lies on her side:

> [g]ravity is pulling her face down toward the ground, and we needed to sell that weight and gravity. That sag might not make it onto the controls for a smile or a squint. Snap Solve made the difference. All the nuances made it into the shot. (Robertson, 2016, p. 38).

However, ILM and Weta Digital are among the very few VFX companies capable of creating such soft tissue work that can hold up to scrutiny in a close-up[72] (Robertson, 2016, p. 40). With these improvements to facial capture on “War for the Planet of the Apes”, Weta Digital has shifted its rhetoric surrounding performance capture from accurately reproducing physical movement to matching what a performance ‘feels’ like. Instead of focusing on a direct translation from the PeCap data to the digital character, “it is the individual animator’s choice as to how much of the solve data they might use in the process of building up the performance they are matching. For them, it’s about the result, not the process.” (Failes, 2017, paras. 13-15). Here, animators are given greater artistic license in terms of how to craft

---

[72] Since these solutions are proprietary and significant points-of-difference for these companies, it is unclear when similar solutions will become available for broader usage within standard 3D software packages such as Maya.

the finished digital performances, so long as those performances still "honour" the source-actor's choices.

6.5 Complex Acting – Dramatic Performance

Chekhov's "The Seagull" epitomised the conflict over the necessity for a paradigm shift in the theatrical conventions of the late 19th century. I believe Treplev's monologue about "new forms"[73] also serves as an apt reference and metaphor for the new forms of acting that result from performance capture technologies, and therefore I chose it for my dramatic performance in this research. I also have previous experience of this role, having played the character Treplev in a production of "The Seagull" at Albion College in 2002. The monologue begins with Treplev plucking flower petals to the line "She loves me, she loves me not" and ends with "I imagined I could read their thoughts, and I was going through agonies." (Chekhov, pp. 7-8). The full monologue is reprinted in Appendix X.

6.5.1 Summary of the Dramatic Performance

This monologue features Konstantin Treplev, an impoverished playwright, confiding about his artistic struggle against the conventions of the "modern" Russian theatre of the late 19th century. Treplev's speech eventually centres on how poorly he feels he was treated growing up around his mother, Arkadina, a famous-but-fading actress. (A more in-depth overview of "The Seagull" is provided in Appendix VI.) As indicated in Section 5.4, I rehearsed this scene with Dawn Glover in the week prior to the performance, and she directed me on the day of the performance. Her presence on set also provided me with a tangible scene partner for the character Nina.

Prior to each take, Glover stood at one corner of the MoCap studio and pretended to be

[73] It is worth pointing out that Treplev is primarily concerned with producing heavily symbolic plays, whereas the new forms of Chekhov and Stanislavski were embedded in realism.

a campaigner for various causes that I oppose and which cause me anxiety. The goal of this exercise was to trigger within me a headspace (or behavioural pattern) similar to Treplev's emotional response to the state of the late 19th century theatre. I also participated in a "Vesuvius" exercise taught at the Kacie Stetson Studio. As described in Stetson's acting classes, a Vesuvius is an eruption of a particular "core emotion" (within Stetson's teaching, these include love, hurt, anger, and fear). A Vesuvius is most often used to channel anger, which involves the actor jumping up and down while screaming, yelling, roaring, spouting profanities, and any other techniques that quickly affect a behavioural pattern consistent with anger. However, a Vesuvius can be modified to channel any emotion, including anxiety (a combination of hurt and fear). Glover only decided on this approach during the final two takes (of eight). My "anxiety Vesuvius" produced a palpable sense of anxiety throughout my body and mind while I tore the corners off a piece of paper (substituting for flower petals) during the start of each of these takes. Glover encouraged me to "act faster" (e.g. to quicken my pace and eliminate pauses) and to abandon the introspective approach to the scene that I originally prepared. Speeding up the monologue produced a better flow of ideas and emotions, as well as a greater sense of urgency. While serving as my scene partner for Nina, Glover provided me with specific hand gestures whenever she wanted me to alter the tempo of the performance (usually in the case of speeding it up). Being able to work so closely with the director during the actual performance recording is quite unlike other forms of acting. Because Glover's presence was not recorded by the MoCap system, she was able to immediately direct the performance to degrees not normally achievable on stage or in film. By working with real-time feedback from a director, this processes provided me with an opportunity to engage with this research beyond just myself as the sole actor.

Glover advised that I should have a distinct emotive reaction to each of the "loves me, loves me not" moments. This part of the scene specifically refers to Treplev's fraught relationship with his mother. Experiencing the behavioural pattern of anxiety helped to physically pull my body inward: my arms tucked close to my body with my elbows touching my torso. This physicality naturally resulted from my authentic experience of anxiety, which made me feel small and that I should not take up much space. This is consistent with my experience of anxiety in general, where my whole body balls up; I curve my spine and I look down, not wanting to see anything or look at anyone. In the performance, my shoulders are hunched and rotated forward. There is little energy in supporting my body as I try to disappear into myself. These physical qualities are a direct result of the emotional preparation exercises I performed prior to each take. Such preparatory work helped to emotionally "jumpstart" me into the scene to match as closely as possible the needs of the character at this moment in the story. For the final "she loves me not", Glover suggested that instead of making the moment into a sudden discovery that Treplev's mother doesn't love him, I should think of it as a forgone conclusion that Treplev has long entertained, and for which this moment simply provides an excellent example for Nina's sake. In our production, this moment involves Treplev being honest with Nina about how he feels toward his mother. Glover consistently provided advice for me to breathe, take in an "emotional hit" (e.g. authentically experience an emotion), and let it affect me before moving on to the next part of the scene. Originally, I had jumped too quickly between one moment to the next without producing clear emotional sincerity, but in the take used for this research, I was able to communicate my thoughts and emotions to the viewer at each moment. Unlike in previous takes where I said the final "she loves me not" with a sense of humour to mask Treplev's hurt, in this take Glover advised that I should simply state it matter-of-factly. She indicated that I should let myself sit with the emotions that result from this admission for a moment before realising that perhaps Treplev had revealed too much too much to Nina. The purpose for the next line ("You see, my mother doesn't love me") then becomes an attempt to deflect from the awkwardness that Treplev's admission creates.

I performed the line "When I'm not around, she's only thirty-two" with a pause, turning it into "When I'm not around, she's… only thirty-two." When I performed these lines, the delivery felt natural to me. In fact, the pause came as a result of me taking a moment to think about what age Arkadina actually behaves like. However, when watching the performance, I feel this delivery seems more like I momentarily forgot the line, which was not the intended effect. This is an example of emotional authenticity versus sincerity in action: while I felt authentically connected to what I was saying in the moment, the outward signs (including timing) of the delivery didn't accurately convey that to the audience, and instead incurs an unintended alternative meaning (whether this counts as a form of hypersemioticisation is debatable). The question is whether I could edit out this pause and maintain a consistent performance. This also raises the question that if a director believes an actor's emotion could have been pushed deeper or held longer, or the signs of the actor's breathing made more or less apparent, to what extent can the existing performance be altered through animation while preserving a continuous sense of performance? Such conceptions of acting essentially equate it to an element of animation, implying that every component of the performance is capable of being adjusted or perfected in some way.

In animation, as with acting in general, any movement needs to be justified: there needs to be a reason that underpins each movement. Such reasons are often internal motivations, but the actor (and animator) must be aware of the reason and not just produce movement for movement's sake. Sometimes an actor will produce motions, especially through the head or hands, that deflect his or her emotional authenticity and/or sincerity, which diminishes the clarity of the actor's intention. In the take used in this research, I produced a lot of hand and arm movement (starting from "When I'm not around, she's only thirty-two"). It is conceivable that if given the opportunity, I would look at how to pull back some of this movement to create more stillness. This is a big difference between working in film versus on stage: in film, there is more need to remain still due to the actor's containment within the frame, as well as the need to remain within the camera's focus. When approaching this scene, I didn't worry about any framing and performed it more as if I were on stage. The result is a physically larger performance. This could work fine if the director is comfortable with producing all of the camera framing in postproduction. However, this would be problematic if the framing had already been locked down in the storyboard. In such a scenario, the actor should be given an indication of the size of each shot so as to modulate his or her performance. When shooting on location versus in a MoCap studio, it is important to be cognizant of what the final shot will be before producing the motion capture, even if the cinematography is still being virtually explored through the Omniscient Frame.

6.5.2 The Dramatic Performance Process

The dramatic performance varied from the BOE in several technical ways. The primary difference is that rather than using the MotionAnalysis passive optical system, the body motion data in this experiment was recorded using a Nansense Indie full-body motion capture suit with Nansense Pro gloves. The Nansense system is an inertial system (discussed in Section 2.8) and the gloves provide accurate motion tracking of hands and fingers that is superior to what is possible with a passive optical system. However, while a passive optical system is able to constantly determine the exact location of tracking markers in 3D space, an inertial system relies upon sensors in the form of gyroscopes, accelerometers, and magnetometers. The inertial system does not measure position but rather the rate of change between sensors, which can lead to inconsistencies in the data whereby some part of the performer's body appears to move ("drift") in an arbitrary direction (Solberg & Jensenius, 2016, p. 470).

Due to the issues encountered with the Vicon Cara in the BOE performance, I elected to use just the Faceware PeCap system in order to maintain at least one form of technological consistency across both experiments. The Faceware camera was mounted to a standard Faceware helmet, which means that it was correctly positioned to be in-line with the performer's nose.

The following subsections build on the principles of performance tracking, clean-up, and retargeting in Analyzer and Retargeter that are addressed in subsection 6.2.3. As a result, these subsections specifically address only insights and technical issues that have not been previously discussed.

6.5.2.1 The Dramatic Performance in Faceware

The dramatic performance is three minutes and twenty-six seconds in length. However, based on the amount of time that the BOE took to complete, I realised I would not be able to produce the full dramatic performance as an animation. To this end, I selected three sections of the performance that I felt demonstrate a significant range of emotions and facial expressions. I tracked these sections separately and saved their tracking results. When it came to retargeting the data, I ended up retargeting only the first section because finishing even one section on my own required hundreds of hours of work. The first section is 39 seconds long and demonstrates a nice range of subtle dramatic performance. The other two sections would have provided more insights into a broader emotional range, but I believe that I have gathered all meaningful insights from the first section alone. The other sections were 39 seconds and 19 seconds, respectively.

Unlike the BOE where I preserved and animated all of the facial data at the original 60fps, I converted the dramatic performance to 24fps. This was to reduce the amount of overall work and because there was no clear benefit to working at 60fps within the scope of this project.

My method for creating the poses via Retargeter was to work from the controls that had the greatest overall effect, and then to gradually work my way to controls with more localised effects. The challenge with the lips is ensuring they have the same overall shape, curvature, and position as in the source-footage. This is difficult because the lips have a high degree of natural flexibility but the rig is limited by its controls; producing a perfect match is not always possible. There is also the need to achieve deformations around the lip corners (especially when the lips are pulled wide), which is a product of the mouth movement combined with compression in the cheeks along the naso-labial fold. The mouth must also appear to wrap around the teeth, so it is important that the lips don't move in such a way that they feel detached from the gums. A further challenge is in achieving the appearance of lip compression, especially when they purse together. The lips easily intersect each other[74] and the appearance of intersections needs to be minimised when animating compression.

---

[74] There is no automatic system in Maya that prevents such intersections. However, Kozlov et al. describe a physical simulation method that corrects lip self-collisions (2017, p. 82). Like other advanced facial solutions described in this research, these results have been primarily adopted within proprietary VFX industry software and are not available within off-the-shelf software, such as Maya.

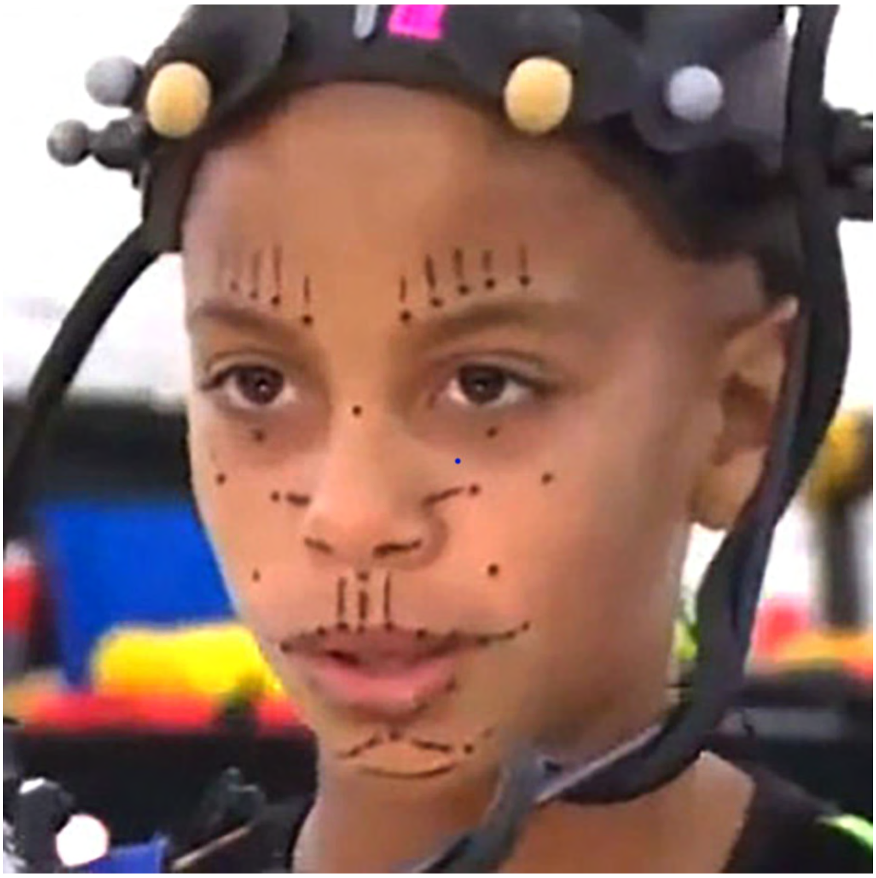


Figure 36. A facial marker pattern used by Electronic Arts when recording performances with Faceware. From *Faceware 2018 Reel*, Faceware Technologies, Inc, 2018.

It was ultimately necessary to create a high-resolution colour texture for the face, especially to help with the animation of the lips. Without the texture, it is difficult to know precisely where the *vermilion border* occurs, which has a large influence on visualising how compressed, flat, or curved the lips are. This proved especially important on the upper lip. The colour distinction between the lip and the surrounding skin is the most important visual factor in this regard.

One of the more difficult aspects of posing the eyeballs was determining the position/rotation of the pupil, which is often obscured by the eyelids and eyelashes. As mentioned in subsection 6.2.3.2, it is also critical to get the correct proportional distance between the eyelid and the iris/pupil to correctly register an expression. Adjusting the eyelids is not done in isolation and needs to factor in movement from the *orbicularis oculi* muscle, especially. It is also important to match the dilation of the pupils in the 3D character with the pupils in

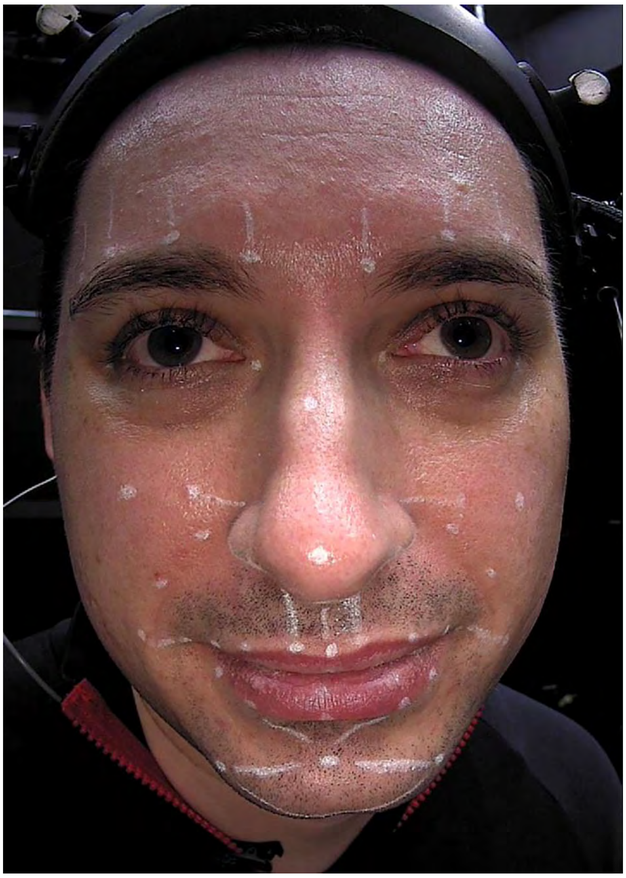


Figure 37. A facial marker pattern similar to the Electronic Arts example applied to my face in the production of the dramatic performance.

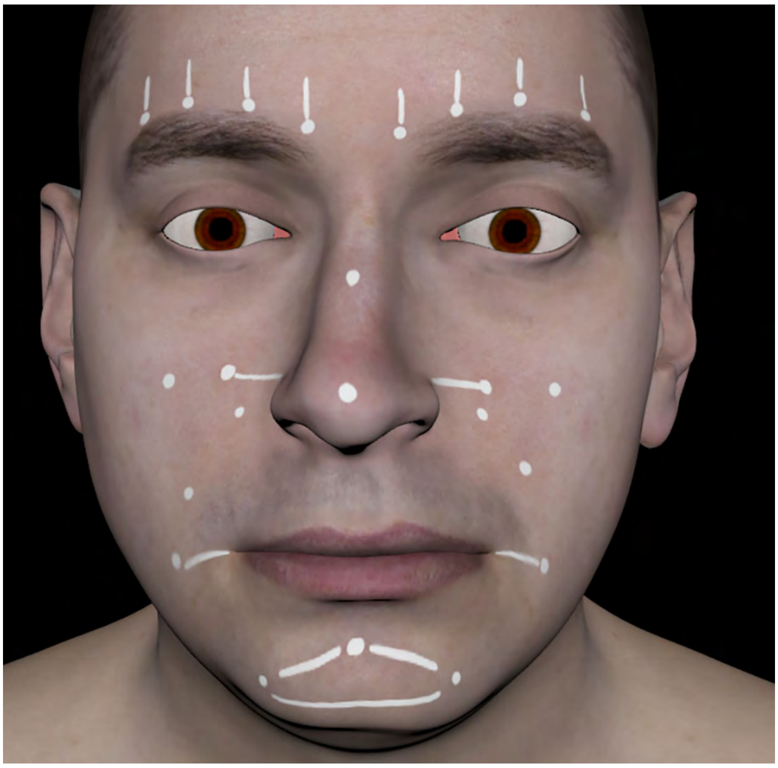


Figure 38. The Electronic Arts facial marker pattern replicated as part of the facial texture in Maya.

the source performance because this affects the proportional distances between the eyelids and pupils.

Prior to recording the dramatic performance, I used eyeliner to apply a facial marker pattern utilised by EA Games (Figure 36) to myself (Figure 37). This pattern features a series of dots and lines that are specifically arranged to highlight movement in key facial regions. The dots on the cheeks were especially helpful to gauge the amount of movement and displacement in those regions, since there was little other detail in those areas to reliably compare against. Initially, I did not include these dots in the texture file I applied to my model, but I later updated the texture to superimpose a replica of the real-world dot pattern (Figure 38). This proved invaluable in matching the expressions with greater reliability, especially in the area immediately around the *oral commissures*.

I discovered that when posing the eyes, I was often unable to achieve the same shape with my rig as in the source performance. I determined this was due to the fact that when I create the blend shapes for the eyelids, I modelled these while looking in a mirror. When compared to the source video, the eyelids in the original performance open wider and with slightly different shapes than what I achieved with just the mirror. I suspect that the difference in size and shape was due to an increase in tension in the eyes when attempting to maintain an expression in front of a mirror. I learned that it was a better idea to sculpt the blend shapes based on the range of motion in the source video than just relying on creating expressions in front of a mirror. Also, ZBrush defaults to an orthographic camera for modelling, which does not provide satisfactory depth information. When I remodelled the eyelid blend shapes, I used a perspective camera in ZBrush and matched its focal length to that of the original Faceware camera (approximately 45-50mm). I also resculpted the base shapes of the eyelids, which led to easier and more accurate posing in Maya.

Mirrors are typically useful as a technique for modelling more cartoon-style animated characters. However, modern practices for modelling realistic human models frequently utilise photogrammetry or other forms of digital scanning. This project has evolved through a number of stages, having started with digital sculpting from fully photographic reference, then moving to photogrammetry. While I captured photogrammetry for all of the FACS shapes, I did not create blend shapes from the FACS data (due to time considerations and the fact that I had already produced blend shapes using the mirror method). It is arguable that had I used photogrammetry from the start of the project, including for creating the blend shapes, I may have achieved more accurate physiological results. However, AUT did not possess the photogrammetry technology at the start of my project and only deployed it after I had completed the BOE animation. Modelling from video reference is also quite different from modelling with a mirror. This is because when modelling with a mirror, it is difficult to achieve some facial shapes at their extreme limits (such as looking down or looking up) because it is not possible to produce those shapes while looking at the mirror. However, it is easy to see these shapes in the source video due to watching a recording of the action rather than producing the action while trying to model from it. When working from a video, it is possible to study a shape for as long as needed without worrying about muscle fatigue. For instance, when creating a series of facial shapes with my eyes using the mirror method, I often noticed my eyelids tiring from overuse of the associated muscles, which obviously impacted my ability to perform a shape to its fullest extent. Blend shapes have to be "road-tested" – that is, tested through animation in order to determine their effectiveness. Normally, the rigging artist would test this out before handing the rig off to the animator, but in this production, I had to serve both roles and did not get a chance to create test animations. Therefore, many of these fixes were made throughout production rather than in a preproduction period.

I produced the first retargeting from a combination of 15 poses, all of which featured poses across the three face groups (brows, eyes, and mouth). The result produced a lot of jitter, most notably in the eyes group. The jitter became more pronounced whenever the eyes looked down. It is unclear why there is so much jitter in this retargeting given that extensive care was taken to accurately track and retarget the performance. I examined the entire performance and made note of frames where there existed significant variances between the retargeted shapes and the original performance. I updated the existing poses to reflect these changes, and made new poses at other problematic frames (generally for individual face groups rather than a full-face pose on a single frame). Then I retargeted again to get an updated result. I repeated this process several times as I continued to refine the pose shapes and add new poses wherever necessary – a total of 14 new shapes. This also helped to identify facial shapes that required custom blend shapes to achieve. Altogether there were 29 facial poses over a 39-second interval that define the performance range of this animation. The amount of noise in the data made it somewhat challenging to evaluate the quality of the overall performance. This noise was likely due to the clip length and may have been reduced if the animation interval had been divided into shorter segments. Once I got the animation to a point where I felt I achieved what would be the best possible result via an automated solution, I proceeded to manually clean-up the rest of the animation.

I evaluated the retargeted performance on every frame and made manual adjustments to all of the controls on a frame-by-frame basis in order to ensure a "close-follow" to the original performance. The result is that the Faceware PeCap data did not provide an exact translation of facial performance (at least in this instance). Instead, it provided a close approximation of the original performance[75], but one which required significant adjustment. A case could be made that it would have been faster, easier, or more accurate to have simply used keyframe animation from the start. The advantage of using the PeCap system is that once trained, it automatically makes adjustments to all of the facial controls that place them "near" the correct shapes on every frame, which functions as a guide for the animator and could be time-saving in some respects. However, a sufficiently talented animator may find this as much of a burden as a boon, and simply opt to delete all of the data to create keyframes from scratch. Where the Faceware PeCap approach may have greater merit would be if the training poses created for Retargeter were added to a shape library that can be used as the basis for tracking other takes, whereby the poses created in one take are used as a basis for automatically retargeting other takes (with other custom

[75] In my talk-aloud recording of my process, I described the close approximation as "in the same neighbourhood, but a few blocks over."

poses added to fill in any gaps in the range of performance). This makes more sense in terms of how a system like Faceware is typically implemented across much larger productions, especially video games, which are worked on by many artists across hundreds or thousands of takes. My use of the technology in this project is perhaps niche and not necessarily representative of its more common applications.

6.5.2.2 Summary of the Dramatic Performance

Due to the sheer volume of work of cleaning the animation on every control, I eventually decided to alter my methodology by introducing one animation assistant into the project. This person had just completed her second year of the animation pathway within the Digital Design major at Auckland University of Technology. She consistently performed at the top of her animation papers and was even awarded a summer internship with Weta Digital following her first year with AUT (a unique achievement). She had also taken a facial performance capture paper I taught in the previous semester, and so already had experience working with Faceware in a similar (albeit, less demanding) workflow. I believed that she would work well as an assistant and could communicate her process to me in a similar fashion to how I had been tracking my own work during that time. I obtained funding from the university that paid for 100 hours of her work. By engaging with an assistant, this also provided an opportunity to open this research to go beyond just my own experience.

I provided my assistant with an overview of the rig and specific instructions as to what I was trying to achieve. I instructed that her role was to specifically help complete the animation clean-up process with the goal of creating a one-to-one correspondence between the facial expressions/motions in the source video and the 3D character. I indicated that at no stage should she "don her animation hat" and try to interpret the resulting motion in order to make the result *feel* more like the original performance.[76] I outlined specific controls that I had not yet cleaned up. Having my assistant perform the remaining clean-up allowed me to

[76] This was a task I wanted to reserve for my own interpretation alone.

focus on writing the thesis component of this research. I also asked her to make note of any instances when she encountered a moment in her workflow that she considered noteworthy and to produce a video reflection on it. The full documentation of instruction provided to my assistant is included in Appendix VII. I periodically provided feedback about her work-in-progress to ensure that it was consistent with the standards I employed in my own work.

6.5.2.3 Cleaning the Body Data

When I received the body data from my assistant, I noted the fact that the arms interpenetrated the character's belly to a severe degree. I initially suspected this error may have derived from inconsistencies within the inertial PeCap system, but as my assistant indicated, cleaning the data resulted in adjustments to the overall positions of the arms. I also recognised that the positions of the arms were consistent with a thinner version of myself (that is, with less belly to interpenetrate) which varied from the BOE performance. In the dramatic performance animation, I was still using the original model of my body that was created earlier in the PhD when I weighed more. To make the character consistent with my current weight, I made some basic proportional and shape adjustments to the character's midsection. This prevented the arms from interpenetrating the belly somewhat but did not provide a complete solution. I also realised that I had not created a blend shape for the arms pointing straight down along the sides of the torso. When I created this, the shoulders still felt hunched rather than slack. To account for some of these differences, I translated the clavicle controls down to produce slouched shoulders consistent with my pose throughout the original performance.

To guide me in terms of which parts of the body to concentrate on at any given time, I repeated the process used in the BOE of first establishing my cinematography using Maya's Camera Sequencer. This also has the effect of heightening the drama at certain moments, and provides opportunities to examine how the character's acting is received in a variety of compositions. It also demonstrates some challenges when cutting between compositions for certain actions. For instance, in a tight composition, the actor would normally confine his or her movements. However, when a performance is solely recorded via the Omniscient Frame, the actor may move without regard to any particular framing. As a result, if a tighter composition is wanted later, the actor's movements may be too broad for that type of shot. It then becomes the director's decision whether the animators should try to reduce the overall degree of body motion to better fit a tighter composition.

While PeCap allows actors to be more stage-like in their conception of a performance space, the reality is that there must still be a filmic sense of delivery. While it may not harm the performance for an actor's eyes to wander on stage, when it comes to film, the direction of the actor's gaze is all-important. Eyelines are critical, and any small divergence of focus is scrutinised for meaning. Screen actors are constantly aware of their position within the camera's framing. To get optimal results in the performance data, the director needs to have a clear sense of the final cinematography and communicate this to the actors. While eyelines can be adjusted in postproduction, this introduces other artists' judgments into shaping the final performance.

So much of the vitality of the dialogue lives inside the mouth; without the tongue's movement, there is no liveliness to the speech. I remodelled, rigged, and keyframe-animated the tongue to create a clear sense of speech being produced within the mouth. Even though it is barely perceptible, the slight movement of the tongue inside the mouth is pivotal to the believability of words emanating from the digital character; it makes the performance of the digital character feel more rooted in reality. This is especially true for any sounds that involve the tongue visibly touching the roof of the mouth, such as the sounds of D, J, L N, S, T, and Z in English.

There were also issues with the *oral commissures* where the digital character's mouth was spaced too far away from its teeth. This was, in part, due to the teeth not properly conforming to the specific shape of my mouth. The original teeth and tongue models were third-party assets that I included (without modification) in the scene. Adjusting the overall proportions of the teeth to fit my skull resolved the issue of the distance between the *oral* *commissures* and the teeth. A few instances persisted where the lips did not appear to fully close in order to produce a particular sound, especially F, M, and V sounds. The mouth required adjustment to ensure the lips fully closed and remained "sticky" for these shapes, as well as some other expressions.

6.5.3 Analysis of the Finished Dramatic Performance

Once the digital performance closely matched the source performance, I analysed the result to determine where the performance seemed uncanny, or where expressions could be pulled back or pushed further. While the data preserves what I actually did in the performance, the reality is that the performance could easily be improved upon. This results in a less *truthful* but better overall performance. For instance, the character's attention is not well directed in the final shot presented in this experiment. The character looks all over the place and I feel the emotional sincerity would be enhanced if a consistent eyeline is maintained from the line "When I'm not around" until the end. This is a good example of how when acting more from a theatrical point-of-view, with less focus on specific shot framing, the actor's gaze may wander.

A few other notes I made were to create nostril flares at a moment when there is an audible inhalation of breath through the nose in order to better ground the audio to the actions of the scene. At a cut between shots three and four, there was a lot of physical movement within the character's body, which appeared out of place at the start of shot four. This occurred mostly during just the first four frames of shot four, so I elected to move the problematic body actions four frames earlier to the end of shot three.

I also noted several frames at which the overall shape of the lips needed to be adjusted. For instance, a mouth shape approximately 9.5 seconds into the original animated performance involves the left and right corners of the lips moving at different times – offset by only a few frames, but enough to be noticeable. This is consistent with the source performance, although this idiosyncrasy feels more prominent and less believable in the digital character.

As I evaluated this moment in the original acting, I realised that it is an example of a chiral expression. This likely resulted from me experiencing multiple feelings at that moment during the performance – or possibly not being emotionally connected. It could be argued that leaving this moment in the final animation may communicate that the character is dealing with conflicting emotions. Conflict is good, but an expression shouldn't lead to the audience's confusion. It is possible this moment would read as an error and I feel the performance would be better if the expression were singular and clear. I elected to retime the right mouth corner to match the movement of the left mouth corner in order to eliminate the chirality of the expression.

Modifying the acting in this example is not as ethically problematic because the chirality was an unintended aspect of the expression. It is similar to an actor watching his or her performance and recognising a missed eyeline, or too much blinking during a critical moment, or an unwanted involuntary reflex: the rest of the acting may have been excellent, but one peculiar moment may spoil the scene. As an actor, I am fully comfortable to remove the offending element so long as it does not impact on the rest of the acting in the scene. I do not feel like it hurts the acting – in fact, I feel like it improves it. In this case, it is easy to obtain the actor's permission to make such a change because I am the director, animator, and actor. However, in more common situations where these roles are separately performed, the director would usually make this call – possibly without the actor's permission (or even knowledge of the change). Animators almost certainly would never participate in a feedback loop with the actors to inquire about their thinking and intentions at any given moment. This does raise the question of what kind of ethical responsibility a director (and animator) has when modifying an actor's performance in this manner. Is this ethically any different from the history of modifying screen performances through other filmic techniques? This question is further explored through subsequent sections.

6.5.3.1 Feedback from Acting Director

As the sole actor and animator in this research, my analyses of my own performances are inherently subjective. As a result, upon completion of the dramatic performance animation, I enlisted Glover's help to provide feedback about the animated performance. Glover's feedback was valuable because she was already familiar with the scene, as well as with her intentions as the director and her experience of directing an actor through performance capture for the first time. I provided her with a copy of the finished animation as a single video file, as well as a version of the finished animation side-by-side a video recording of the PeCap source performance. We met via a video chat using Zoom. I instructed her to respond to what she felt was retained and/or lost between the filmed performance and the animated performance. I also asked her to specifically indicate any elements that seemed truthful or untruthful to her, but I did not provide her with the definition of *truthfulness* that I describe in this thesis; instead, I wanted her to respond based on her own professional understanding of what truthfulness means in acting. I explained that the character was rendered with a simple grey material in order for the viewer to focus more on the acting than on the quality of simulated realistic skin. Only the eyes have a basic colour texture applied to them to allow the irises and pupils to stand out from the *sclera* (the whites of the eyes), which I believe helps the viewer to better follow the character's gaze.

Glover was initially surprised by the degree to which the animated character was able to retain the qualities of acting from the source performance. She did not expect the animation to be able to capture the fine and nuanced details of the human facial poses. She indicated that due to the lower camera angle, this made it more difficult to see how much of the performance "was coming through the eyes". She noted that there was a lot of expression around the eye area, but the one area she felt was "only a little bit off" was the "depth of the eyes", but that "it almost feels there" – indicating that the result closely approximates a realistic performance result in the eyes. She was especially intrigued by the degree to which the character's "breathing" was evident. She felt that this was critical because "it probably bridges whatever [she's] slightly catching about the eyes. If the breathing wasn't there, that would be even more *off*."

I pursued her insight about the eyes further and asked her if she felt that the animated eyes resulted in a reduction in the sense of the character's screen presence. She agreed with this and indicated that metaphysically, it came down to a sense of whether the character had a soul. I discussed with her how the notion of "soulless" eyes was a common criticism of early hyper-realistic CG humans, especially the characters from "The Polar Express". She described her experience of how "when you get an actor who's tired, or you get an actor who is too into their head, that the eyes convey less." Having also performed using PeCap herself, Glover felt that the requirements of acting within the PeCap space produce

> an extra level [for the actor] to dismiss reality in the present moment because you've got so much going on. Your thinking mind is really loud because it has to be, because you've got all these technical things that you have to make sure you're doing. I guess that kind of just shows you how amazing Andy Serkis' work was[77], because I think that Gollum definitely had a lot of soul going on in his eyes, even though they were weird and freaky and not human.

Together we explored whether the soulfulness of CG eyes is due more to an accurate capture of the original performance, or if it is related more to accuracy in the modelling, texturing, and rendering of the eyes. This was a question that we ultimately left open, although the answer to it likely entails a combination of these factors. She also questioned to what degree her *a priori* knowledge of the performance as an animation rather than live-action played a role in her subconsciously looking for and identifying synthetic elements.

We concluded with a discussion about the moments in which the eyes felt "more alive" and "soulful". For Glover, these moments more often occurred when the eyes were in a state of motion rather than being held still, as well as when the character visibly breathed. When the character was "in a still stare", she felt that the performance became more "off". This is consistent with my observation about the deadening effect specific to stillness in 3D character animation. Overall, Glover's feedback is meaningful in that it validates my own observations and conclusions about my animated performance.

[77] I wish to point out that in this conversation with Glover, I did not mention Serkis. Her opinion here is solely based on her own knowledge of motion capture and Serkis' popular role within it.

6.6 Authorship Within Constructed Performances

One of the most hotly contested topics in regard to acting in motion capture is the degree to which the source-actor is responsible for the finished performance of the digital character on screen. As mentioned in Chapter 2, this becomes an especially sensitive issue during the awards season for film and television. Source-actors are frequently heralded by production companies as fully responsible for the digital characters' performances, but MoCap performances historically have not been considered for best actor or best animation categories. The animators themselves are almost entirely left out of this discussion – their contributions being pushed so far below the line that the best an animation/VFX team can hope for is their supervisors receiving a Best Visual Effects award on their behalf.

How should we understand authorship in the context of digital counterpart performances created using performance capture? Certainly, PeCap does not provide the first examples of multiple authors producing a single character performance in film. Indeed, it is common for a single film character to involve a lead actor, stuntpersons, body doubles, and dubbed voiceovers (Hosea, 2012, p. 68). Such attendant performers are rarely understood to provide the same degree of performance as the lead actor, and are usually left out of the discussion of the character altogether. In the bizarre case of the CG resurrection of Tarkin from "Rogue One", Washington Post writer Michael Cavna boldly declares: "[o]ne of the best performances in 'Rogue One' is by an actor who died in 1994" (2016). Cavna is guilty of conflating the image of an actor with the source of the acting. This reflects a confusion about actors versus the characters they perform. The performance of Tarkin is a combined effort between Guy Henry (the source-actor) and the MoCap team, and it is performed through a synthespian recreation of Peter Cushing. However, it is not a Peter Cushing performance, despite every effort made by Henry and the production team to produce that precise illusion. At best, it approximates a Peter Cushing performance. This is subtly different from the case of the digital (re)incarnation of Paul Walker as Brian O'Connor in "Furious 7", in which case his facial performances were derived from archival footage of Walker shot over the many years of the Fast & Furious franchise's production. In this case, we could more legitimately say that O'Connor's screen performance is actually a Paul Walker performance. Balcerzak indicates that filmmaking is always a fragmented process that "never allow[s] for a complete approximation of [either] Stanislavski [or] Strasberg. In a potent sense, mo-cap only aggressively heightens this disempowering of the actor." (2013, p. 202).

Is it then the case that within a performance capture context, we can only recognise a particular actor's performance if it is enacted through a synthespian double of the same actor? Or, if it is a matter of attenuation, how much of the source actor's physiology must be recognisable in the digital counterpart before we detect the aura of the source actor? As discussed in Chapter 4, within a PeCap context the aura can be understood as the kinaesthetic trace of the source-actor's movements. This would suggest that the physiology of the digital counterpart is irrelevant so long as it is able to fully translate the kinaesthetic trace, which is in itself partly determined by the source-actor's physiology. How many points of similarity do their need to be in order to preserve the aura, and are there specific parts of the body that are more important as the loci of recognisable performance? Serkis again proves a useful example because he has provided the source-performances for so many digital counterparts over the years. It is arguable that if someone was familiar with Serkis as an actor but somehow did not know he provided the source-performance for Gollum in "An Unexpected Journey", that person may not recognise Serkis as Gollum. This would likely be due to the profound differences in their physiologies (as well as the difference in voice Serkis used for the character). However, in order to make Serkis more recognisable as the source-actor in the recent "Apes" trilogy, the character Caesar was progressively resculpted over the course of the films to better reflect Serkis' physiology around the eyes and brow regions (these changes were cleverly concealed to occur at certain points when Caesar aged throughout the trilogy). This led to greater emotional connection with the character (the eyes/brows became subtly more human-like) and also made Serkis more recognisable in the character (D. Lemmon, personal communication, January 31, 2018). It is unsurprising that the more physical similarities there are between source-actor and digital counterpart, the more the source-actor will be recognisable within the character – and it may be that the eye and brow regions are the most critical to promoting recognisability.

It is interesting how the various creative people responsible for a PeCap performance vie for a greater share of responsibility for a role: actors assert that animators must honour every acting choice, while animators decry the purported fidelity of PeCap technology and seek acknowledgement for their roles in modifying/perfecting/supplementing/supplanting an original acting performance into the final screen product. It appears that all creative contributors involved feel disenfranchised by the PeCap process to various degrees. Serkis recounts how the one of the most frequent questions he is asked by journalists is whether he believes PeCap source-actors should be considered under a different awards category than standard live-action actors. For him, this question really asks whether a source-actor's performance in a PeCap role is a true form of acting when it is enhanced by other artists: "[m]y answer is that there isn't a single screen performance, whether it be live action, animation or performance capture, that isn't enhanced by other people's artistic endeavor." (2016, paras. 9-10). Yet, it is the nature of those artistic endeavours, and the contexts in which they are enacted, that forms the substance of this debate. For instance, in the case of Serkis' promotion of the term "digital makeup", Shaw-Williams notes that the controversy has been led by animators who do not work for Weta Digital, the VFX company responsible for producing the majority of Serkis' PeCap performances. Instead, without necessarily understanding the relationship between Serkis and the Weta PeCap team, they claim offense on behalf of those artists – leading to inflammatory headlines such as "Andy Serkis Does Everything, Animators Do Nothing, Says Andy Serkis" (Amidi, 2014b). While their criticism may be founded if it were aimed at other VFX companies,

> Serkis, [director Matt] Reeves and [VFX supervisor Joe] Letteri have all indicated that Weta now has a very unique, very specific approach geared towards honoring the actors as much as possible. If that's the case, then it's possible Weta's visual effects team might actually concur with Serkis' statement. (Shaw-Williams, 2014, para. 3).

In the interviews later in this section, both Serkis and Weta VFX supervisor Dan Lemmon address this process as it specifically relates to Weta's workflow. What this reveals is that animators' grievances need to be contextualised within a film's visual style and technological underpinnings. Workflows vary between VFX companies as they deliver a range of visual styles, from highly realistic creatures and characters to more stylised aesthetic choices. It could be argued that the animators who take issue with “digital makeup” have more ground to stand on in the context of a stylised role like Captain Haddock from “Tintin” than a realistic design like Supreme Leader Snoke from “The Force Awakens” and “The Last Jedi”. Serkis even admits that the “cartoon element” of Tintin affects the way in which his source performance is matched to Haddock (Ktorides, 2014, para. 28). Other characters Serkis plays tend to exist in a space between these stylistic extremes, where realism is paramount but animators must also use their artistic judgment when matching the source-performance to physiologies of non-human characters. To better clarify the different roles involved with creating the performance of Gollum in “The Two Towers”, Bay Raitt uses the following analogy: “Andy Serkis wrote the music, then [animation director] Randy Cook conducted it, and the animators are the ones playing the music that you see in the film.” (Singer, 2003, para. 3). Unfortunately, this degree of subtle distinction rarely finds its way into the discussions about performance, and it isn’t clear whether the creative individuals involved are fully aware of the nuances themselves. The result is that Serkis is perceived as talking up his acting contribution in an attempt to gain more acceptance and recognition for his craft (especially during the awards season) (Cohen, 2014, para. 26), while animators are perceived as naïve for seeking greater credit in a field where it is standard for below-the-line performance contributions to go unrecognised.

Do animators deserve more credit, especially for scenes in which they fully create performances for characters that could not be physically (or safely) produced by human actors?[78] Certainly – but the source-actor and the animators are all working toward producing a character with a seamless performance, and the source-actor is the primary driver of that performance. He or she provides the template upon which the animators will base any subsequent animation. The danger of having ten authors of a single character is that you may get ten different styles of performance for a character who should convey a consistent performance. It’s like asking ten different actors on stage to invoke the

---

[78] A couple meaningful examples of this type of supplementary animation for characters played by Serkis include Gollum climbing down the rock faces of Cirith Ungol in “The Return of the King” and Kong fighting a tyrannosaurus in “King Kong”.

mannerisms and behaviour of a single actor. As Raitt describes the problem: "[w]e didn't want him warping from his cousin to his brother to his grandpa – he had to be Gollum the whole time." (Serkis, 2003, p. 79). In this way, the source-actor provides a consistent range of performance choices for the animators to work from and uphold.

It is noteworthy how Serkis' opinion regarding his degree of contribution to a performance-captured role has changed with time. In a book he wrote about his experience of working on the "Lord of the Rings" films, Serkis describes the animators responsible for Gollum as "managing to capture the essence of what I was doing and then augment it, taking it to another level, amplifying the underlying psychology of the acting with their phenomenal talent." (2003, p. 100). This sounds far more generous to animators than Serkis' later rhetoric about honouring the source performance via digital makeup. However, many aesthetic and technological innovations occurred in the decade between these two performance appraisals, and the concept of digital makeup gradually became more accurate with time – but within the specific context of realistic films produced through the VFX pipeline developed by Weta Digital. However, as animation supervisor Randall William Cook points out, Serkis' more recent attempts to describe the animators' contributions to Gollum in "Lord of the Rings" as digital makeup is heavily flawed (Amidi, 2014a, para. 8). For instance, there are numerous examples of Gollum's acting performances that were either reworked or completely made from scratch by the animators and without any input from Serkis, often as a result of the poor reliability of the MoCap equipment at the time (2014, paras. 14-16). Cook is careful to level his criticism solely with his experience of working on "Lord of the Rings" with Serkis, and purposefully does not try to address any of Serkis' later PeCap performances. He summarises by stating that

> Andy really should be considered the principal author of Gollum's performance, but there's a hell of a difference between principal author and sole author. […] [T]he animators on THE LORD OF THE RINGS [sic] were most certainly not 'digital makeup artists', and nobody has any business saying that they were. (2014, paras. 18-19).

Cook exposes an underlying reality within performance capture that is rarely made public: that on certain PeCap films, there are some shots that are either mostly or entirely constructed by animators. It is less clear the degree to which this happens in more recent films (although Lemmon's interview below provides some small degree of insight), and it may legitimately be argued that this is a much rarer occurrence now. What seems clear is that Serkis believes that the entire performances of his digital counterparts are solely derived from his source acting. This impacts his understanding of animators as actors:

> [Animators] are actors in the sense that they create key frames and the computer will join up the dots, carefully choreograph a moment or an expression and accent it with an emotion. But that's not what an actor does. An actor finds things in the moment with a director and other actors that you don't have time to hand-draw or animate with a computer. (Appelo, 2011, para. 8).

What is left from Serkis' understanding is any time in which animators must create some aspect of a performance for which there is not usable data or reference. If the animators have to create their own performance reference and animate from that, then they are responsible for acting in the same capacity as an actor, with the additional responsibility of interpreting that performance through the animated character[79] (and possibly to match the performance style of the source-actor).

Ultimately, Serkis acknowledges that the degree to which animators are expected to interpret a performance comes down to individual projects and the will of the director (personal communication, June 11, 2015). However, Serkis' position regarding authorship in performance capture is directly related to his experience of working on big-budget productions that can afford large teams of experienced artists to work on his performance data. Such teams have the personpower and time to reconstruct any aspect of performance that was not accurately captured. As a result, Serkis can advise to play without exaggeration because even if the PeCap system is unable to fully capture his subtlety, there are sufficient artists on hand who can animate his subtlety from the video reference. By contrast, smaller productions do not necessarily have the resources to make such touchups. As a result, the actors for smaller-budget (or quick turnaround) productions may be directed to exaggerate

[79] Such as adjusting the arm movements to accommodate the dishrack actions in the BOE.

their performances (or "play to the backrow", in theatre-speak) in order to ensure their expressions and gestures are unequivocally captured.

6.6.1 Interview with Andy Serkis

I conducted the following interview with Andy Serkis on 11 June 2015 at the Sanderson Hotel in London, UK. For context, this interview takes place after the release of "Dawn of the Planet of the Apes" and before filming had commenced on "War for the Planet of the Apes". What follows is my summary of the interview, and the full transcript is available in the appendices (Section 9.7). All quotes below are from this interview (A. Serkis, personal communication, June 11, 2015) unless otherwise cited.

Serkis sees it as a misconception that in order to play a role using PeCap, you must perform any differently from the standard methods used by a screen actor. He is adamant that he makes no adjustment to his acting practice when working with PeCap rather than live action: "I've always maintained that there is absolutely no difference in the acting process between playing a role where you're authoring a role wearing a costume and makeup, and when you're authoring a role wearing a suit with facial markers and a head-mounted camera." He acknowledges that actors have a wide range of methods to approach their work, and that all those myriad ways apply when using PeCap. He cautions actors new to PeCap against the

> perception that you have to pantomime, or over-perform, or demonstrate the character. It's actually the opposite of that. It's very much about being the character and allowing the subtlety of your movement, of your physiology, of your physiognomy to play as you would if you were giving any live-action performance. So, there's not a necessity to project the character in any sort of theatrical way.

Serkis acknowledges the need to build a physical language for certain characters, such as the apes from the Planet of the Apes franchise, but that this approach to developing this physicality is the same between live-action and PeCap. The notion of over-performing or demonstrating a character runs counter to Serkis' experience of playing performance-captured roles, as he believes the stillness he holds on screen is just as important in PeCap as it is in live-action. This stillness is evident in his portrayal of Caesar, the leader of the apes in the recent "Planet of the Apes" trilogy. Caesar's actions are often understated yet remain clear and effective, defined by a meaningful glance or expressive nuance. While there are times in these films where Caesar must perform large actions, the sincerest acting occurs during these quieter and subtler moments.

While Serkis' approach to playing a role in a PeCap context may be no different from a live-action context, there are key differences in the way that performance-captured characters are intermediated through the animation process. In turn, this draws into question who is/are the author(s) of a given PeCap performance. Serkis cites a variety of performance types – including Wayang shadow puppetry, Greek mask work, *Comedia del Arte*, interpretive dance, and physical theatre – to highlight multiple ways to author a role and imbue it with life. For Serkis, animation is merely another method for actors to embody characters that possess markedly different physiologies from themselves, as well as a means for actors to transmute themselves through those characters. As Bode points out, these types of performance analogies (especially to mask work) serve to reframe MoCap as a form of "digital prosthesis", and in turn attempt to ground MoCap as an extension of existing and accepted forms of "legitimate acting" (2015, p. 90). She indicates that critics "trying to understand performance capture in relation to what is familiar, locate Serkis on screen 'underneath' Caesar or as providing Caesar's 'skeleton and soul'." (2015, p. 93). It is little wonder that despite Robert Zemeckis' efforts to pioneer performance capture in the early 2000s, it is actually Serkis who became the technology's chief ambassador. Bestor suggests that this is mainly due to the fact that while Zemeckis promotes performance capture as a "radical departure from conventional acting", Serkis instead grounds PeCap in more traditional and familiar rhetoric about acting (e.g. less likely to exacerbate the anxieties of the filmmaking establishment) (2016, pp. 183-184).

Serkis reflects on the debate about the roles of actors versus animators in the authorship of PeCap performances. Serkis has been lambasted by animators on the Internet who take issue with some of his comments that they regard as minimising the creative role of animators in PeCap performances. In his discussion with me, Serkis admits that the question of authorship is a grey area in the debate about PeCap. Does the authorship of a performance-captured character begin and end with the actor? How much of the creation/interpretation of performance is handed over to the animator? Serkis quickly points out that there is no one-size-fits-all answer to these questions. Instead, the degree of shared authorship depends on the nature of the production and the type of performance the director desires.

> [… A]nimators, like all human beings, have the potential to act, have the ability to act, to place themselves mentally in the mind of another being – another person, another creature, another character – as all human beings do, if they allow their imagination to go in that way. […] But I think the difference between animators and actors is that actors have the ability to completely bury themselves in a character […], to have extended periods of time getting under the skin of the character emotionally, and truthfully allowing themselves to lose themselves within that role. And I think that's the differentiation with animators, because I believe animators do that, but it's not required of them to do that – to actually take a script and to work on a character from page one to 120, and create the emotional arc and journey, and so on. Because that's not part of, generally, what animators are asked to do. They're asked to work on short segments of a character. Or if they're working on one character, then they're often not aware of the entire story. As an actor, you have to have a full understanding of the function of your role within the whole – and obviously the emotional journey and the emotional arc.

Many large-budget animation and VFX companies divide acting shots across animators without necessarily providing any one animator with the full context of a character in a film. As a result, animators in this situation cannot comprehend the character with the same fullness that an actor would, and the animators must rely on their supervisors to convey an understanding of the totality of the character's performance instead. Serkis believes that an animator working in isolation lacks what he calls the "chemical exchange" between two or more actors who play off each other in a scene:

> The chemical exchange […] between actors is generally what informs the freshness of that moment, and finding in each different take and series of takes – and if you're working for a director who's looking for the heart of a scene, who is looking for different ways of interpreting the moment – the arrival at those moments that feel real and engaging are because of that.

In Serkis' estimation, an animator is expected to choreograph a scene, but when it comes to

PeCap, the contribution of the actor rules over a staged moment.

PeCap introduces a degree of intermediation between an actor's original performance and the final performance that is applied to a digital character. That final performance retains an undeniable indexicality to the actor's original performance, but does the degree of intermediation in PeCap change our understanding of what is acting in any way? Serkis doesn't believe it ought to because, historically, an actor's authorship has been retained despite numerous postproduction augmentations that could alter the actor's original intention in order to satisfy the whims of a director.

> As screen actors, we don't control, ultimately, the emotional arc of the character because of the edit, because of the choice of music, because of the choice of shot, because of the choice of so many different things. So, we are augmented. Every single actor's performance is augmented from a minor way – from an edit point, to having a blink removed, or a tear, and so on and so forth. So, unless you're talking about stage acting where you purely see the actor on stage, then I would say all forms of screen acting are augmented.

Although there is a precedent in film to augment actors' performances through a variety of direct and indirect editing techniques, in PeCap animators have at least the potential to make dramatic expressive changes to a character, or to substitute part of the captured performance with some keyframe-animated alteration. However, Serkis indicates that at least with the productions he's been a part of, the animators must fully take their cues from the actor's performance and preserve its integrity:

> I don't believe that the animator actually changes anything to do with the authorship of the performance emotionally. What the actor gives off emotionally cuts through and touches the audience. […] Then again, you have directors who might want to, for a specific project, have that performance enhanced in some way by the animator. If it's a slightly more cartoon-like, slightly more heightened or cartoon physical rendering of the character.

Here Serkis points out that he believes that any kind of animator-driven enhancement or alteration occurs solely at the technical level: "I don't believe that the actual emotional intensity or authorship is altered at all." For Serkis, the role of an animator is to preserve and clarify the essence of his acting through a digital character with a different physiology. He provides an example from his experience of working with director Matt Reeves on Dawn on the Planet of the Apes:

> [Reeves] directs the scenes with the actors. We are shot, we are filmed. He then takes the performances and cuts them. And, if there's anything that varies from absolutely, quintessentially what he is given by the performance of the action on the day – and then he lives with that cut by the way, and screens that cut to many, many people [sic]. He creates the arc of the story, he creates the journey, understands the dramatic principles of the film, of the movie, from watching those performances. Nine months later on, the visual effects shots come in and the actors are gradually replaced by those shots. But he's already made the film, he's already told the story with those performances. So, he is looking to basically have them clothed, or make-upped if you like. Hence this term, which really angers animators: "digital makeup."

This is a direct reference to the controversies in the online animation community provoked by Serkis' use of the term "digital makeup". Animators who take issue with Serkis believe that he assumes undue total credit for the finished performances of the digital incarnations of the characters he's played, and actively diminishes the role of animators to mere button pushers whose contributions are technical rather than creative. He distances himself from this controversy at the first opportunity: "And [this term "digital makeup"] certainly, by the way, did not come from me." He wants to be understood that he made use of this term because it was the language Weta Digital provided him to refer to the process of preserving and translating performance from human actors onto their digital counterparts.

Due to the process of translating performance from a human actor to a digital character, as well as the intermediation of the animators in interpreting the performance, it is possible that in any given moment a performance may be preserved or altered. How does Serkis feel about watching digital counterpart performances, and especially his own performances translated through digital counterparts? Does he ever feel anything is missing from the final performance that was present in the original acting?

> Sometimes a performance can be slightly misread and the intentions slightly misread. But I have to say that Weta have established a pipeline over the years which is increasingly being able to understand and interpolate the performance to such a high level now – and "Dawn [of the Planet of the Apes]" is a great example of that. With every iteration of technology, with the pipeline, with the workflow and all the people that – and not to say, it's not just about technology, it's about people. It's about people having an understanding what is required, and not seeking to augment

> unnecessarily or to add their flourish. I think if you're an animator working with performance capture, you have to for these types of films – you've got to want to honour the performance and that's what it comes down to. Because that is why all of the trouble has been gone to in the first place: of shooting on set with the director who wants to have that relationship with actors. They want to have that sense of the drama they create in the moment on the day, on the film set, honoured.

On the subject of whether animators should be considered actors, Serkis provides greater nuance than what is often reflected in other interviews. He believes that some keyframe animators "absolutely" are actors, but that "some of them are definitely not." He distinguishes between these extremes based on the degree to which an animator possesses a full understanding of a character:

> [As an actor] you're finding… you're using emotional recall or you're using parts of yourself – using a lot of parts of yourself – I'm putting that under the microscope. And I'm sure animators do that, too. But I think, again, I go back to this notion of – if you asked an animator to play a role over the course of the length of the play of an evening, they might find that challenging. Whereas, what they are brilliant at doing is taking a moment and making it crystalline. Sometimes the good ones, the good [animator-actors], make that moment crystalline but also subtle and emotionally engaging, and the ones that are not-so-good make it look flashy.

Serkis asserts that while the physical blocking may be the same between different animators, he believes he can always tell when he is watching the work of an animator who values subtlety and sincerity versus an animator who wants to show off his or her skills without necessarily respecting the source-performance.

It is important to point out that Serkis cites these examples as stemming only from his own experience, and he is careful to indicate that directors and animators working on other types of films may have different approaches to this same process. That is, he recognises that his experience of acting in performance capture may not be universal. This provides a meaningful counterpoint to the fact that his comments are often interpreted as pertaining to MoCap as a whole due to his prominence in the field.

6.6.2 Interview with Dan Lemmon

I conducted the following interview with Weta Digital VFX supervisor Dan Lemmon on 31 January 2018 via Skype video chat. Lemmon spoke with me from his office at Weta Digital in Wellington, New Zealand, while I was at my office at Auckland University of Technology in Auckland, New Zealand. For context, Lemmon was the VFX supervisor for all three films in the recent "Planet of the Apes" trilogy, each of which he was nominated (along with other supervisors, variously including Daniel Barrett, Joe Letteri, R. Christopher White, Erik Winquist, and Joel Whist) for an Academy Award for Best Visual Effects. He won his first Academy Award for Best Visual Effects for "The Jungle Book" at the 89th Academy Awards in 2017. In each of the "Planet of the Apes" films, Lemmon worked directly with Serkis, and he therefore possesses valuable insights about both the acting and technological considerations that underpin the final PeCap screen performances in those films.
What follows is my summary of the interview, and the full transcript is available in the appendices (Section 9.8). All quotes below are from this interview (D. Lemmon, personal communication, January 31, 2018) unless otherwise cited.

A central part of this interview focuses on Lemmon's insights regarding how acting is produced via performance capture on the films he has worked on, and specifically how this affects the notion of authorship within a PeCap performance. He notes that in response to the "Planet of the Apes" movies, the critics often indicate how they believe the performance of the ape characters more than the human characters. He attributes this to how well the animators finessed the acting performances of the ape characters to make them more life-like and believable:

> [n]ot necessarily because of its appearance, but emotionally the performance seems more authentic than the performance that the human actor is putting in, and I think this is an astute observation. I think it comes down to the quality of a number of things, but one is specifically the quality of the actors that are playing different roles. How many movies have we seen where somebody is fantastic in a scene and then you cut away to the actor who is maybe less talented or a little bit less engaged, and the whole scene falls over? And this is something that the performance capture process is really good at, is really useful in giving filmmakers and actors the tools with which to craft scenes […]. When you've got a grip waving a stick around with a tennis ball

> at the end of it – and, you know, saying that this is a scary monster, react to it – you get a totally different kind of performance from the human than if you actually had something or somebody in there playing that character. […] [I]t's not just about making the apes look better, it's also about giving the humans something to react to and to play off of as well.

Similar to Serkis' discussion about the director's process of editing the PeCap takes into a version of the film that tells the whole story without the source-actors replaced by their digital counterparts, Lemmon also elaborates on this process for both "Dawn of the Planet of the Apes" and "War for the Planet of the Apes". Working together from the production dailies, Reeves and his editors William Hoy and Stan Salfas created storytelling decisions based on what the PeCap performers were doing on set.

> And so, [Reeves is] making specific choices from cut to cut – when to cut, all the pacing, all the mood, all of the tone is based on what Andy did. And so, there's a version of the movie that goes together where there's no apes at all: it's just the dailies with Andy and Kieran and Terry – all the ape actors wearing their funny grey suit with the little dots on them and the scenes play out with them in those funny grey pyjamas.

Lemmon notes that as soon as the story starts to unfold, it is a natural human instinct to invoke a suspension of disbelief and the audience forgets they are watching actors "wearing funny grey pyjamas [… and] just accept that that's an ape now, that's the big gorilla." All of the cues to invoking this belief are embedded in what an actor does in terms of his/her face, voice, and posture. In order to make these performances as believable as possible, Terry Notary, who plays the apes Rocket and Bright Eyes, explains that the actors learned "the principles of ape locomotion – allowing the front limb to absorb one's weight and squaring the shoulders before moving the back limb – and the ways apes focus intently on one object at a time." (Bode, 2015, p. 103). Lemmon describes how Reeves spends many takes trying to get the believability of the ape performances just right, and that ultimately both his and Reeves' approaches come down to asking "do I believe it?" If yes, they move onto the next scene, but if not, then they pick apart what inhibits the believability of a shot – and specifically whether the character's performance seems emotionally consistent, motivated, and sincere.

Lemmon elaborates on a scene in "Dawn of the Planet of the Apes" in which the character

Koba, played by Toby Kebbell, encounters two human soldiers drinking whiskey in an ammunition depot. Kebbell, who grew up in the UK and spent time in pubs in London, mused if he could play the character like an English tough who

> act[s] like they're your best friend, but as soon as you drop your guard they come around and stick a knife in you. And he recounted a story where something very much like that happened to a friend of his. That's where he was bringing his own background, his own story into that scene. That scene played out, when we were watching it on the set on the day just from the video monitor, it was chilling – it was amazing how well [sic] it was and how scary that character became because you trusted him for a minute. You're laughing along with him, he's clowning around, and all of a sudden, he murders these two guys.

Lemmon sees this type of improvisational actor decision-making as one of the key justifications for engaging with performance capture. With PeCap, actors are able to interact with each other and their environments in a way similar to shooting a standard live-action film, while capturing the details of a performance to map onto characters who look nothing like them, and who will ultimately replace the source-actors in the scene.

It is only here that Lemmon comes full-circle to address my initial question as it relates to the authorship of PeCap performances. He compares the source-actor's performance to the "DNA of what's going into [a] character emotionally". This is then passed on to artists and technicians who "go to a tremendous amount of work to service that performance, to stay as true as possible to those decisions – to that intentionality – to keep everything firing in the same way it was fired before." He explains that since the edits are all based on the source-actor's original performances,

> [t]he cut doesn't work anymore if we start mudding it up if we start making different decisions about taking the face to a different place emotionally – changing the blinks. The cut functions on very delicate, specific things that happen on the actor's face, on the actor's body – and if you alter those things even in a small amount, it won't have the same… Well, in the best case, the emotional impact is diluted and in the worst case, the whole cut and the scene fall over because it just doesn't work anymore.

Lemmon notes that there were some ape characters that were keyframe animated rather than performance-captured, but these were primarily background characters and actions that human actors could not physically produce. He relates the story of how the production needed to record brachiation – a type of movement that an ape or monkey performs when it swings from one arm to the other. He explains that when apes perform this movement and release from one arm, they become airborne for a moment before catching a branch with the other arm. Importantly, they continue their body movement forward without double-pumping their arms after each catch.

> We had guys that could release, and then they'd pump. Basically, when they caught with the next hand they'd have to pump before they could go again. And this guy, Jason Chu, he could do it [without pumping], and he did it three times and then he tore his chest muscle. He was no good to us for the rest of the shoot – it was a real shame. But even those three times, we used them over and over again as good solid motion capture data we could build cycles from. […] [T]here are some things just physically that humans can't do, and in those cases the animators sort of take-over, either 90 percent or 100 percent.

One area in which the animators had a significant impact was interpreting how the source-actors' performances would be meaningfully represented through the non-human physiologies of the ape characters. Lemmon notes that the apes' legs are much shorter than human legs, which meant that the way an ape could ride a horse was significantly different from a human. Likewise, ape arms are long, stronger, and often more massive than human arms. Maurice, the orangutan, is especially massive, and animators often needed to dampen her movement, although Lemmon notes that actress Karin Konoval did a good job of moving as if she were much larger. In order to assist with the facial animation, distinct features from a source-actor's face were modelled into his or her digital counterpart, such as the shape of Serkis' eyebrows and folds of his upper eyelids built into Caesar's face. However, the overall shape of the ape muzzle is so different from the human mouth region that it is not possible to create a one-to-one mapping. Lemmon describes how "fairly substantial sections of the facial rig" were rebuilt in order to accommodate Serkis' emotional performance in a given scene: "where he took the kind of level of conflict and extremity with what he was doing with his face was so different from what he'd ever done before, so we ended up rebuilding substantial portions of it." Even the nasolabial fold region (the crease that goes from the corner of the nose to the corner of the mouth) was pushed further on Caesar than a normal ape in order to better transmit Serkis' sneering. This process takes the strong acting template provided by the source-actor and passes it through the animator's eye to ensure it reads properly on the ape character. Despite any of these artistic contributions from animators, Lemmon still sees the authorship of the character as belonging to the source-actor. He compares this to how a muralist may have a concept for a mural, with planning done through sketches and a template created for what the finished work should look like. While the muralist may hire dozens of painters to actually execute the painting, the muralist remains the author of the work. Similarly, so too does the source-actor – so long as the animators can be seen as merely executing the creative template and not adding creative choices of their own.

In a Cinefex article about "War for the Planet of the Apes", Lemmon discusses the challenges imposed when animating a character whose physiology does not closely match a source-actor. He mentions that when trying to animate the likeness of an actor, "it's more about the change than it is the rest pose. If you can get the change to read, the character's performance tends to feel like the actor" (Fordham, 2017b, p. 19). That is, despite the great care taken with creating meaningful blend shapes, the real key to achieving a believable performance is in how those shapes and motions transitions between each other. I ask Lemmon to further elaborate on this idea and he explains that after Paul Walker died during the production of "Furious 7", Weta used Paul's brother Caleb, who is a decent look-alike, to test lighting and performance.

> The funny thing that happened was we would take a scene where maybe Caleb was sitting on the driver's seat, driving and looking out the window, reacting to something – and we would replace Caleb's face with Paul's face, doing the same thing that Caleb was doing. And what was really interesting was that it didn't look like Paul anymore. Like, a still frame still looked like Paul, but when you played it back, it looked like Caleb. Because they looked close enough to one another that suddenly it wasn't about the actual still appearance of their face – it was the performance that was driving it, and the performance itself bore the signature of Caleb, not Paul.

Lemmon's comments suggest that within PeCap, the indexical quality of an actor's performance is found more in the nature of the movements between key facial poses rather than in the poses themselves. When Lemmon and his team examined Paul's old performances, they noted a number of qualities specific to Paul's performance that varied from Caleb, even though they looked similar. These include certain hesitancies, moves, ticks, and "star swagger that Paul had developed over years of being a movie star and inhabiting that skin – that role was totally different from Caleb". He elaborates further:

> it's easy to point at a photograph of a person and say "that's the person", and when you look at a video of a person, often you think of the collection of images as being the person. But, in fact, it's the movement between the images – that going from one spot to the next spot and the timing of those movements – that actually is the performance. And the performance is as much the signature of the actor as the facsimile of the actor – the static visual appearance of the actor. And what we're doing with performance capture – we're completely erasing the visual facsimile and preserving the performance.

Returning to a quote used in Section 2.4, the notion that a character's performance occurs between key facial poses is similar to animator Norman McLaren's famous declaration that

> [a]nimation is not the art of drawings that move, but the art of movements that are drawn; what happens between each frame is much more important than what exists on each frame; animation is therefore the art of manipulating the invisible interstices that lie between the frames. (Thain, 2016, p. 168).

Here we find parallels between animation and performance capture which, through their dissection of performance at an individual frame level, suggest that acting is recognised more in the transitions between expressions than in the individual expressions themselves. This appears to be supported by research into the identification of people by analysis of their individual gait patterns (Sarkar et al., 2005) and body actions (Loula, Prasad, Harber, & Shiffrar, 2005). There is plenty of room for future research that tests the validity of this conjecture.

6.7 Conclusion

This chapter explored the various contributions of source-actors, PeCap technicians, and PeCap animators to the finished screen performance of digital counterparts. While most academic literature on this subject is written from the perspective of outsiders (e.g. people who are not personally familiar with the practice of acting and animating for PeCap), the research presented here provides an insider view from the perspectives of both actors and animators in the creation of PeCap performances. This variety of perspectives helps to more accurately articulate the strengths and limitations of PeCap workflows, and provides insights into making sense of the controversies between actors and animators surrounding authorship in performance capture.

If we are to read between the lines of the Hollywood rhetoric about performance capture, we may come to the same conclusion as Variety senior editor David Cohen:

> [w]ith performance capture, specifically, the involvement of animators is a bug, not a feature. It's a necessary bug, because apes and N'avi and the like aren't a neat match for the actors who play them, but the better and more precise the process gets, the more the actors will translate to the screen. Progress in the field is *defined* by steps that shift control from the animator to the actor. (2014, para. 15, *original emphasis*).

The question is whether the process will ever be able to create such a perfect match between extreme physiologies. This is a debatable point, and it is conceivable that animators will always need to be on hand to make aesthetic choices for digital nonhuman characters. No matter how good a source performance is, if the PeCap technology is unable to transfer all of its nuances, the skills of animators will be needed to achieve the final effect – especially in the case of nonhuman characters. What seems clearer is that as the technologies improve in capture fidelity, animators are less called upon to make creative performance choices. Some people regard cinema as not an actor's medium – at least in the sense that an actor does not have control over how the final product of his or her acting is constructed on screen (Cohen, 2014, para. 27). It may be similarly argued that performance capture is not an animator's medium in the sense that animators have far less control over creating animated character performances, and instead work from authoritative templates performed by source-actors. Pizzo sees the digital counterpart as "a text with many authors, whose jobs may be interdependent" (2016, p. 50). Likewise, he suggests that

> the motion capture actor does not provide a performance per se, but a text to be played and interpreted by the animators. The data the actors provide can be used in a different setting from the one imagined at the moment of shooting, the pace of the movement can be changed, the length can be increased. But there is one thing that will be retained: the emotional mood of the action. Just as, although a score or a

> script may be read or performed in different way, unless the data is radically changed, it will hold its original sense. (2016, p. 59).

Based on the interviews with Serkis and Lemmon, it is unlikely that PeCap data would be wielded so cavalierly, at least in films where the final edit of the film is heavily determined by the timing and pacing of the source-actors' performances.[80] Bestor cautions against taking the promotional rhetoric of directors (e.g. Cameron and Zemeckis) and source-actors (e.g. Serkis) at face value, as he believes that such individuals "can be a poor source of information on the state of the technology", and that the perspectives of digital artists can often prove more insightful (2016, p. 171). Assuming that Bestor is correct in this summation, the practice-led research I conduct in this PhD helps to fill a knowledge gap specific to the below-the-line labour in performance capture.

---

[80] Coincidentally, Pizzo compares MoCap process to the interpretation of a score or script, which is consistent with Raitt's metaphor of a MoCap source-actor as the composer of a performance, the animation director as the conductor, and the animators as the musicians playing the piece (Singer, 2003, para. 3).

7

CONCLUSION: A NEW PARADIGM OF ACTING

7.1 The Nature of Acting in Performance Capture

Performance capture has long been the pariah of both animators and actors who see it as a threat to their respective crafts and livelihoods, and which only offer impoverished results in return. The popular rhetoric around PeCap by film producers such as Jon Landau often directs the discussion toward the preservation of the actor through the technology. For instance, while speaking about his experience of producing "Avatar", Landau indicates that "[o]ur goal in using performance capture was not to replace the actor with our computer animated character, but to *preserve* the actor – because what a great actor does and what a great animator does are antithetical to each other." (Duncan, 2010, pp. 70-72, *original emphasis*). He suggests that great actors withhold information while capturing the audience's attention through presence alone, whereas animators would never "animate a character to just sit there and do nothing."[81] (2010, p. 72). In a separate interview, Landau indicated the "Avatar" producers' decision to affirm the technological approach as *performance* capture rather than *motion* capture:

> We wanted the creative choices of performance to be made by actors and not made by animators, and that takes nothing away from great animators, because the animators are a very important step in the process, but the decisions that are made of when to stare and not blink and when to twitch – those are made by actors because that's what they do. (Menache, 2011, p. 73).

According to Landau, the role of the animators is ensuring that the source-actors' performances are fully transmitted through their digital counterparts, although he provides

---

[81] It is highly debatable whether an actor ever truly "does nothing" even when he or she appears still, nor is it clear why an animator would not be able to animate stillness – although, as I highlighted in Section 6.4, the 3D animator would need to provide some minor sense of movement in order to overcome the deadening effect specific to 3D characters.

no details about how this is accomplished. He is quick to point out that PeCap was not used to save on animation or money, but rather on *performance*. That is, at least in a full PeCap production such as “Avatar”, the process more immediately captures an actor’s performance without worrying about lights, cameras, cuts, or many other attendant aspects of standard film production. In this way, the actors’ performances are not interrupted and the actors can more naturally play off of each other (2011, p. 73). Actor John Malkovich speaks of his experience performing in “Beowulf” in similar terms:

> Well, to me, it was remarkably reminiscent of doing plays […] [A] lot of the things that might have come into play in normal filmmaking, don’t come into play [in performance capture]. You don’t wait for lights, you don’t wait for camera, and you don’t wait really for anything. Continuity doesn’t matter too much. And, you just act all day. […] That, for most of us, is quite liberating because, at times, the process of making a regular film has remained quite medieval, in some ways, especially with the amount of time it can take. (Roberts, 2007, para. 55).

It can be argued that performance capture has the strongest relationship to theatre out of all approaches to filmmaking, especially when used within a fully virtual environment. In the case of “Beowulf”, where an ensemble cast is assembled on a MoCap stage to create continuous takes without regard to cameras or lighting, the actors are able to encounter each other in a similar manner to performing on stage. Rozik indicates that a distinctive feature of theatre is the audience’s immediate experience of a human actor on stage (2002, p. 110). Likewise, stage acting more reliably provides actors with an immediate experience of each other – whereas in filmmaking, actors in the same scene (or sometimes even in the same shot) are not always present together at the time of filming.

How does performance capture affect the nature of stardom? King notes that Hollywood stars are dependent on “the key supports of marketing, promotion, and publicity to create a personal monopoly of attention”, which are used to justify increased compensation due to recognisability (2011, p. 249). The danger of becoming “hidden” within a PeCap role runs counter to recognition, and a famous actor may understandably fear that such a role would bring neither the financial reward nor publicity needed to maintain star status. Serkis serves as a telling example. He has played iconic or leading roles in numerous blockbuster films (i.e. Gollum, Kong, Captain Haddock, Caesar, Snoke), but due to the performance capture nature these roles, he remains unrecognisable to many filmgoers. People who do not recognise Serkis by name are often familiar with the roles he has played, but would not be able to pick his face out of a line-up. In fact, in my experience, people are most likely to remember his face from his role as Ulysses Klaue in "Avengers: Age of Ultron" and "Black Panther" (Coogler, 2018). As mentioned in Section 4.1, Lupita Nyong'o was encouraged by the prospect of disappearing into the digital incarnation of the alien heroine Maz Kanata in "The Force Awakens", but she is not betting the rest of her career on invisible PeCap roles. No other actor has made such a prominent performance career out of PeCap as Serkis. However, with PeCap acting training becoming more available, this may soon change. And as Serkis advises, PeCap provides opportunities for actors who may otherwise not be traditionally castable, due to either looks or disabilities (Sharf, 2019, para. 1). The question is, how possible would it be for an unattractive or disabled actor to gain star status through acting alone if they are never "seen" in the traditional film sense? There are examples of stars who may not be considered conventionally attractive (Steve Buscemi, John C. Reilly, Glenn Close, Toni Colette)[82] but they manage to attract audiences through their charisma, presence, humour, or other qualities – or sometimes simply through fortunate casting in an unexpectedly popular role.

Goldstein and Levy argue that when it comes to awarding acting performances, critics are not interested in creative performances, but rather in providing a recognisable representation of reality. In fact, it is unclear precisely what the notion of a "creative actor" would entail (2017, p. 154). For the sake of argument, I describe a creative actor as one who has a keen ability to clearly imagine fictional environments, circumstances, characters, and physiologies, and who can act to these imagined elements while performing in an austere space.[83] Based on this definition, it is reasonable to suggest that good motion

---

[82] Attractiveness is an admittedly subjective criterion and someone may legitimately take issue with the examples I've provided here. What does seem inarguable is that there are more male than female stars who are not conventionally attractive. If stardom is accessible to performance capture actors, it may help to level the gender bias between attractiveness and stardom.

[83] Goldstein and Levy also suggest that intuition may play a role in an actor's creativeness (2017, p. 154).

capture actors are creative, as MoCap requires the ability to imagine and retain fictional circumstances while performing – most often in excess of the fictional circumstances a stage or film actor encounters.

Within the Stanislavskian approach to acting, the actor functions as "an autonomous artist who freely collaborates with the director in analysing and creating the character on stage." (Balcerzak, 2013, p. 202). However, in filmmaking, it is the director, not the actor, who fulfils the role of *auteur*, which fits a more Strasbergian view of acting. Strasberg believed that the director had the final authority over a character, and "stat[ed] that he considers the 'problem of acting solved' when the actor 'is capable of giving to the directory anything that he wants'." (Carnicke, as cited in Balcerzak, 2013, p. 202). While the film actor is subject to editing to produce a performance that was never given, PeCap functions as an even more aggressive form of Strasbergian control: every aspect of the source-actor's performance is subject to revision by the director (2013, p. 202). Visual effects artists privately admit that changing actors' expressions is common practice:

> opening or closing eyes; making a limp more convincing; removing breathing signs; eradicating blinking eyelids from a lingering gaze; or splicing together different takes of an unsuccessful love scene to produce one in which both parties look like they are enjoying themselves. (Hoyle, 2017, para. 7).

One VFX artist admits that making such changes to an actor's performance places him in a difficult moral position, but that ultimately visual effects artists are hired by directors, not actors, and it is the directors who they are obliged to satisfy (2017, para. 8). Speaking about the character Clu in "Tron: Legacy", who is a digital recreation of Jeff Bridges from 28 years earlier, director Joseph Kosinski notes that "[w]hen you age, not only does your look change, but your mannerisms change, as well." (Duncan, 2011, p. 41). The VFX team encountered numerous difficulties when transferring Bridges' performance onto his younger digital counterpart, and it was often necessary to adjust his expressions to better match his younger performance as a 35-year-old actor in the original "Tron" (Lisberger, 1982).

One unresolved question is whether animators are actors, or at least whether they are actors within specific contexts. Blair's neurocognitive understanding of "character" as not an identity but rather a strategy of performance helps guide us toward an answer. Within the Alba acting model, "character becomes a set of choices and behaviors – a process, rather than a discrete entity, a motivated movement, rather than a gloss of feeling –, supported by what the actor brings to the role." (2008, p. 82). As a result, an actor engaging with a character is not necessarily that different from an animator engaging with a character. The actor engages with a particular process of choices and behaviours as strategies to produce a sense of a character, but not in the sense of trying to become that character. All of the actor's movements are motivated to convey a sense of who the character is, but this is not necessarily at the whim of any particular feeling. In fact, this is similar to how animators have worked to produce characters all along. Blair indicates that Stanislavski's method of active analysis "create[s] an embodied, conscious, coherently articulated being [in which the actor] manipulates principles of action, behavior, imagination, attention, emotion, and memory to […] engage the work intellectually, physically, and emotionally." (2002, p. 181). Although Stanislavski likely saw the conscious and unconscious as discrete entities rather than intertwined neurophysiological processes, Blair believes it is possible to see in his work an embodied consciousness that can escape a Cartesian dualist conception of the self/actor (2002, p. 181). The core point of difference between actors and animators in this regard is that an actor utilizes both his/her own mind and body, whereas an animator is given only a separate body to work from. It is only through moving the animated body that an animator can create the illusion of a thinking and feeling character.

7.2 Detailed Significance of Research

This research drew on a wide range of disciplines and sources to provide a comprehensive overview of how we understand the practice of acting within performance capture. These disciplines include acting, animation, computer science, film studies, celebrity studies, performance studies, popular culture criticism, philosophy, psychology, sociology, and kinesiology. This thesis linked theory with practice and functioned as a research nexus between these various fields in order to address several critical questions within the current state of performance capture research:

- What does acting become when the product of acting starts as data and finishes as computer-generated images that preserve, to varying degrees, the source-actor's "original" performance?
- What is the nature of acting within the contexts of animation and performance capture?
- What is the potential for a knowledge of acting to have on the practice of animating, and a knowledge of animation to have on the practice of acting?
- What is the role of the animator in interpreting an actor's performance data and how does this affect our understanding of the authorship of a given performance?

To move beyond a purely theoretical understanding of performance capture, I explored these questions through practical experiments in which I inhabited the combined role of an actor-animator. A digital counterpart involves a highly mediated form of acting that often blurs the distinction between the roles of actors and animators in the authoring of a final performance. By fulfilling all of the roles normally executed by specialists in the production of performance capture, I was able to create a one-to-one correspondence between a single animator and a single 3D character's performance, while serving as the sole creative actor.

This research provides a unique contribution to the study of performance capture in that it establishes a unified language of experience across the various creative roles responsible for the final screen performance of a digital counterpart. This is important because academic animation literature provides a meaningful contextualisation of ideas, but is less well-informed by the pragmatic and creative experiences from animators themselves. This research helps to fill this gap by contributing my process as both an actor and an animator to the literature. In turn, this research articulates the strengths and weaknesses of various PeCap workflows and provides a voice to both actors, animators, and below-the-line film workers in terms of making sense of authorship in performance capture. My perspective as an independent practitioner also provides an alternative narrative to the promotional rhetoric of popular film directors and actors, whose opinions some may see as tarnished by self-serving propaganda.

From a practice-informed theoretical perspective, I traced a history of acting styles that influence modern approaches to acting in performance capture. I also revealed how animators rarely encounter any professional acting training and instead often succumb to producing either emotionally underplayed or overplayed reference performances. From this, I suggested that the Alba Technique may a) provide a more accessible (and ethically responsible) means for animators to produce sincere reference performances, and b) form a basis for understanding how and when animators function as actors.

I demonstrated how a knowledge of psychology and physiology of human expressions affects how an animator understands and interprets data when working from realistic performances. This has a variety of implications from distinguishing between genuine micro expressions and signal noise in the performance data, to understanding how facial chirality affects an audience's empathic response to characters.

I provided a comprehensive explanation and overview of my technological methods when working with PeCap software, including making practice-informed modifications to the modelling, texturing, and rigging of a realistic 3D human character. I provided solutions to specific problems encountered during the animation process, including my approach to managing data noise, inaccurate facial retargeting, and manual pose adjustments. I also discussed situations in which PeCap animation may need to be modified or fully replaced by keyframe animation and how to successfully match this to the existing PeCap data.

This research clarified modern production strategies for PeCap and how this both aligns with and challenges the critiques lodged against pro-actor PeCap proponents such as Andy Serkis, James Cameron, and Peter Jackson. Additionally, I described how MoCap animators' artistic judgment is deployed in the interpretation of dissimilar physiologies between source-actors and digital characters, as well as in resolving problematic performance data – including the ability to perceive, interpret, and translate both large-scale and small-scale performance details. These concerns help us better address how we understand authorship in the context of digital counterpart performances created using motion capture, as well as in the context of digital posthumous performances. I clarified how critiques of PeCap authorship need to be contextualised within the nature of a particular VFX company's production workflow; some companies privilege the actor's input whereas others privilege the animator's artistic interpretation. Likewise, I indicated how such critiques must also account for differences in production strategies for realistic versus stylised approaches to performance capture. This research made a strong case for how, within a PeCap context, our understanding of acting needs to be rethought to include the contributions of a collective rather than just an individual.

I presented and analysed in-depth interviews with Andy Serkis and Dan Lemmon that examine performance-specific complexities within acting and animating for performance capture. This provides insight into the thinking of leading PeCap practitioners about the nature of authorship of digital characters and the methods used by actors and animators to create them. My interview with Serkis also provides greater nuance to his controversial understanding of the authorship of digital characters, highlighting that the nature of production and the desires of a director greatly influence his opinion. He also acknowledges that he considers some animators to be actors based on how fully they understand the characters they animate.

7.2.1 Mapping Specific Details

Within the larger contexts of the research outlined above, numerous additional questions were pursued throughout this research that resulted in significant insights. These include:

- In what ways do actors, PeCap artists, and PeCap animators contribute to the final screen acting of performance-captured characters?

- How are we to understand the nature of acting as an embodied process when performance capture separates an actor's performance from his or her body?
- Does performance capture serve as a type of digital prosthesis (or "digital makeup") through which the actor performs autonomously?
- What is the relationship between the flesh-and-blood performer and the digital character that performs on screen?
- Should acting in performance capture be seen as equivalent to standard live-action and therefore be eligible for Academy Award consideration for Best Actor and Best Supporting Actor categories?

Significant outcomes resulting from these questions include robust definitions for performance terms that historically lack clear and distinct meanings among actors, animators, and directors, including: *authenticity*, *sincerity*, *truthfulness*, *honesty*, *aura* (Section 4.1), *emotion*, *feeling*, *behaviour*, *expression*, and *empathy* (Section 4.2). Additionally, I provide a proposal for standardised definitions of virtual performance terms that are often mistakenly used interchangeably, including: *vactors*, *synthespians*, and *digital doubles*. I provide a taxonomy of these terms, including sub-classifications such as *protean doubles*, *digital doppelgängers*, and *digital fantoccini*. A core innovation in the definitions of these terms is that they are based on performance rather than visual style (subsection 4.4.1).

Additional meaningful outcomes include:

- An overview of the ontology of digital counterparts that combine Deleuze's concept of the *geste*, Bazin's "change mummified", Balcerzak's "ghosting of the actor", and Craig's über-marionette (subsection 2.7.3).
- Adapting Bode's concept of a screen performance continuum to chart a variety of technologically mediated performance types and examples of associated film characters (Section 3.3).
- An explanation of how the concept of empathy can be extended to our experience of identifying with the emotions of animated characters, as well as a description for

how animators engage with empathic attunement when kinaesthetically performing their characters' actions through their bodies (subsection 4.2.3).

- A reclassification of vactors/synthespians that is exclusively limited to profilmic contexts, and the creation of a new term, *ex-filmic*, to describe cinema that is solely recorded through software-based cameras (subsection 4.4.1).
- A detailed analysis and taxonomy of vactor performance styles (subsection 4.4.2).
- A theorisation of an actor's experience while performing via motion capture in terms of three poles that circumscribe his or her bodily perception: Leib, Körper, and Doppelgänger. These poles are represented graphically and describe a post-Merleau-Ponty phenomenological understanding of the embodied performance experience in terms of auto-empathy (body-as-experiencer), physicality (body-as-experienced), and the virtual (body-as-simulation) (Section 5.2).
- An account of issues faced by actors when encountering a motion capture environment for the first time, as well as strategies for managing these issues (Section 6.1).
- Clearly identifying the roles and responsibilities of source-actors (Section 6.1), MoCap technicians, MoCap artists (Section 6.3), and MoCap animators (Section 6.4).
- Articulating the animator's role within performance capture, especially in terms of how the animator completes actions that are not part of a captured performance, as well as creating seamless transitions between custom-animated performances and source performances (subsection 6.2.3.3).
- A process for quantifying how much the original PeCap data drives the finished performance versus how much is augmented by manual keyframing (subsection 6.3.1).
- Ethical considerations surrounding the modification of an actor's "original" performance data in order to modify acting versus technical issues, such as continuity (Section 6.5).
- Unique performance-specific insights from interviews I conducted with two highly regarded industrial professionals (Serkis and Lemmon) (Section 6.6).

- An account for a meaningful description of "digital makeup", including in which situations it is appropriately used (Section 6.6).

7.2.2 Utility of This Research

This research is intended to be useful to researchers in a variety of disciplines, each of whom may find benefit from distinct insights raised here. This thesis provides motion capture artists with greater visibility for their craft and their roles in producing digital character performances. Additionally, motion capture artists may gain a clearer insight into an actor's process for producing the source performance, which may conceivably affect how a MoCap artist understands and modifies the performance data.

Actors who feel threatened that their craft will be replaced by cyberstars will find some solace in the fact that this research suggests that such digital replacements are likely still a long way off, and may remain prohibitively expensive even when they finally emerge. Alternatively, actors may find additional opportunities via performance capture to play roles that sit well outside their standard "types", including characters that are physiologically different from the source actor. A core message that emerges from this research is the need for actors to begin to see their work not as the product of an individual, but rather as a composite of multiple performers who may each contribute vital elements to a single role.

I offer to animation theorists an ontology of digital performers that continues to question whether indexicality is an appropriate measure by which to understand digital performances. With an eye toward animation, I provide new definitions for several acting terms that are frequently used by producers, directors, and actors, but which often lack specificity. Furthermore, I describe an academically rigorous understanding of what is acting as it relates to animated characters, especially motion-captured characters. This extends beyond Hooks' "Acting for Animators" (2011), which serves as the current industry-standard manual for applying acting principles to animation. Animators may also appreciate specific acting strategies for achieving better reference performances. This research provides animators with a clearer understanding of how performance capture is produced and the role that animation plays within the creation of the finished performance.

7.3 Room for Future Research

This research explores numerous theoretical and practical considerations surrounding acting in performance capture. However, my approach to the practical components of this research is not representative of the standard team model in the VFX industry. Instead of specialising on one aspect of the creation of the final performance, I performed all of the roles from acting to 3D modelling, texturing, rigging, PeCap clean-up, and PeCap animation. While this provides me with a clear account of the performance decisions that underpin the entire process, it means that my progress was both slow and affected by my own limitations in any of these areas. That is, while I may be a capable 3D generalist, I lack some of the deep insights and techniques of someone who specialises in a particular area of 3D production. Now that I have made an account of how performance is constructed via PeCap, it would be meaningful to continue this research by working with artists who are experts in these various specialities. Such future research could focus on how the knowledge of acting in PeCap articulated in this research affects a team-based approach to producing PeCap performances. Additionally, there are elements of advanced 3D production workflows that I was unable to engage with in this research due to the limitations of my own knowledge. It would be meaningful to consider how facial muscle simulations affect the perceived sincerity of expressions in a 3D human – for instance, by directly comparing PeCap animation on identical character models, with one that includes muscle simulation and one that does not. Finally, it would be useful to directly compare how different PeCap systems record and retarget facial performance data, and to devise a system of best practices for PeCap actors and artists to consider when engaging with one type of system versus another.

There are numerous avenues to research how we understand animators as actors. For instance, it is worth investigating whether emotional sincerity in character animation can be achieved from a solely physical approach to generating reference performances – that is, an approach lacking in emotional authenticity. Within this regard, it would be worthy to test how animators respond to the Alba Technique as a method for creating their reference performances, and whether the Alba Technique leads to reference performances that are well-suited for the purposes of a variety of animation styles. The ethics of adjusting the acting of a source-performer within a PeCap context is also an emerging and fascinating area that demands deep consideration and a clear response.

Another potential research area considers the phenomenon of mirror neurons to inquire to what degree an animated character must express an emotion in order for a viewer's brain to trigger empathy with the character. Also, how does the visual/graphic style of an animated character affect how a viewer empathises with the character? Butler & Joschko argue that photorealistic animation provides viewers with limited capacity to engage their imaginations in creating a unique visual experience, whereas more stylised animation allows greater "visual space to enrich the animation with their own experiences", which in turn establishes a stronger sense of empathy (2007, p. 62 [citing Riley, 2005]). While a detailed discussion about the relationship between animation style and empathy is beyond the scope of this paper, it is a worthy subject for future research.

Ed Hooks currently provides the most thorough acting training specific to animators, yet it would be worth investigating what other aspects of acting could be beneficial for PeCap animators to engage with. Dan Lemmon's observation that the indexical quality of a source-actor's performance exists more in the movements between poses than in the poses themselves is also a prime candidate for further inquiry that tests the validity of this conjecture.

There are a few areas I originally intended to include as part of this research but excluded due to space limitations. These include:

- a history of acting through Western animation and motion capture

- the role and potential of the virtual camera in the construction of acting within a PeCap context
- individual case studies of the performances of digital counterparts

In separate publications, I am producing a comprehensive catalogue of all vactor performers in feature films, which taxonomizes each vactor according to the categories of performance styles outlined in this research (subsection 5.4.2).

Finally, from a purely technical point of view, it would be extremely useful to conduct pragmatic research that demonstrates how to implement many of the high-level computer science insights about facial animation and simulation (e.g. Bermano, et al., 2014; Bermano, et al., 2015; Kozlov, et al., 2017; McDonagh, et al., 2016; Ravikumar, 2017; Zollhöfer, et al., 2018) into standard industry software, such as Maya, for democratised usage.

7.4 Conclusion: And the Award Goes to...

Producers have been at odds with how to "educate" audiences and awards bodies to understand motion-captured performances in a fashion that is sympathetic to the achievements of both actors and animators. Starting with Jackson's "King Kong", filmmakers have often produced publicity materials and postproduction diaries that position MoCap technologies within familiar discourses of celebrity, acting, and emotion. In part, this suggests the process creates a one-to-one translation of gesture and expression from a source-actor to a digital character. In so doing, the filmmakers conflate the source-actor with the character, which conveniently avoids the ontological problems of the digital counterpart (Allison, 2011, p. 333).[84]

[84] Allison's use of the term "cyberstar" here seems more consistent with my definition of *synthespian* rather than *cyberstar*.

Despite these efforts, confusion about how to understand motion capture performances still abounds. Fourteen years after Gollum received his first Oscar snub, Serkis lamented that “within our own industry, there is still a woeful lack of knowledge about the art and craft of performance capture; many people still don’t consider it to be ‘real’ acting.” (2016, para. 4). For example, James Lipton, host of “Inside the Actor’s Studio”, is quoted: “I’m not sure that motion capture, while it captures the flicker of an eyebrow, the twist of a mouth, a gesture of a hand, equally captures emotion.” (Block, 2010, para. 8). Serkis indicates that many actors remain anxious about the technology replacing them or rendering their performances unrecognisable, as well as animators who are afraid of losing their jobs (2016, para. 5). Regarding the actors in “Avatar” not receiving any Oscar nominations, producer Jon Landau called it “a disappointment, but I blame ourselves for not educating people in the right way.” (Block, 2010, para. 1). Ten years later, apparently just what the right way is still remains to be seen.

7.4.1 Embracing a New Realism

In the leadup to the 84th Academy Awards (2012), the producers of two movies that heavily featured Serkis campaigned for Oscar recognition. The producers of “The Adventures of Tintin”, which won the Golden Globe for best animated film, implored Academy voters to consider it for the Best Animated Feature Film, despite the fact that the character performances, including Serkis as Captain Haddock, were entirely produced using performance capture. That same year, Serkis’ portrayal of the ape Caesar in “Rise of the Planet of the Apes” (Wyatt, 2011) earned him critical acclaim and the hope for his first Best Actor Oscar nomination. Twentieth Century Fox advertised in industry trade papers and created behind-the-scenes footage comparing Serkis’ on-set performance to his digital ape counterpart “in a bid to convince Academy voters that his embodiment of the role was every bit as worthy of consideration as a traditional performance from the likes of Meryl Streep or Brad Pitt.” (Snyder, 2012, para. 2). Co-star James Franco even wrote a Deadline Hollywood article supporting Serkis: “What we need is recognition for him now, not later when this kind of acting is *de rigueur*, but now, when he has elevated this fresh mode of acting into an art form." (2012, *original emphasis*). In the article, Franco sees performance capture as the future of Hollywood – not a threat to traditional acting and actors, but as a necessary element within VFX-heavy films and a superior alternative to the common substitution of acting opposite tennis balls on sticks (Franco, 2012).

Ultimately, neither film received the nominations their producers vied for. Snyder speculates that Serkis' "digital skin" functioned as a deterrent to Academy voters accepting him "as a conventional leading man." (para. 6). He continues:

> The two biggest Oscar disses of 2012 leads one to call the true values of the Academy into question. Is the priority here to acknowledge the performers and the films that are pushing the boundaries and advancing the art form? Or is all this pageantry more about protecting the status quo? (para. 8).

Despite the Academy's reluctance to accept performance capture as either a recognisable acting or animation technique, many other awards bodies are willing to do so. At the 2003 MTV Movie Awards, Serkis is featured in a pre-recorded video to accept "on behalf" of Gollum the Award for Best Virtual Performance for "The Two Towers". With help from Weta Digital, Gollum interrupts Serkis midway through his acceptance speech. Gollum steals the Golden Popcorn award from Serkis' hands and proceeds to take full credit for the role in a profanity-laden diatribe that excoriates Serkis, Peter Jackson, Weta Digital animators, MTV and even Dobby from the Harry Potter franchise (O'Connell, 2014). At the 11th Annual Critics' Choice Awards, the Broadcast Film Critics Association presented a special award for Distinguished Achievement in Performing Arts for the film "King Kong". The award was accepted by Serkis, animation director Christian Rivers, and animation supervisor Joe Letteri on January 9, 2006 (Debruge, 2006, para. 6). Serkis was also presented with a Saturn Award at the 2012 Academy of Science Fiction, Fantasy & Horror Films for Best Supporting Actor in "Rise of the Planet of the Apes" (38th Saturn Awards, 2019). As I mentioned in subsection 2.1.2, the Austin Film Critics Association now recognises a Best Motion Capture/Special Effects Performance, with Josh Brolin being the first recipient of this award in 2018.

As evidenced by the proliferation of VFX-heavy blockbusters, as well as an increasing number of lower budget productions that embrace motion capture to varying degrees, Hollywood is home to "a new breed of film director" who is less concerned about capturing reality as it is, and more focused on constructing a reality that *could be* (Geralt, 2010, p. 3). Why is it so important to focus on Academy Award recognition? The award, considered to be the highest level of achievement in the film industry, belies fear, prejudice, and discrimination toward an increasingly popular method of acting, as well as a rift within the industry about who should receive recognition and in what manner. There are fewer opportunities for individual achievements to be recognised, the result of which is a push toward reducing PeCap contributions to mere cinematic components within a larger matrix of constructed performance (Balcerzak, 2013, p. 197). This issue also draws into question whether recognition for performance-captured acting should be awarded to just the source-actor, or if principal animators and supervisors should accompany the actor at the podium in order to observe proper recognition and fairness. This runs counter to historical conceptions about film acting as an activity performed by a single person; the recipients of the Best Actor awards never share the stage with another person as the award is understood to represent the achievement of an individual, not an ensemble, regardless of how many behind-the-scenes people help the actor achieve his or her on-screen portrayal. There is no award for Best Animator, either – only Best Animated Feature Film and Best Animated Short Film. Animation is understood to require the contributions of many individuals to achieve a final result, and the award is generally received by the film's director(s) and executive producer(s) on behalf of the film's crew. At the heart of the issue are the questions pursued through this thesis: how is acting produced, constructed, and preserved in performance capture? Who shares the stage with Gollum?

7.4.2 The Future of Acting in Performance Capture

The history of performance capture has been defined by the limitations of the various technological tools used to digitise performance as input and map it within a virtual space. Delbridge anticipates that the future of PeCap will be less defined by a tool-centred approach, that “the dominance of the twelve principles of animation as guiding principles to capturing performance will diminish, and the problematic confusion between key roles in the studio will cease as performance is privileged and captured.” (2014, p. 162). This carries with it the assumption that the quality of the capture will be identical enough to the source-performance that no data clean-up or modification will be required – and, in turn, less potential for the twelve principles of animation to play a role in the interpretation of a performance. This may become truer in the sense of realistic digital counterpart performances, but mapping a human actor’s source-performance onto a stylised character’s physiology will likely still require an animator’s artistry guided – at least in part – by a knowledge of the twelve principles of animation.

So long as cyberstars remain more costly to create than employing flesh-and-blood actors, there is little threat of actors being “replaced” by digital actors. The notion of replacement is also embedded in the belief that digital actors are able to achieve sentient performances similar to Victor Taransky’s Simone (examined in subsection 2.7.1). While this may one day be the case, it remains far from reality today. A fully digital replacement for an actor would require advanced artificial-intelligence capable of passing not only a Turing test but also able to believably create expressions and physical gestures – all while being impeccably rendered and without relying on the input of a flesh-and-blood actor (Perry, 2014, para. 26). While the technical challenges are immense, the exact process of how an audience buys into an actor’s fictional performance remains unclear, let alone how to represent such concepts within artificial intelligence. While this research does not attempt to provide an answer to the big question of how acting is created and transmitted on a psychophysiological level, it hopefully provides meaningful insights into how we conceive of the process of acting within a PeCap environment, and illuminates a path toward resolving these larger problems.

Instead of wholesale replacement of living actors with digital counterparts in film and television, we are witnessing an attenuation of the actor’s presumed control over his or her performance and image. While this may seem a distressful development to actors who are thoroughly invested in the status-quo of profilmic performance, younger actors are much more accepting of PeCap as a valid form of acting (Stern, 2014, para. 23) that provides opportunities beyond their natural physical assets. While King cites this as a source of anxiety for young actors (2011, p. 258), this runs counter to my experience of introducing fellow actors at the Kacie Stetson Studio to performance capture at Auckland University of Technology. These actors are impressed with the possibilities the technology affords them, and encouraged by opportunities that go beyond type-casting or conforming to a particular Hollywood beauty standard. Within the realm of PeCap-specific acting, perhaps a new form of stardom will emerge; it could be argued that Serkis is the first exemplar. However, this new star is no longer fetishized as a whole actor but rather as a commodity of performance parts, each used by degrees to produce a director's vision of a role that displaces the actor's physical body from the digital performance body. King notes that celebrity is embedded within an actor's claim to ownership of a particular physiology and presence, but what remains within PeCap is at best a vestigial presence (2011, p. 259). Since its conception, film has provided a commodification of the actor's image. PeCap now provides a commodification of the actor's performance, which can be repurposed to myriad images of a digital counterpart; the actor and image are no longer intimately linked.

Young actors may be hurt as much as helped by performance technologies. For instance, Kyle Smith of the New York Post notes that Lola VFX's technology used to de-age Robert Downey, Jr to a copy of his 20-year-old self in "Captain America: Civil War" likely "cost some young actor a chance at precious screen time." (2016, para. 13). Even when cast into a role, a PeCap actor is likely to earn less than his or her live-action co-stars (2016, para. 13). A question yet to be answered is whether an actor can become "type-cast" as a PeCap actor, specifically. Since acting agents primarily promote their talent based on visual evidence (photograph and acting show reels), an actor whose career begins in PeCap may find it difficult to compete with profilmic actors. An actor's performance becomes even more entangled when he or she performs through the digital counterpart of another actor – such as Guy-Henry-performing-Peter-Cushing-performing-Tarkin. Talent agents, casting agents, directors, and producers need to quickly become conversant in the range of performance possibilities that PeCap affords actors, as well as the level of artistry required by actors to perform such roles. Likewise, these same representatives need to understand how PeCap performances are constructed within the VFX pipeline: the degree to which a source-actor's performance is adjusted through animation largely depends on the visual style of a film and the workflows of individual VFX studios.

The future of performance capture will entail a further emphasis on character-driven performances. While MoCap has historically emphasised the recording and retargeting of motion as its primary purpose, PeCap's increasing ability to capture and transfer nuanced human performances will lead to a more intense focus on character acting that is supplemented by physicality, rather than the opposite. This will require more PeCap-specific training throughout the world – especially training that combines character acting with physicality. One of the forerunners in this area is The Mocap Vaults, co-founded by actor Oliver Hollis-Leick and based in the UK. The few other PeCap training programmes that exist are often one- or two-day workshops, and include The Mocap Academy (UK) and Captivate (UK and Australia). While short-courses are practical, I believe it would be valuable for actors to engage with PeCap-specific training over longer periods of time, similar to standard acting schools.

7.4.3 I'd Like to Thank All the Little People

As I have explored, a performance-captured role cannot be seen to exist as the source-actor's contribution alone; PeCap technicians, artists, and animators – not to mention the many 3D artists who construct the digital body – play significant roles within the creation of any digital character performance. The degree to which the actor maintains authorship of the final onscreen performance greatly depends on the nature of the production in a given VFX studio, as well as the degree of stylisation of a digital character. Whenever animators have to "interpret" the source-acting onto a digital character that varies from the source-actor's physiology, the source-actor's authorship of the final performance reduces according to the degree of physiological variance. How this difference in authorship is apportioned among 3D artists is likely impossible to measure. The question is whether authorship remains a meaningful metric by which to judge performance-captured roles, and how a shift in thinking around this problem would lead to alternative means of recognising acting within PeCap.

8

WORKS CITED

8.1 Bibliography

38th Saturn Awards. (2019). In *Wikipedia*. Retrieved December 9, 2019, from https://en.wikipedia.org/wiki/38th_Saturn_Awards

Aldred, J. (2011). From synthespian to avatar: Re-framing the digital human in final fantasy and the polar express. *Mediascape*, 1-12, http://www.tft.ucla.edu/mediascape/Winter2011_Avatar.pdf

Allison, T. (2011). More than a man in a monkey suit: Andy Serkis, motion capture, and digital realism. *Quarterly Review of Film and Video, 28*(4), 325-341.

Allison, T. (2015). Blackface, *happy feet*: the politics of race in motion capture and animation. In D. North, B. Rehak, & M. Duffy (Eds.), *Special effects: new histories/theories/contexts*. (pp. 114-126). London, UK: BFI Palgrave.

Amidi, A. (2014a). *'Lord of the Rings' animation supervisor Randall William Cook speaks out on Andy Serkis*. Retrieved November 4, 2017, from http://www.cartoonbrew.com/motion-capture/lord-of-the-rings-animation-supervisor-randall-william-cook-speaks-out-on-andy-serkis-99439.html

Amidi, A. (2014b). *Andy Serkis does everything, animators do nothing, says Andy Serkis*. Retrieved March 13, 2020, from https://www.cartoonbrew.com/motion-capture/andy-serkis-does-everything-animators-do-nothing-says-andy-serkis-98868.html

Anderson, J. (2004/5). Note & comment: what's wrong with this picture?: dead or alive: protecting actors in the age of virtual reanimation. *Loyola of Los Angeles Entertainment Law Review, 25*, 155-201.

Anzovin, S., & Anzovin, R. (2005). *3d toons: creative 3d design for cartoonists and animators*. Lewes, UK: Ilex.

Appelo, T. (2011). *Andy Serkis on who send him hate mail: why Oscar should rethink animated category (Q&A)*. Retrieved December 1, 2019, from: https://www.hollywoodreporter.com/news/andy-serkis-oscars-adventures-of-tintin-268750

Artaud, A. (1976). *Antonin Artaud: Selected writings*. Los Angeles, CA, USA: University of California Press.

Auckland University of Technology. (2020). *Applying for ethics approval: guidelines and procedures*. Retrieved January 21, 2020, from https://www.aut.ac.nz/research/researchethics/guidelines-and-procedures#6

Auslander, P. (2008). *Liveness: performance in a mediatized culture*. New York, USA: Routledge.

Balcerzak, S. (2013). Andy Serkis as actor, body and gorilla: motion capture and the presence of performance. In S. Balcerzak & J. Sperb (Eds.), *Cinephilia in the age of digital reproduction, vol. 1.* (pp. 195-213). London, UK: Wallflower Press.

Bancroft, T. (2014). *Directing for animation: everything you didn't learn in art school.* Burlington, MA, USA: Focal Press.

Barish, J. (1969). Exhibitionism and the antitheatrical prejudice. ELH, 36(1), 1-29.

Barnett, D. (2021). *Realism*. Retrieved January 9, 2021, from https://brechtinpractice.org/theory/realism/

Bartlett, W. (2013). *Galaxy chauffeur*. Retrieved February 13, 2020, from https://www.framestore.com/work/chauffeur

Bates, J. (1994). The role of emotion in believable agents. *Communications of the ACM*, 37(7), 122-125.

Bazin, A., & Gray, H. (1960). The Ontology of the Photographic Image. *Film Quarterly, 13*(4), 4-9. DOI:10.2307/1210183

Beiman, N. (2010). *Animated performance: Bringing imaginary animal, human and fantasy character to life*. Lausanne, Switzerland: AVA Publishing.

Benjamin, W. (1969). The work of art in the age of mechanical reproduction. In H. Arendt (Ed.), *Illuminations*, (pp. 1-26), New York, New York, USA: Schocken Books.

Bermano, A., Bradley, D., Beeler, T., Zünd, F, Nowrouzezahrai, D., Baran, I., . . . Gross, M. (2014). Facial performance enhancement using dynamic shape space analysis. *ACM Transcriptions on Graphics, 33*(2): 13. DOI: 10.1145/2546276

Bermano, A., Beeler, T., Kozlov, Y., Bradley, D. Bickel, B., and Gross, M. (2015). Detailed spatio-temporal reconstruction of eyelids. *ACM Transactions on Graphics, 34*(4), Article 44, pp.1-11. DOI: 10.1145/2766924.

Bestor, N. (2016). The technologically determined decade: Robert Zemeckis, Andy Serkis, and the promotion of performance capture." *Animation: An Interdisciplinary Journal, 11*(2), 169-188.

Biggs, M. A. R. (2004). Learning from experience: approaches to the experiential component of practice-based research. *Forskning-Reflektion-Utveckling*. Stockholm, Sweden: Swedish Research Council, Vetenskapsrådet, 6-21.

Billington, A. (2007). *From motion capture to 3D: the technology of Beowulf.* Retrieved September 13, 2007, from http://www.firstshowing.net/2007/from-motion-capture-to-3d-the-technology-of-beowulf

Bishko, L. (2007). The uses and abuses of cartoon style in animation. *Animation: An Interdisciplinary Journal, 2*, 24-35.

Bishko, L. (2014). Our empathic experience of believable characters. In M. Nixon (Ed.), *Nonverbal communication in virtual worlds* (pp. 47-59). Halifax, Nova Scotia, Canada: ETC Press.

Blair, R. (2002). Reconsidering Stanislavsky: feeling, feminism, and the actor. *Theatre Topics, 12*(2), pp. 177-190.

Blair, R. (2008). *The actor, image, and action: acting and cognitive neuroscience*. Routledge: New York.

Bloch, S. (1993). Alba emoting: a psychophysiological technique to help actors create and control real emotions. *Theatre Topics, 3*(2), pp. 121-138. DOI: 10.1353/tt.2010.0017

Bloch, S., Orthous, P., and Santibáñez-H, G. (1987). Effector patterns of basic emotions: a psychophysiological method for training actors. *Journal of Social and Biological Structures, 10*, pp. 1-19.

Block, A. (2010). *Oscars snub 'Avatar's' motion-capture actors*. Retrieved January 10, 2016, from http://www.hollywoodreporter.com/news/oscars-snub-avatars-motion-capture-20207

Bode, L. (2005). Digital doppelgängers. *M/C Journal, 8*(3). Retrieved from http://journal.media-culture.org.au/0507/07-bode.php

Bode, L. (2006). From shadow citizens to Teflon stars: reception of the transfiguring effects of new moving image technologies. *Animation: An Interdisciplinary Journal, 1*(2), 173-189.

Bode, L. (2008, July). *Making faces: Hybridity, animation and the screen actor*. Paper presented at the 20th Society for Animation Studies Conference, Bournemouth, UK. Retrieved from the author.

Bode, L. (2010a). No longer themselves?: framing digitally enabled posthumous performance. *Cinema Journal 49*(4), 46-70.

Bode, L. (2010b). Avatar, "e-motion capture", and the shifting industry rhetoric around performance/animation hybrids. *Society for Animation Studies Conference*. Edinburgh, Scotland.

Bode, L. (2015). Fleshing it out: prosthetic makeup effects, motion capture and the reception of performance. In D. North, B. Rehak & M. Duffy (Eds.), *Special effects: new histories, theories, contexts* (pp. 32-44), London, United Kingdom: British Film Institute, Palgrave Macmillan.

Bode, L. (2017). *Making believe: Screen performance and special effects in popular cinema.* New Brunswick, New Jersey, USA: Rutgers University Press.

Bouldin, J. (2001). The body, animation and the real: race, reality and the rotoscope of

Betty Boop. *Proceedings of the Affective Encounters: Rethinking Embodiment in Feminist Media Studies Conference*, University of Turku, 48-54.

BreathXpress. (n.d.). *Alba emoting*. Retrieved January 30, 2020, from https://breathxpress.wordpress.com/alba-emoting/

Brigg, M., & Bleiker, R. (2010). Autoethnographic International Relations: exploring the self as a source of knowledge. *Review of International Studies, 36*, 779-798.

Broadway World. (2010). *Rules approved for 83rd Academy Awards*. Retrieved December 7, 2019, from https://www.broadwayworld.com/los-angeles/article/Rules-Approved-for-83rd-Academy-Awards-20100708

Brody, J. (2017). *Acting, archetype, & neuroscience: superscenes for rehearsal & performance*. Abingdon, UK: Routledge.

Burrill, D. (2005). Out of the box: performance, drama, and interactive software. *Modern Drama, 48*(3), 492-512.

Busso, C., & Narayanan, S. (2006). Interplay between linguistic and affective goals in facial expression during emotional utterances. In *7th International Seminar on Speech Production (ISSP 2006)*, 549-556.

Butler, M., & Joschko, L. (2007). Final fantasy or the Incredibles. *Animation Studies, 4*(55), pp.55-63.

Caesar, E. (2016). *Sir Mark Rylance: the greatest actor of his generation*. Retrieved February 1, 2020, from https://www.gq-magazine.co.uk/article/mark-rylance-interview-bbc-wolf-hall-thomas-cromwell

Candy, L. (2006). Practice based research: a guide. *CSS Report, 1*, 1-19.

Carnicke, S. (2009). *Stanislavsky in focus: an acting master for the twenty-first century (2nd ed.)*. New York, NY, USA: Routledge.

Carter, K. (2015). *Why Lupita Nyong'o didn't want to be seen in "star wars"*. Retrieved November 12, 2017, from https://www.buzzfeed.com/kelleylcarter/lupita-nyongo-didnt-want-you-to-see-her-face-in-star-wars

Cash, J. (2020). *Realism and naturalism theatre conventions*. Retrieved January 9, 2021, from https://thedramateacher.com/realism-and-naturalism-theatre-conventions/

Caton, C. (2017). *Animation for animatronics part 2: animating servos & virtual mechanisms*. Retrieved December 23, 2019, from https://www.stanwinstonschool.com/tutorials/animation-for-animatronics-2-animating-servos-virtual-mechanisms

Cavna, M. (2016). *One of the best performances in 'rogue one' is by an actor who died in 1994*. Retrieved February 2, 2020, from https://www.washingtonpost.com/news/comic-riffs/wp/2016/12/15/one-of-the-best-performances-in-rogue-one-is-by-an-actor-who-died-in-1994/

Chaikin, J. (1991). *The presence of the actor*. New York, NY, USA: Theatre Communications Group, Inc.

Chekhov, A. (1997). *The seagull* (J-C van Itallie, Trans.). New York, NY, USA: Dramatists Play Service.

Chimielewski, D. (2005). *Meet Sunny's digital doppelganger*. Retrieved December 20, 2018, from http://www.theage.com.au/news/Film/Meet-Sunnys-digitaldoppelganger/2005/01/04/1104601340883.html

Chubbuck, I. (2005). *The power of the actor: the Chubbuck Technique*. New York, NY, USA: Gotham Books.

Cohen, D. (2014). *Andy Serkis vs. visual effects animators: the wrong fight for both sides.* Retrieved January 10, 2016, from http://variety.com/2014/film/columns/andy-serkis-vs-visual-effects-animators-the-wrong-fight-for-both-sides-1201260274/

Collins English Dictionary. (2018). *Definition of 'synthespian'*. Retrieved December 18, 2018, from https://www.collinsdictionary.com/dictionary/english/synthespian

cpn_admin. (2012). *Performance captured*. Retrieved November 23, 2016, from http://www.creativeplanetnetwork.com/news/news-articles/performance-captured/380975

Crafton, D. (2013). *Shadow of a mouse: performance, belief, and world-making in animation*. Los Angeles, CA, USA: University of California Press.

Creed, B. (2000). The cyberstar: digital pleasures and the end of the unconscious. *Screen, 41*(1), 79-86.

Crivelli, C., Jarillo, S., Russell, J. A., & Fernández-Dols, J. M. (2016). Reading emotions from faces in two indigenous societies. *Journal of Experimental Psychology: General*, *145*(7), 830-843.

da Silva, R., dos Santos, M., & Iurgel, I. (2011). Developing virtual actors inspired by real actors' practice. *2011 Workshop on Digital Media and Digital Content Management*, 146-149.

Dale, D. (2003). *Can Gollum win an Oscar?* Retrieved December 5, 2019, from https://www.theage.com.au/entertainment/movies/can-gollum-win-an-oscar-20030109-gdv1fy.html

Debruge, P. (2006). *Is it actor, ape or computer?* Retrieved January 10, 2016, from http://variety.com/2006/film/awards/is-it-actor-ape-or-computer-1117935630/

Delbridge, M. (2014). *The cooling steam of the Polar Express: historical origins, properties and implications of performance capture* (Doctoral dissertation, Queensland University of Technology, Brisbane, QLD, Australia). Retrieved from http://eprints.qut.edu.au/72193/

Delbridge, M. (2015). *Motion capture in performance: an introduction*. London, UK:

Palgrave MacMillan.

Deleuze, G. (1969). *The logic of sense* (M. Lester, Trans.). New York: Columbia University Press.

Deleuze, G., & Guattari, F. (1987). *A thousand plateaus: Capitalism and schizophrenia* (B. Massumi, Trans.). Minneapolis, MN: University of Minnesota Press.

Dixon, S. (2007). *Digital Performance: A History of New Media in Theater, Dance, Performance Art, and Installation*. Cambridge, Massachusetts: MIT Press.

Duncan, J. (2009). The unusual birth of Benjamin Button. *Cinefex, 116*, 70-118.

Duncan, J. (2010). The seduction of reality. *Cinefex, 120*, 68-140.

Edwards, G. (2017). A lonely god. *Cinefex, 152*, 58-83.

Leyda, J. (Ed.) (1988). *Eisenstein on Disney* (A. Upchurch, Trans.). New York, NY, USA: Methuen.

Ekman, P. (1979). About brows: emotional and conversational signals. In M. von Cranach, K. Foppa, W. Lepenies, & D. Ploog (Eds.), *Human ethology: Claims and limits of a new discipline* (pp. 169-202). London, UK: Cambridge University Press.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion, 6*(3-4), 169-200.

Ekman, P. (2004). Emotional and conversational nonverbal signals. In J. Larrazabal & L. Pérez Miranda (Eds.), *Language, knowledge, and representation* (pp. 39-50). Dordrecht, Netherlands: Kluwer Academic Publishers.

Ekman, P. (2016). *Micro expressions*. Retrieved March 29, 2016, from http://www.paulekman.com/microexpressions/

Ekman, P., & Friesen, W. (1986). A new pan-cultural facial expression of emotion. *Motivation and Emotion, 10*(2), 159-168.

Erickson, R. (1995). The importance of authenticity for self and society. *Symbolic Interaction, 18*(2), 121-144.

Faceware Technologies, Inc. (2017). *Faceware Analyzer: professional tracking software*. Retrieved October 14, 2017, from http://www.facewaretech.com/products/software/analyzer/

Faceware Technologies, Inc. (2020). *Free assets*. Retrieved April 11, 2020, from https://www.facewaretech.com/learn/free-assets

Failes, I. (2017). *From performance capture to creature: how the apes were created in 'War for the Planet of the Apes'*. Retrieved November 29, 2019, from http://www.cartoonbrew.com/vfx/performance-capture-creature-apes-created-war-planet-apes-152357.html

Failes, I. (2019). *'Computer pajamas': the history of ILM's IMocap.* Retrieved January 4, 2021, from https://beforesandafters.com/2019/09/10/computer-pajamas-the-history-of-ilms-imocap/

Find an Expert. (2017). *A/Prof Matthew Delbridge*. Retrieved May 2, 2017, from https://findanexpert.unimelb.edu.au/display/person359180

Fischer, P. (2005). *Andy Serkis for "King Kong"*. Retrieved November 26, 2016, from http://www.darkhorizons.com/andy-serkis-for-king-kong/

Fleming, D. (2012). Digitalising Deleuze: the curious case of the digital human assemblage, or what can a digital body do? In Martin-Jones, D. (Ed.), *Deleuze and Film*. (pp. 192-209). Edinburgh, UK: Edinburgh University Press.

Fonteyn, M. E., Kuipers, B., & Grobe, S. J. (1993). A description of think aloud method and protocol analysis. *Qualitative Health Research, 3*(4), 430-441.

Fordham, J. (2017a). Rogue one: Sacred ground. *Cinefex, 151*, pp. 1-37.

Fordham, J. (2017b). Endangered species. *Cinefex, 154*, pp. 12-33.

Franco, J. (2012). *Oscar exclusive: James Franco on why Andy Serkis deserves credit from actors*. Retrieved December 9, 2019, from https://deadline.com/2012/01/oscar-exclusive-james-franco-on-why-andy-serkis-deserves-credit-from-actors-211525/

Freedman, Y. (2012). Is it real… or is it motion capture? The battle to redefine animation in the age of digital performance. *The Velvet Light Trap, 69*, 38-49.

Freud, S. (2004). The uncanny. In Sandner, D. (Ed.), *Fantastic literature: a critical reader* (pp. 74-101). Westport, CT, USA: Praeger Publishers.

Frijda, N., & Tcherkassof, A. (1997). Facial expressions as modes of action readiness. In. J. Russell & J. Fernández-Dols (Eds.), *The psychology of facial expression* (pp. 78-102). Cambridge, UK: Cambridge University Press.

Furniss, M. (1999). *Motion capture*. Retrieved December 23, 2019, from http://web.mit.edu/comm-forum/legacy/papers/furniss.html

Garrido, P., Valgaerts, L., Wu, C., and Theobalt, C. (2013). Reconstructing detailed dynamic face geometry from monocular video. *ACM Transcriptions on Graphics, 32*(6), 1-10.

Garrido, P., Zollhöfer, M., Wu, C., Bradley, D., Pérez, P., Beeler, T., Theobalt, C. (2016). Corrective 3D reconstruction of lips from monocular video. *ACM Transcriptions on Graphics, 35*(6): 219. DOI: 10.1145/2980179.2982419

Giardina, C. (2015). *How 'Furious 7' brought the late Paul Walker back to life*. Retrieved February 2, 2020, from https://www.hollywoodreporter.com/behind-screen/how-furious-7-brought-late-845763

Giardina, C. (2016). *How 'Captain America: civil war' turned Robert Downey Jr. back into a teen*. Retrieved February 1, 2020, from https://www.hollywoodreporter.com/behind-screen/how-captain-america-civil-war-892387

Giralt, G. (2010). Realism and realistic representation in the digital age. *Journal of Film and Video, 62*(3), 3-16. DOI: 10.1353/jfv.2010.0000

Glebas, F. (2009). *Directing the story: professional storytelling and storyboarding techniques for live action and animation*. Burlington, MA, USA: Focal Press.

Goldstein, J. (2003). The construction of emergent order, or, how to resist the temptation of hylozoism. *Nonlinear Dynamics, Psychology, and Life Sciences, 7*(4), 295-314.

Goldstein, T. R., & Levy, A. G. (2017). The constricted muse: acting. In J. Kaufman, V. Glaveanu, & J. Baer (Eds.), *Creativity across domains* (pp. 145-160). Cambridge, UK: Cambridge University Press.

Gosselin, P., Kirouac, G., & Doré, F. (2005). Components and recognition of facial expression in the communication of emotion by actors. In P. Ekman, & E. Rosenberg (Eds.), *What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system (FACS) (2nd ed)* (pp. 243-267). New York, NY, USA: Oxford University Press.

Grady, C. (2018). *Mark Rylance is one of our greatest living actors*. Retrieved February 1, 2020, from https://www.vox.com/culture/2018/3/23/17093508/mark-rylance-greatest-living-actor-explained

Gray, A. (2014). *A brief history of motion-capture in the movies*. Retrieved January 6, 2020, from https://za.ign.com/the-polar-express/76815/feature/a-brief-history-of-motion-capture-in-the-movies

Gray, D. (2004). *Doing research in the real world*. Los Angeles, USA: Sage.

Gunning, T. (2004). What's the point of an index? Or, faking photographs. *Nordicom Review, 5*(1/2), 39-49.

Gunning, T. (2007). Moving away from the index: Cinema and the impression of reality. *Differences: A Journal of Feminist Cultural Studies, 18*(1), 29-52.

Hadjioannou, M. (2008). How does the digital matter? Envisioning corporeality through Christian Volckman's Renaissance. *Studies in French Cinema, 8*(2), 123-136. DOI: 10.1386/sfc.8.2.123_1

Hagen, U. (1991). *A challenge for the actor.* New York, NY, USA: Scribner.

Hagen, U., and Frankel, H. (1973). *Respect for acting*. New York, NY, USA: Wiley Publishing.

Harper, D. (2020). *Authentic (adj.)*. Retrieved March 25, 2020, from https://www.etymonline.com/word/authentic

Hammond, W. (1895). Hylozoism: a chapter in the early history of science. *The Philosophical Review, 4*(4), 394-406.

Hampton, D. (2014). *Written all over your face*. Retrieved December 9, 2017, from https://thebestbrainpossible.com/written-all-over-your-face/

Hayes, D., & Webster, D. (2013). *Acting and performance for animation*. Burlington, MA: Focal Press.

Heathfield, A. (2004). *Live: Art and performance.* London, UK: Tate Publishing.

Henderson, L., Henderson, M., Grant, S., & Huang, H. (2010). What are users thinking in a virtual world lesson? Using stimulated recall interviews to report student cognition, and its triggers. *Journal of Virtual Worlds Research, 3*(1), 3-22.

Hethmon, R. (Ed.). (1965). *Strasberg at The Actors Studio*. New York, NY, USA: Theatre Communications Group, Inc.

Hiatt, B. (2014). *How Andy Serkis became the king of post-human acting*. Retrieved January 10, 2016, from http://www.rollingstone.com/movies/news/how-andy-serkis-became-the-king-of-post-human-acting-20140714

Hobbes, T. (2011). *Leviathan*. Seattle, WA, USA: Pacific Publishing Studio.

Hoffer, T. (1981). *Animation, a reference guide*. Westport, CT, USA: Greenwood.

Hooks, E. (2005). *Acting in animation: a look at 12 films*. Portsmouth, NH, USA: Heinemann Drama.

Hooks, E. (2011). *Acting for animators (3rd ed.)*. New York, NY, USA: Routledge.

Hooks, E. (2017). *Craft notes for animators: A perspective on a 21st century career*. New York, NY, USA: Routledge.

Hosea, B. (2009). Performativity, post-animation and how I became a cartoon character. *Illuminance, 4*, 1-19.

Hosea, B. (2012). *Substitutive bodies and constructed actors: a practice-based investigation of animation as performance* (Doctoral dissertation, University of the Arts London, London, UK). Retrieved from http://ualresearchonline.arts.ac.uk/3437/

Hoyle, B. (2007). *And for my acting Oscar, I thank the special effects*. Retrieved October 16, 2017, from https://www.thetimes.co.uk/article/and-for-my-acting-oscar-i-thank-the-special-effects-8qfmg9cjphr

Huston, H. (1992). *The actor's instrument: body, theory, stage*. Ann Arbor, MI, USA: The University of Michigan Press.

Ings, W. (2013). Narcissus and the muse: supervisory implications of autobiographical, practice-led PhD design theses. *Qualitative Research, 14*(6), 675-693.

InspiredMag. (2013). *The various applications of motion capture software*. Retrieved October 9, 2017, from http://inspiredm.com/various-applications-motion-capture-software/

Internet Movie Database (IMDb). (2020). *The internet movie database*. Retrieved January 22, 2020, from http://www.imdb.com

Itzkoff, D. (2016). *How 'rogue one' brought back familiar faces*. Retrieved February 2, 2020, from https://www.nytimes.com/2016/12/27/movies/how-rogue-one-brought-back-grand-moff-tarkin.html

Izard, C. E. (1997). Emotions and facial expressions: a perspective from differential emotions theory. In J. Russell, & J. Fernández-Dols (Eds.), *The psychology of facial expression* (pp. 57-77). Cambridge, UK: Cambridge University Press.

Jensenius, A. (2013). Some video abstraction techniques for displaying body movement in analysis and performance. *Leonardo, 46*(1), 53-60.

Jones, A., & Oliff, J. (2007). *Thinking animation: bridging the gap between 2d and CG*. Boston, MA, USA: Thomson Course Technology PTR.

Joon, J. (2008). *Human motion based on actor physique using motion capture: A case study.* Saarbrücken, Germany: VDM Verlag.

Keane, S. *How gemini man's vfx whiz made two versions of Will Smith.* Retrieved January 4, 2021, from https://www.cnet.com/news/how-gemini-man-vfx-whiz-made-two-versions-of-will-smith/

Kehr, D. (2004). *A face that launched a thousand chips*. Retrieved November 23, 2016, from http://www.nytimes.com/2004/10/24/movies/a-face-that-launched-a-thousand-chips.html

Kemp, R. (2012). *Embodied acting: what neuroscience tells us about performance*. Abingdon, UK: Routledge.

Kennedy, J. (2014). Character acting: a case for better animation reference. *Proceedings of the 4th Annual Conference of the Popular Culture Association of Australia and New Zealand (PopCAANZ)*, 11-20.

Kennedy, J. (2015a). Gauging meaningful reference performance in animation and motion capture. *Proceedings of the Cumulus Conference, Milano 2015*, 941-948.

Kennedy, J. (2015b). Beyond the mirror: producing emotionally-authentic facial performance for animation reference. *Proceedings of the 7th International Conference on Illustration & Animation (CONFIA)*, 101-114.

Kennedy, J. (2017). The animator's (missing) hand: how practice informs seeing in 3d animation. *Proceedings of the 7th International Conference on Illustration & Animation (CONFIA)*, 298-309.

Kennedy, J. (2019a). Acting-centred definitions of vactors, synthespians, and digital doubles. *Proceedings of the 25th International Symposium on Electronic Art (ISEA), 25*, 297-305.

Kennedy, J. (2019b). Critiquing the screen presence of synthespian counterparts. *Proceedings of the 9th International Conference on Illustration & Animation (CONFIA)*, 190-202.

Kennedy, J. (2020). Vactor ontologies: framing acting within a motion capture context. *In press*.

Kilday, G. (2018). *Academy postponing new popular Oscar category*. Retrieved December 7, 2019, from https://www.hollywoodreporter.com/news/academy-postponing-new-popular-oscar-category-1140423

King, B. (2011). Articulating digital stardom. *Celebrity Studies, 2*(3), 247-262.

Kirby, M. (1972). On acting and not-acting. *The Drama Review: TDR, 16*(1), 3-15.

Kitagawa, M., & Windsor, B. (2008). *MoCap for artists: workflow and techniques for motion capture*. Burlington, MA: Elsevier, Inc.

Kiwitt, P. (2012). What is cinema in a digital age? Divergent definitions from a production perspective. *Journal of Film and Video, 64*(4), 3-22.

Klein, N. (1993). *7 minutes: the life and death of the American animated cartoon*. New York, NY, USA: Verso.

Kleinman, J. (2019). *'Gemini man' cgi isn't like marvel's de-aging – it's way better.* Retrieved January 4, 2021, from https://www.inverse.com/article/58444-gemini-man-movie-will-smith-junior-cgi-technology-explained

Kozlov, Y., Bradley, D., Bächer, M., Thomaszewski, B., Beeler, T., Gross, M. (2017). Enriching facial blendshape rigs with physical simulation. *Computer Graphics Forum, 36*(2), 75-84.

Krasner, D. (2010). Strasberg, Adler, and Meisner: Method acting. In A. Hodge (Ed.), *Actor training (2nd ed.)*. New York, NY, USA: Routledge.

Ktorides, C. (2014). *Dawn of the planet of the apes' Andy Serkis: "it's so much easier working on primates"*. Retrieved December 1, 2019, from: http://diymag.com/2014/07/17/dawn-of-the-planet-of-the-apes-andy-serkis-interview

Landay, L. (2012). The mirror of performance: kinaesthetic, subjectivity, and the body in film, television, and virtual worlds. *Cinema Journal, 51*(3), 129-136.

LASALLE College of the Arts. (2016). *Professor Steve Dixon*. Retrieved October 3, 2017, from http://www.lasalle.edu.sg/academics/professor-steve-dixon/

Le Boeuf, P. (2010). On the Nature of Edward Gordon Craig's Über-Marionette. *New Theatre Quarterly, 26*(2), 102-114.

Leder, D. (1990). *The absent body*. Chicago, IL, USA: University of Chicago Press.

Lederer, J. (2012). *When will a motion-capture actor win an Oscar?* Retrieved December 1, 2019, from https://www.wired.com/2012/01/andy-serkis-oscars/Lexico. (2020). Doppelgänger. Retrieved January 19, 2020, from https://www.lexico.com/definition/doppelganger

Li, H., Weise, T., and Pauly, M. (2010). Example-based facial rigging. *Proceedings of ACM SIGGRAPH 2010, 29*(4), pp. 1-6. DOI: 10.1145/1833349.1778769

Liverman, M. (2004). *The animator's motion capture guide*. Newton, MA, USA: Charles River Media.

Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance, 31*(1), 210-220. DOI: 10.1037/0096-1523.31.1.210

Konijn, E. (2000). *Acting emotions*. Amsterdam, Netherlands: Amsterdam University Press.

Maibom, H. (2007). The presence of others. *Philosophical Studies, 132*, 161–190.

Manovich, L. (2001). *The language of new media*. Cambridge, MA, USA: MIT Press.

Manovich, L. (2016). What is digital cinema? In. S. Denson & J. Leyda (Eds.), *Post-cinema: Theorizing 21st-century film* (pp. 20-50). REFRAME Books. Retrieved January 6, 2021 from https://mediarep.org/bitstream/handle/doc/14379/Denson_Leyda_2016_Post-Cinema_.pdf?sequence=4#page=43

McDonagh, S., Klaudiny, M., Bradley, D., Beeler, T., Matthews, I., and Mitchell, K. (2016). Synthetic prior design for real-time face tracking. *Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV)*, pp. 639-648.

McNary, D. (2011). *Motion-capture tests acting boundaries*. Retrieved January 10, 2016, from http://variety.com/2011/digital/awards/motion-capture-tests-acting-boundaries-1118047652/

Medeiros, J. (2014). *How Andy Serkis changed the face of film-making*. Retrieved December 1, 2019, from https://www.wired.co.uk/article/planet-of-the-apes-andy-serkis

Menache, A. (2011) *Understanding motion capture for computer animation (2nd ed.)* Burlington, MA, USA: Elsevier.

Merleau-Ponty, M. (2014). *Phenomenology of perception*. Abingdon, UK: Routledge.

MetaMotion. (n.d.). *Motion capture uses*. Retrieved December 23, 2019, from https://metamotion.com/motion-capture-uses/motion-capture-applications.html

Merriam-Webster. (2020). *Presence*. Retrieved February 15, 2020, from https://www.merriam-webster.com/dictionary/presence

Merriam-Webster. (2021). *The origin of 'hypocrite'*. Retrieved January 8, 2021, from https://www.merriam-webster.com/words-at-play/hypocrite-meaning-origin

Misselhorn, C. (2009). Empathy with inanimate objects and the Uncanny Valley. *Minds & Machines, 19*, pp. 345-359. DOI: 10.1007/s11023-009-9158-2

Morgan, A. (1889). *An hour with Delsarte: a study of expression*. Lee and Shepard: Boston, MA, USA.

Mori, M. (1970). The uncanny valley (K. F. MacDorman & T. Mianto, Trans.). *Energy, 7*(4), 33-35.

Nansense. (2019a). *Pricing & discounts*. Retrieved January 4, 2020, from https://www.nansense.com/pricing

Nansense. (2019b). *Gloves*. Retrieved January 6, 2020, from https://www.nansense.com/gloves/

Navarro, J. (2016). *Chirality: a look at emotional asymmetry of the face.* Retrieved April 7, 2017, from https://www.psychologytoday.com/blog/spycatcher/201605/chirality-look-emotional-asymmetry-the-face

Ngunjiri, F. W., Hernandez, K. C., & Chang, H. (2010). Living autoethnography: connecting life and research. *Journal of Research Practice, 6*(1), Article E1. Retrieved October 22, 2017, from http://jrp.icaap.org/index.php/jrp/article/view/241/186

Nikolai, J., Bennett, G., Marks, S., and Gilson, G. (2019). Active learning and teaching through digital technology and live performance: 'choreographic thinking' as art practice in the tertiary sector. *The International Journal of Art & Design Education, 38*, 137-152. DOI: 10.1111/jade.12181

North, D. (2008). *Performing Illusions: Cinema, Special Effects and the Virtual Actor.* London, UK: Wallflower Press.

O'Neill, R. (2016). *Digital character development: Theory and practice* (2nd ed.). Boca Raton, FL, USA: CRC Press.

Ortony, A. and Turner, T. (1990). What's Basic About Basic Emotions? *Psychological Review, 97*(3), 315-331.

Osborn, K. (2015). *Cartoon character animation with Maya: Mastering the art of exaggerated animation*. London, UK: Bloomsbury.

Osipa, J. (2010). *Stop staring: Facial modeling and animation done right* (4th ed.). Indianapolis, IN, USA: Wiley Publishing.

Oxford Living Dictionary. (2017). *Hyperreal*. Retrieved October 29, 2017, from https://en.oxforddictionaries.com/definition/hyperreal

Oxford Living Dictionary. (2018). *Synthespian*. Retrieved December 18, 2018, from https://en.oxforddictionaries.com/definition/synthespian

Pallant, C. (2012). Digital dimensions in actorly performance: the aesthetic potential of performance capture. *Film International, 10*(3), 37-49. DOI:10.1386/fiin.10.3.37_1

PaulEkmanGroup. (2017). *About Paul Ekman*. Retrieved December 1, 2017, from https://www.paulekman.com/about/paul-ekman/

Peirce, C. (1906). Prolegomena to an apology for pragmaticism. *The Monist, 16*, 492-546.

Perception Neuron. (2020). *Perception Neuron*. Retrieved January 6, 2020, from https://neuronmocap.com/

Perry, T. (2014). Digital actors go beyond the Uncanny Valley. *IEEE Spectrum*. Retrieved December 1, 2019, from: https://spectrum.ieee.org/computing/software/digital-actors-go-beyond-the-uncanny-valley

Pessino, A. (2004). Mistaken identity: a call to strengthen publicity right for digital personas. *Virginia Sports and Entertainment Law Journal, 4*(1), 86-118.

Phelan, P. (1996). *Unmarked: the politics of performance*. New York, NY, USA: Routledge.

Pizzo, A. (2016). *Actors and acting in motion capture* (A. Pizzo, Trans.). *Acting Archives Review, 11* (May 2016), pp. 38-69.

Pollick, F. (2010). In search of the uncanny valley. In P. Daras, & O. Mayora (Eds.), *User centric media: first international conference, UCMedia 2009*, (pp. 69-78). Berlin, Germany: Springer.

Power, P. (2009). Animated expressions: expressive style in 3d computer graphic narrative animation. *Animation: An Interdisciplinary Journal, 4*(2), 107-129.

Preston, S., & de Waal, F. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*, 1-72.

Psychology Today. (2020). *Joe Navarro M.A.* Retrieved February 14, 2020, from https://www.psychologytoday.com/nz/experts/joe-navarro-ma

Purves, B. (2008). *Stop motion: Passion, process, and performance.* Amsterdam, Netherlands: Focal Press.

Ravikumar, S. (n.d.). *Home*. Retrieved December 11, 2019, from http://shridharravikumar.com/

Ravikumar, S. (2017). *Performance driven facial animation with blendshapes* (Doctoral thesis, University of Bath, Bath, UK). Retrieved from https://purehost.bath.ac.uk/ws/portalfiles/portal/187906933/Thesis_Electronic.pdf

Redhead, L. (2012). *An exploration of the idea and importance of practice-led research in the current climate*. Retrieved March 3, 2017, from https://artspages.wordpress.com/tag/practice-based-research/

Ressner, J. (2013). *Far sighted*. Retrieved February 4, 2020, from https://www.dga.org/Craft/DGAQ/All-Articles/1302-Spring-2013/DGA-Interview-Robert-Zemeckis.aspx

Riley, J. (2010). *Brave new world*. Retrieved November 22, 2016, from http://www.backstage.com/interview/brave-new-world/

Roberts, S. (2007). *Cast of Beowulf interview*. Retrieved January 9, 2016, from http://www.moviesonline.ca/movienews_13362.html

Robertson, B. (2016). When characters collide. *3D World, 210*, 36-41.

Royle, N. (2003). *The uncanny: an introduction*. Manchester, UK: Manchester University Press.

Rozik, E. (2002). Acting: the quintessence of theatricality. *SubStance, 31(2&3)*, 110-124. DOI: 10.1353/sub.2002.0039

Russell, J., & Fernández-Dols, J. (1997). What does a facial expression mean? In. J. Russell & J. Fernández-Dols (Eds.), *The psychology of facial expression* (pp. 3-30). Cambridge, UK: Cambridge University Press.

Salter, C. (2010). *Entangled: technology and the transformation of performance*. Cambridge, MA, USA: MIT Press.

Sarkar, S., Phillips, P., Liu, Z., Vega, I., Grother, P., & Bowyer, K. (2005). The humanoid gait challenge problem: Data sets, performance, and analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 27*(2), 162-177. DOI: 10.1109/tpami.2005.39

Schechner, R. (2006). *Performance studies: An introduction* (2nd ed.). Abingdon, UK: Routledge.

Serkis, A. (2003). *Gollum: how we made movie magic*. Boston, MA, USA: Houghton Mifflin.

Serkis, A. (2016). *Empire: from Gollum to TinTin*. Retrieved October 23, 2016, from: http://www.cavemanfilms.co.uk/from-gollum-to-tintin/

Seymour, M. (2018). *Making Thanos face the Avengers*. Accessed March 11, 2019, from: https://www.fxguide.com/featured/making-thanos-face-the-avengers

Shacklock, Z. (2016). Two of a kind: revaluing the work of acting doubles in Orphan Black. *Journal of Film and Video, 68*(3), 69-82.

Sharf, Z. (2019). *Andy Serkis says motion capture should give disabled actors 'great opportunities'*. Retrieved November 16, 2019, from https://www.indiewire.com/2019/09/andy-serkis-motion-capture-disabled-actors-typecasting-1202174645/

Shaw-Williams, H. (2014). *Where does Andy Serkis end & animation begin?* Retrieved July 1, 2016, from http://screenrant.com/andy-serkis-motion-capture-oscars-controversy/

Singer, G. (2003). The two towers: Face to face with Gollum. *Animation World Magazine.*

Silvio, T. (2010). Animation: the new performance? *Journal of Linguistic Anthropology, 20*(2), 422-438. DOI: 10.1111/j.1548-1395.2010.01078.x.

Simi Reality Motion Systems. (2014). *Markerless motion capture an analysis with Simi Shape 3D*. Retrieved January 6, 2020, from http://www.simi.com/en/products/movement-analysis/markerless-motion-capture.html

Sito, T. (2013). *Moving innovation: a history of computer animation*. Cambridge, MA: MIT Press.

Smith. K. (2016). *'Rogue One' proves stars don't have to be alive to still be on screen*. Retrieved from http://nypost.com/2016/12/17/rogue-one-proves-stars-dont-have-to-be-alive-to-still-be-on-the-screen/

Snyder, S. (2012). *The Oscars vs motion capture: Tintin and Apes snubs raise big questions about the Academy*. Retrieved December 1, 2019, from http://entertainment.time.com/2012/02/08/the-oscars-vs-motion-capture-tintin-and-apes-snubs-raise-big-questions-about-the-academy/

Sobchack, V. (2006). Final fantasies: Computer graphic animation and the [dis]illusion of life. In S. Buchan, D. Surman, & P. Ward (Eds.), *Animated 'worlds'*. New Barnet, UK: John Libbey Publishing Ltd.

Sobchack, V. (2009). Animation and automation, or, the incredible effortfulness of being. *Screen, 50*(4), 375-391.

Russell, J., & Fernández-Dols, J. (1997). What does a facial expression mean? In. J. Russell & J. Fernández-Dols (Eds.), *The psychology of facial expression* (pp. 3-30). Cambridge, UK: Cambridge University Press.

Sober, E., & Wilson, D. (1998). *Unto others: The evolution and psychology of unselfish behavior*. Cambridge, MA, USA: Harvard University Press.

Solberg, R., & Jensenius, A. (2016). Optical or inertial? Evaluation of two motion capture systems for studies of dancing to electronic dance music. *Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2016)*, 469-474.

Stanchfield, W. (2013). *Drawn to life: 20 golden years of Disney master classes* (D. Hahn, Ed.). New York, NY, USA: Focal Press.

States, B. (1985). *Great reckonings in little rooms*. London, UK: University of California Press.

Stern, M. (2014). *Motion capture maestro Andy Serkis on 'Dawn of the Planet of the Apes' and Revolutionizing Cinema*. Retrieved December 1, 2019, from https://www.thedailybeast.com/motion-capture-maestro-andy-serkis-on-dawn-of-the-planet-of-the-apes-and-revolutionizing-cinema

Sturman, D. J. (1994). A brief history of motion capture for computer character animation. *SIGGRAPH 94*.

Synthespian Studios. (2015). *About Synthespian Studios*. Retrieved December 18, 2018, from https://web.archive.org/web/20150201014739/http://www.synthespianstudios.net/about/

Taylor, C. (2000). *A painstaking journey*. Retrieved October 7, 2017, from http://content.time.com/time/magazine/article/0,9171,50566,00.html

Tenney, M. (2012). *Microsoft Kinect – hardware*. Retrieved January 6, 2020, from http://gmv.cast.uark.edu/scanning/hardware/microsoft-kinect-resourceshardware/

The New York Times. (2011). *Where the deer and the chameleon play*. Retrieved February 3, 2020, from https://www.nytimes.com/2011/02/27/movies/27rango.html

Thain, A. (2016). In the blink of an eye: Norman McLaren between dance and animation. In D. Rosenberg (Ed.), *The Oxford handbook of screendance studies* (pp. 167-186). New York, NY, USA: Oxford University Press.

Thomas, F., & Johnston, O. (1981). *The illusion of life: Disney animation.* New York: Disney Editions.

Thompson, A. (2008). *What are 'Wall-E's' Oscar chances?* Retrieved December 7, 2019, from https://variety.com/2008/digital/awards/what-are-wall-e-s-oscar-chances-1117989134/

Thompson, K., & Bordwell, D. (2010). *Motion-capturing an Oscar*. Retrieved November 12, 2017, from http://www.davidbordwell.net/blog/2010/02/23/motion-capturing-an-oscar/

Tinwell, A. (2015). *The Uncanny Valley in Games and Animation*. Boca Raton, FL: CRC Press.

Tinwell, A., Grimshaw, M., Nabi, D., & Williams, A. (2011). Facial expression of emotion and perception of the Uncanny Valley in virtual characters. *Computers in Human Behavior, 27*(2), 741-749.

Tisseron, S., Todro, F., and Baddoura, R. (2015). Testing empathy with robots: a model in four dimensions and sixteen items. *International Journal of Social Robotics, 7*, 97-102. DOI: 10.1007/s12369-014-0268-5

Tyler, K. (2000). *Virtual humans*. Retrieved February 11, 2020, from https://www.pbs.org/wgbh/nova/specialfx2/humans.html

Variety Staff. (2008). *'Button' pushes Oscar's digital debate*. Retrieved February 15, 2020, from https://variety.com/2008/music/news/button-pushes-oscar-s-digital-debate-1117997315/

Wells, P. (1998). *Understanding animation*. London, UK: Routledge.

Wells, P. (2006). *The fundamentals of animation*. London, UK: AVA Publishing.

Whitaker, H., and Halas, J. (2002). *Timing for animation.* Burlington, MA, USA: Focal Press.

Whittaker, R. (2018). *Austin Film Critics release 2018 awards nominees lists*. Retrieved December 9, 2019, from https://www.austinchronicle.com/daily/screens/2018-12-28/austin-film-critics-release-2018-awards-nominee-lists/

Wild About Movies. (2019). *Wild about movies*. Retrieved January 22, 2020, from https://www.wildaboutmovies.com/

Williams, R. (2002). *The animator's survival kit: a manual of methods, principles and formulas for classical, computer, games, stop motion and internet animators*, 2nd ed.. London, UK: Faber and Faber.

Wilson, E., and Denis, L. (2018). *Kant and Hume on morality (The Stanford Encyclopedia of Philosophy)*. Retrieved January 28, 2020, from https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=kant-hume-morality

Winchester Hospital. (2019). *Phantom limb syndrome*. Retrieved January 17, 2021, from https://www.winchesterhospital.org/health-library/article?id=96857

Wolf, M. (2003). The technological construction of performance. Convergence: *The International Journal of Research into New Media Technologies, 9*(4), 48-59. DOI:10.1177/135485650300900405

Wood, Aylish. (2014). Behind the scenes: a study of Autodesk Maya. *Animation: An Interdisciplinary Journal, 9*(3), 317-332.

World Encyclopedia of Puppetry Arts. (2018). *Fantoccini*. Retrieved December 20, 2018, from https://wepa.unima.org/en/fantoccini/

Worthen, W. (1984). *The idea of the actor: drama and the ethics of performance*. Princeton, NJ, USA: Princeton University Press.

Zarrilli, P. (2004). Toward a phenomenological model of the actor's embodied modes of experience. *Theatre Journal, 56*(4), 653-666. DOI: 10.1353/tj.2004.0189

Zollhöfer, M., Thies, J., Garrido, P., Bradley, D., Beeler, T., Pérez, P., . . . Theobalt, C. (2018). State of the art on monocular 3d face reconstruction, tracking, and applications. *Computer Graphics Forum, 37*(2), 523-550. DOI: 10.1111/cgf.13382

8.2 Filmography and Image List

Abrams, J. (Director). (2015). *Star wars: episode vii – the force awakens* [Motion picture]. United States: Lucasfilm Ltd.

Animation Mentor. (Producer). (2013, June 25). *Creating weight in four legged walks with Tom Gibbons* [Video file]. Retrieved from https://vimeo.com/ondemand/animationmentor/

Apted, M. (Director). (2010). *The chronicles of Narnia: Voyage of the Dawn Treader* [Motion picture]. United Kingdom: Walden Media.

Bay, M. (Director). (2011). *Transformers: Dark of the moon* [Motion picture]. United States: Hasbro.

Bird, B. (Director). (1999). *The iron giant* [Motion picture]. United States: Warner Bros. Feature Animation.

Bird, B. (Director). (2018). *The Incredibles 2* [Motion picture]. United States: Walt Disney Pictures.

Branagh, K. (Director). (2011). *Thor* [Motion picture]. United States: Marvel Studios.

Burton, T. (Director). (2010). *Alice in Wonderland* [Motion picture]. United States: Walt Disney Studios.

Cameron, J. (Director). (2009). *Avatar* [Motion picture]. United States: Lightstorm Entertainment.

Chow, S., & Kwok, D. (Directors). (2013). *Journey to the west: Conquering the demons* [Motion picture]. China: Bingo Movie.

Columbus, C. (Director). (2002). *Harry Potter and the Chamber of Secrets* [Motion picture]. United States: Warner Bros. Pictures.

Coogler, R. (Director). (2018). *Black panther* [Motion picture]. United States: Marvel Studios.

Cooper, M., & Shoedsack, E. (1933). *King Kong* [Motion picture]. United States: Radio Pictures.

Cuarón, A. (Director). (2013). *Gravity* [Motion picture]. United Kingdom: Warner Bros. Pictures.

DeBlois, D. (Director). (2019). *How to train your dragon: The hidden world* [Motion picture]. United States: DreamWorks Animation.

Derrickson, S. (Director). (2016). *Doctor Strange* [Motion picture]. United States: Marvel Studios.

Disa, M. (Director). (2011). *Hoodwinked too! Hood vs. evil* [Motion picture]. United States:

Kanbar Entertainment.

Docter, P. (Director). (2001). *Monster's, inc*. [Motion picture]. United States: Walt Disney Pictures.

Dudok de Wit, M. (Director). (1994). *The monk and the fish* [Short animation]. France: Patrick Eveno.

Dudok de Wit, M. (Director). (2000). *Father and daughter* [Short animation]. Netherlands: Claire Jennings.

Edwards, G. (Director). (2016). *Rogue one: A star wars story* [Motion picture]. United States: Lucasfilm Ltd.

Faceware Technologies, Inc. (Director). (2013, August 8). *How to use 'neutral frames' in Faceware Analyzer* [Video file]. Retrieved from https://www.youtube.com/watch?v=tANFU_4Ak78

Faceware Technologies, Inc. (Director). (2018, March 19). *Faceware 2018 reel* [Video file]. Retrieved from https://vimeo.com/260855305

Favreau, J. (Director). (2010). *Iron man 2* [Motion picture]. United States: Marvel Studios.

Favreau, J. (Director). (2016). *The jungle book* [Motion picture]. United States: Walt Disney Studios.

Favreau, J. (Creator). (2019-2020). *The mandalorian* [Television series]. United States: Lucasfilm.

Fincher, D. (Director). (2008). *The curious case of Benjamin Button* [Motion picture]. United States: The Kennedy/Marshall Company.

Fischinger, O. (Director). (1938). *An optical poem* [Short animation]. United States: Metro-Goldwyn-Mayer.

Gillett, B. (Director). (1933). *The three little pigs* [Short animation]. United States: Walt Disney Productions.

Gosnell, R. (Director). (2011). *The Smurfs* [Motion picture]. United States: Columbia Pictures.

Jackson, P. (Director). (2001). *The lord of the rings: the fellowship of the ring* [Motion picture]. New Zealand: WingNut Films.

Jackson, P. (Director). (2002). *The lord of the rings: the two towers* [Motion picture]. New Zealand: WingNut Films.

Jackson, P. (Director). (2003). *The lord of the rings: the return of the king* [Motion picture]. New Zealand: WingNut Films.

Jackson, P. (Director). (2005). *King Kong* [Motion picture]. New Zealand: WingNut Films.

Jackson, P. (Director). (2012). *The hobbit: an unexpected journey* [Motion picture]. New Zealand: WingNut Films.

Jackson, P. (Director). (2014). *The hobbit: the battle of the five armies* [Motion picture]. New Zealand: WingNut Films.

Johnson, R. (Director). (2017). *Star wars: episode viii – the last jedi* [Motion picture]. United States: Lucasfilm Ltd.

Jones, C. (Director). (1963). *The dot and the line: a romance in lower mathematics* [Short animation]. United States: MGM Animation/Visual Arts.

Jones, D. (Director). (2016). *Warcraft* [Motion picture]. United States: Legendary Pictures.

Kazan, E. (Director). (1954). *On the waterfront* [Motion picture]. United States: Horizon Pictures.

Kleinman, D. (Director). (2013). *Galaxy chauffeur* [Animated advertisement]. United Kingdom: Framestore.

Kleiser, J. (Director). (2009, August 26). *Don't touch me (1989)* [Video file]. Retrieved from https://www.youtube.com/watch?v=8ovn8qRezPA

Kosinksi, J. (Director). (2010). *Tron: Legacy* [Motion picture]. United States: Walt Disney Pictures.

Lee, A. (Director). (2019). *Gemini man* [Motion picture]. United States: Skydance Media.

Leitch, D. (Director). (2018). *Deadpool 2* [Motion picture]. United States: 20th Century Fox.

Leterrier, L. (Director). (2010). *Clash of the titans* [Motion picture]. United States: Legendary Pictures.

Lévêque, B. (2013, June 15). *VFX breakdown – the social network – Lola Visual Effects* [Video file]. Retrieved from https://www.youtube.com/watch?v=Oitr_eYcJk0

Lisberger, S. (Director). (1982). *Tron* [Motion picture]. United States: Walt Disney Productions.

Lucas, G. (Director). (2002). *Star wars: episode ii – attack of the clones* [Motion picture]. United States: Lucasfilm, Ltd.

Lynch, D. (Director). (1980). *The elephant man* [Motion picture]. United States: Brooksfilms.

Malick, T. (Director). (2011). *The tree of life* [Motion picture]. United States: River Road.

Marey, E. (1884). *Walking man* [Photograph]. Retrieved from https://www.researchgate.net/figure/Etienne-Jules-Marey-Walking-Man-chronophotography-1884-This-was-based-on-adding_fig2_252066643

McNamara, B. (Director). (2011). *L.A. noire* [Video game]. United States: Rockstar Games.

McQueen, S. (Director). (2013). *12 years a slave* [Motion picture]. United States: Regency Enterprises.

Miller, B. (Director). (2011). *Moneyball* [Motion picture]. United States: Columbia Pictures.

Miller, T. (Director). (2016). *Deadpool* [Motion picture]. United States: 20th Century Fox.

Niccol, A. (Director). (2002). *Simone* [Motion picture]. United States: New Line Cinema.

Nightingale, N, & Cook, B. (Directors). (2013). *Walking with dinosaurs* [Motion picture]. United Kingdom: Reliance Entertainment.

Nixey, T. (Director). (2010). *Don't be afraid of the dark* [Motion picture]. United States: Necropia.

Nozue, T. (Director). (2016). *Kingsglaive: Final fantasy xv* [Motion picture]. Japan: Visual Works.

O'Connell, J. (2014, November 16). *Gollum wins an award …* [Video file]. Retrieved from https://www.youtube.com/watch?v=lRn_FlDKITU

Pau, P., & Tianyu, Z. (Directors). (2015). *Zhong Kui: Snow girl and the dark crystal* [Motion picture]. China: Village Roadshow.

Reeves, M. (Director). (2014). *Dawn of the planet of the apes* [Motion picture]. United States: Chernin Entertainment.

Reeves, M. (Director). (2017). *War for the planet of the apes* [Motion picture]. United States: Chernin Entertainment.

Rodriguez, R. (Director). (2019). *Alita: Battle angel* [Motion picture]. United States: 20th Century Fox.

Roshan, R. (Director). (2013). *Krrish 3* [Motion picture]. India: Rakesh Roshan.

Ross, G. (Director). (2012). *The hunger games* [Motion picture]. United States: Color Force.

Russo, A., & Russo, J. (Directors). (2016). *Captain America: civil war* [Motion picture]. United States: Marvel Studios.

Russo, A., & Russo, J. (Directors). (2018). *Avengers: Infinity War* [Motion picture]. United States: Marvel Studios.

Sakaguchi, H., & Sakakibara, M. (Directors). (2001). *Final fantasy: the spirits within* [Motion picture]. United States: Square Pictures.

Sanders, R. (Director). (2012). *Snow White and the huntsman* [Motion picture]. United States: Roth Films.

Scorsese, M. (Director). (2019). *The Irishman* [Motion picture]. United States: TriBeCa Productions.

Serkis, A. (Director). (2018). *Mowgli: legend of the jungle* [Motion picture]. United States: Warner Bros. Pictures.

Sharma, M. (Director). (2016). *Fan* [Motion picture]. India: Yash Raj Films.

Shankar, S. (Director). (2010). *Enthiran* [Motion picture]. India: Sun Pictures.

Shyamalan, M. (Director). (2010). *The last airbender* [Motion picture]. United States: Nickelodeon Movies.

Spielberg, S. (Director). (2011). *The adventures of Tintin: the secret of the unicorn* [Motion picture]. United States: Columbia Pictures.

Spielberg, S. (Director). (2015). *Bridge of spies* [Motion picture]. United States: DreamWorks Pictures.

Spielberg, S. (Director). (2016). *The BFG* [Motion picture]. United States: Walt Disney Pictures.

Spielberg, S. (Director). (2018). *Ready player one* [Motion picture]. United States: Warner Bros. Pictures.

Taylor, A. (Director). (2013). *Thor: The dark world* [Motion picture]. United States: Marvel Studios.

Trevorrow, C. (Director). (2015). *Jurassic World* [Motion picture]. United States: Legendary Pictures.

Verbinski, G. (Director). (2011). *Rango* [Motion picture]. United States: Paramount Pictures.

Verbinski, G. (Director). (2013). *The lone ranger* [Motion picture]. United States: Walt Disney Pictures.

Vogt-Roberts, J. (Director). (2017). *Kong: skull island* [Motion picture]. United States: Legendary Pictures.

Wan, J. (2015). (Director). *Furious 7* [Motion picture]. United States: Original Film.

Webb, M. (Director). (2012). *The amazing Spider-Man* [Motion picture]. United States: Columbia Pictures.

Wedge, C. (Director). (2005). *Robots* [Motion picture]. United States: Blue Sky Studios.

Whedon, J. (Director). (2015). *Avengers: age of Ultron* [Motion picture]. United States: Marvel Studios.

Wired. (February 19, 2016). *It took 5 actors to create Deadpool's cgi Colossus |design fx |*

WIRED [Video file]. Retrieved from https://www.youtube.com/watch?v=LncTFEnxRV8

Wyatt, R. (Director). (2011). *Rise of the planet of the apes* [Motion picture]. United States: Chernin Entertainment.

Zemeckis, R. (Director). (2004). *The Polar Express* [Motion picture]. United States: ImageMovers.

Zemeckis, R. (Director). (2007). *Beowulf* [Motion picture]. United States: ImageMovers.

Zwick, E. (Director). (2006). *Blood diamond* [Motion picture]. United States: Bedford Falls Productions.

8.3 *Definitions of Key Terms*

Actor's Double:	The phenomenological experience of the actor's body as a virtual representation (p. 191)
Animation:	Animation is a form of performance by proxy that extends and displaces the body of the performer onto a series of mediated, moving images of a manipulated, artificial construct set at positions with sufficient differences from each other to produce an intervallic illusion of motion (p. 36)
Authenticity:	the degree to which an actor represents a character to him/herself (p. 127)
Basic Emotions:	pan-cultural discrete emotions from which all other emotions are subcategories or mixtures (p. 147)
Behaviour:	the resulting action a person makes based on the experience of a particular feeling (p. 146)
Capture Volume:	the physical space in which the MoCap performance can be recorded by the system (p. 53)
Complex Acting:	acting in which multiple degrees of pretence are used (p. 212)
Cyberstar:	a digital performer that can wholesale replace carbon-based actors (p. 70)
Digital Counterpart	a CG or digitally-augmented character that is capable of producing a screen-based performance and whose performance is derived, at least in part, by a source-actor via motion capture (p. 28)
Digital Doppelgänger:	a digital double that is an exact digital recreation of the actor on which it is based (p. 173)
Digital Double:	a CG character within a profilmic context whose participation in a screen-based story is implied more by its physiology, costume, and surroundings than by its actions (p. 172)

Digital Fantoccini:	a digital double whose performance is solely derived through keyframe animation (p. 173)
Emotion:	a distinct and dynamic functional state of an entire organism, comprising particular groups of effector systems (visceral, endocrine, muscular) and particular subjective states (feelings) (p. 146)
Emotionally Connected	the degree to which an actor truly represents a character to him/herself while truly representing the character to an audience (p. 128)
Empathy:	an emotional reaction to the situation of others which involves feeling the emotions they feel (p. 162)
Ex-Filmic:	a reality, situation, or element solely recorded through software-based cameras (p. 175)
Expression:	a universally understood signal of the triggering of a discrete categorical basic emotion (p. 151)
Facial Chirality:	an asymmetrical expression, suggesting a simultaneously display of multiple emotions (p. 158)
Feeling:	the conscious registration of a new body-state; emotion made conscious (p. 146)
Honour:	to retain the intrinsic aspects and quality of a source performance in the performance of a digital counterpart (p. 351)
Micro Expression:	a rapid facial expression that lasts between 1/25 and 1/15 of a second; micro expressions occur due to emotional repression or concealment (p. 155)
Motion Capture (MoCap):	the process of digitally recording movement in either 2D or 3D physical space and translating that movement onto a model/object in screen-based or physical 3D (p. 49)
Neutral Frame:	a video frame in which every region of an actor's face is expressionless (p. 228)

Oral Commissures:	the edges of the mouth where the top lip transitions into the bottom lip (p. 224)
Performance Capture (PeCap):	a subset of motion capture that includes the total recording of a performance without cuts. Performance capture generally refers to the capture of facial performance in addition to body performance (within this research, PeCap specifically refers to a system that captures both body movements and facial performance) (p. 50)
Practice-Led:	a practice-based methodology in which the research leads primarily to new understandings about practice (p. 186)
Presence:	an actor's unique, personal ability to produce charisma, chemistry, and/or appeal, especially within service to a given role (p. 166)
Profilmic:	a reality, situation, or element recorded through live-action cinematography (p. 175)
Protean Double:	a digital double that does not resemble the actor on which it is based (p. 173)
Realist Acting:	acting that privileges mimesis above other forms of representation (p. 91)
Realistic Animated Character:	the degree to which a character is physically believable and human-like; if the character is not human, the realism refers to how closly the creature resembles its own kind (or a recognisable terrestrial analogue) (p. 138)
Retargeting:	the process of translating a statistical model of facial performance data onto the animation controls of a 3D character rig (p. 51)
Simple Acting:	acting in which only one degree of pretence is used (p. 90)
Sincerity:	the degree to which an actor represents a character to the audience (p. 127)
Sympathy:	an emotional reaction to the situation of others which does not involve feeling the emotions they feel (p. 162)

Synthespian:	a CG or digitally-augmented character within a profilmic context that advances a screen-based story while producing acting that engages in at least one area of pretence (p. 172)
Truthfulness:	the degree to which an actor represents a character consistent with the needs of a story (p. 132)
UV Layout:	a two-dimensional representation of the surface of a 3D model, primarily used as the basis for creating 2D textures to project onto a 3D object (p. 352)
Vactor:	a CG or digitally-augmented character within a profilmic context that is capable of producing a screen-based performance (p. 172)
Vermilion Border:	the outer edge of the lips (p. 223)
Vermilion Mucosal Junction:	the zone of the inner lip where the fleshy outer part of the lips transforms into the wet, mucosal region inside the mouth (p. 224)

Appendices

Appendix I. Analysis of Vactor Acting Styles

This thesis highlights acting and performance considerations specific to motion capture. As part of this account, I identify and define a range of vactor acting styles (subsection 4.4.2). This research is the first attempt to produce this type of robust digital performance taxonomy, and provides a clear delineation of performance styles within motion capture that are irrespective of visual style.

The creation of this performance taxonomy began as a response to questioning how many vactors have appeared in feature films since "Avatar" (December 2009). For me, "Avatar" was a critical moment in the history of film-based vactors due to the unprecedented performance realism of the main Na'vi characters. In the years that followed, films incorporated far more digital characters than in previous years, but a full survey of these characters, including their different methods of performance, had never been conducted. Instead, academic research prioritised already well-documented historical examples – such as in "Final Fantasy: The Spirits Within", "The Lord of the Rings", "The Polar Express", and "Beowulf" – leaving a void of research into an emerging range of meaningful digital character performances and performance styles.

I specifically focus on feature films in order to provide a manageable scope of inquiry. Although fully investigating a worldwide oeuvre of feature films each year is a monumental endeavour, it is still more manageable than performing a thorough yearly analysis of all episodes of all television shows in even a single country. This research surveys primarily

Western feature films (American, Canadian, and Australian), with an attempt to pay service to Bollywood, Chinese, Korean, and European cinema. While vactors are lead characters in some films, many other vactors are supporting characters, often without speaking parts, who may exist for just a few frames in the entire film.

It is unreasonable to expect to watch every feature film released since 2010 in its entirety. To accommodate this, I employ a method to determine whether or not a film is likely to contain vactors. The website "Wild About Movies" (2019) catalogues (with movie trailers, posters, and written summaries[85]) every Western feature film released within a given year. From this list, I am able to exclude certain films based on the lack of a "perceptual likelihood" that those films would utilise vactors as part of their aesthetic or production concerns. This likelihood is based on the visual, genre, and story content derived from the film's trailers and posters. For instance, romantic comedies (romcoms) and indie films are less likely to contain digital characters than are science-fiction, fantasy, and horror films. Certain production needs also warrant the use of vactors, such as films that feature talking animals, films with photorealistic robots or aliens, and films that rely on posthumous performances (Figure 39).

Since this is not a fool-proof assessment, I use the "Internet Movie Database (IMDb)" (2020) to cross-examine the film budget of each movie to determine the "financial likelihood" that a film's production could support the expensive cost of producing vactors. In the case of non-Western films, I consult a variety of online sources to determine the most expensive films produced in major global film centres (such as India, China, and Nigeria) each year, and cross-reference these with the types of roles in each movie. While the majority of films with vactors catalogued by this research have a budget of USD $15 million or greater, a few lower budget films include vactors, but often in much more limited scope. I use USD $500,000 as an absolute minimum budget for a film to include

---

[85] I feel it is important to acknowledge that the written summaries on Wild About Movies often express politically problematic viewpoints. While this is a reason to second guess the appropriateness of this resource, I find that its system of reliably cataloguing feature films by date and year, along with other attendant materials, to facilitate the needs of this research.

Common Vactor Roles

- Aged/de-aged versions of living actors
- Aliens
- Animals enacting specific directions and/or talking
- Animals or groups of animals that would be otherwise difficult to film
- Animals performing stunts or interacting with humans
- Characters engaging with magic
- Characters performing complex or dangerous choreography
- Characters with physical modifications
- Fantasy and horror creatures
- Posthumous performances
- Robots and characters with robotic/mechanical modifications
- Stylised characters within a profilmic context
- Superhero characters

Figure 39. Types of characters in feature films most likely to be portrayed using a vactor.

vactors. As a film budget approaches this minimum, the likelihood that the film includes vactors approaches zero. If I suspect a film may contain vactors, I search Google for discussions and visual breakdowns about the film's VFX production. I also refer to the VFX industry publications *Cinefex* and *3D World*, which discusses in detail the major VFX film productions from each year.

Using the method outlined above, I compile a list of every movie likely to contain vactors. I watch each film using the MPC-HC media player for Windows. MPC-HC enables me to advance ten seconds forward or backward at the tap of an arrow key, which allows me to quickly advance through a film while identifying key moments that may contain vactors, at which point I watch the film at normal speed.

Once I have identified and recorded all of the vactors from a given movie, I compare these roles against other vactor roles from the same year. As this study is concerned with modes of performance, I specifically focus on the nature of each character's performance rather than its appearance. Through cross-analysing observations about vactor performance, I develop a list of vactor acting categories and I assign each film character to one or more acting categories (it is possible that over the course of a film, a vactor participates in more than one of these acting categories).

Appendix II. Detailed Technological Overview of BOE

II.i Analyzer Technological Overview

Several key poses are created for Analyzer to use as a template for tracking the markers on the frames in-between the keys. However, this tracking is never accurate enough on the first pass and requires significant manual adjustment of in-between poses over many iterations. In addition to ensuring that the markers match the correct position of their associated landmarks, it is essential that the markers follow their landmarks at the same speed as the source footage. That is, while the tracking may be accurate overall, during quick transitions (such as the eyes suddenly shifting from looking forward to looking left), the markers may follow a few frames late, even if they eventually align in the correct position with the landmarks. It is vital that the timing of the tracking perfectly matches the source performance in order preserve the integrity of expressions and to honour (that is, to retain the quality of) the source performance as much as possible. This reduces the risk of uncanniness emerging from a performance that is even a few frames out of sync between the actions of different parts of the face and/or body. It is important to keep the markers as stable as possible because any pronounced jitter in the Analyzer performance will likely result in jitter when the tracking data is retargeted onto the character rig in Maya. Likewise, it is important to be as accurate as possible when tracking the facial features so as to create a library of facial shapes that provide meaningful and consistent retargeting. If the facial shapes are not accurately tracked, this creates confusion for Analyzer and Retargeter in terms of how the shape of a face group corresponds to particular values associated with the retargeted controls in Maya. Basically, the more accurate the tracking is in Analyzer, the more accurate the retargeting should be in Maya.

II.ii Retargeter Technological Overview

The Retargeter workflow begins with calculating AutoPoses – that is, frames from the reference video that Retargeter determines to be the most significant in terms of encompassing the full performance range of a given face group within the source video. The general workflow involves generating two or three AutoPoses, creating equivalent poses on the 3D model, and testing how well those additional shapes contribute to accurate retargeting. This final step creates a relationship between the 3D facial poses and the layout of the facial markers on the corresponding frame within Analyzer. It does this for every frame on which I define a facial pose with Retargeter, analyses this data to produce a tracking model, and extrapolates the most likely facial pose for each frame of the animation. With only a few facial poses, the quality of its extrapolation is always inaccurate and misses both large-scale and small-scale details. However, as additional AutoPoses are added, the overall retargeting result improves. Eventually, rather than relying on the software to define AutoPoses, it is more meaningful to heuristically determine which poses I believe are most likely to improve the overall retargeting. Retargeter allows user-defined poses to be defined alongside AutoPoses, and both of these pose-types are used to evaluate the retargeting process.

II.iii UV Layout

Because the original head model (and the blend shapes derived from it) were created without a body attached to it, the original head model had a different UV layout[86] from the more recent version of the head that is attached to the body model (the full body model). The more detailed texture was created for just the original head model used in the BOE.

II.iv Facial Rigging Configuration

I developed a rigging configuration so that any blend shape adjustments made on the original head model were copied onto the full body model via a wrap deformer. Then, the SWB controls of the full body model were animated on top of the full body model to

[86] A UV layout (or a UV map) is a two-dimensional representation of the surface of 3D model, and is primarily used as the basis for creating 2D textures to project onto a 3D object.

produce the final performance. As a result, two heads sat on top of each other and it was important to only see the desired one at any given time. I created a transparency control to turn on and off the visibility of the each of the models, which meant that the more detailed head texture was only visible when the original head model was also visible. The drawback here was that I could only see the detailed texture when adjusting the blend shapes, but not when refining the shapes using the SWB controls. More time spent recreating the textures from scratch for the head and body model would have led to easier animation, but it would have cost time spent texturing rather than animating. At that stage, I decided to push forward with animation instead.

II.v Facial Posing Considerations

Any differences between the source video and the positions of the mouth could be difficult to control on the rig if a corresponding blend shape had not been created to precisely recreate that mouth shape. In a situation where the lips began in a narrow shape, became wide, then returned to narrow, but were ever so slightly wider than the original narrow shape, it was important that the second narrow pose I created was accurate to this difference and did not, for instance, become narrower than the original shape. These types of proportional adjustments and calibrations took place throughout every region in the face to ensure there was consistency across all of the Faceware poses. However, due to the degree of subtlety between each of these poses, it was unclear whether Faceware Retargeter would be able to produce consistent animation data at that level of nuance, or whether the result would become arbitrarily noisy.

The Analyzer data then passes through to Retargeter, which imports the same video feed used in Analyzer into Maya in order to sync the two systems. Through Retargeter, the user chooses key facial shapes from the source video based on only one of the three face groups at a time. For instance, the mouth region may have a high degree of variance during the same time at which the brows barely move. The user selects key facial shapes for just the mouth region in a single pass, which is then followed by selecting key shapes for the eyes, and then the brows. It is rare for the key shapes on all three facial regions to occur on a single frame. Each facial shape should vary from the other facial shapes within the same facial group in some meaningful way. That is, there should not be a repetition of facial shapes that are too similar, as this may potentially confuse the system. I create between 15-20 facial shapes per facial group as a baseline for capturing the full range of performance within a sequence.

In the end, I produced 19 Faceware poses for the mouth, many of which were quite similar to each other. However, I was unsure of how much variance the Faceware system would be able to detect, and I decided to err on the side of more poses rather than fewer. For the eyes, I produced 18 Faceware poses, and just 3 poses for the brows (due to their minimal movement throughout the performance). The low angle of the Faceware camera made the degree to which the eyelids should open in any given pose less obvious, requiring an element of educated guesswork when constructing these poses. In an ideal setup, I wouldn't have to rely on artistic intuition as much. I also added a colour texture to the eyeballs to assist with posing them. Ensuring that the iris and the pupil were true-to-life-size proved important when judging the degree of openness of the eyelids.

II.vi Eye Skin Weight Transitions

Ensuring that there were smooth transitions between the skin weights of neighbouring SWB controls for the eyes was challenging but essential to creating believable shape adjustments without pronounced irregularities. This is a problem that possibly could have been mitigated by using more markers around the eyelids, or by creating two rings of markers – one directly along the eyelid edge, and one around the centre of the eyelids. I implemented this second technique for the upper eyelids and would have likely had benefits if I implemented the same technique for the lower eyelids.

Appendix III. Overview of the Dramatic Performance

III.i Analyzer Technological Overview

Throughout the dramatic performance, there are numerous subtle eye twitches that are a product of me acting in a state of anxiety. However, these types of movements could easily be mistaken for noise in the data and must not be eliminated in the clean-up process. The basic tracking in Analyzer struggles to accurately capture these subtle movements and often has to be guided on a frame-by-frame basis by the MoCap artist to fully track this movement. These eye twitches are very minor movements that may strongly affect how the emotion and the intention of this moment is understood. The variations in position of these movements are so small that they fall within the normal range of error for a tracking point; within the software's calculations, these variations are indistinguishable from noise. This could potentially be solved by having higher resolution video to track from.

For an unknown reason, whenever I tracked the face groups for the dramatic performance in Analyzer, the mouth group produced significant errors if it was combined with the tracking for the eyes and brows. As a result, I had to create two separate files: one for the eyes and brows, and one for just the mouth. I then combined the tracking results from these two files via Retargeter. Fortunately, this does not produce inconsistencies in the performance. Because the mouth tracking and the eyes/brows tracking were in separate files, I had to import and perform the tracking in these regions separately. I did not necessarily pose the full face on any particular frame. Instead, I found the frames most relevant to posing all of the mouth shapes, and I did separate passes using (usually) different frames to create the eyes and brows poses, respectively.

Rotating the SWB-controls for the mouth was essential when creating the correct lip deformations – whether it involved flattening or rounding them out more. A particularly problematic region to perfect was the shape of the *oral commissures*; the SWB-controls were critical to achieving the correct mouth shapes in this region. Likewise, the SWB-controls are important when creating squash and stretch in the cheek region, especially when the lips widen or move up and down. While the SWB-controls help with the shape of the lips, I still needed to create some custom blend shapes for particular lip shapes that I could not achieve with the SWB controls alone. These shapes included flattening the upper lip, rounding out the lower lip, and squaring the lower lip.

Having a neutral frame to work from becomes especially helpful when creating the facial poses as it provides a guaranteed "rest position" which can always be referred to as the basis against which subsequent proportional adjustments are evaluated. When creating the mouth shapes, especially if the mouth was open, I started by opening the jaw in order to create the correct distance between the upper and lower teeth. Then I matched the position of the lips by proportion to the location of the teeth. When the lips were only slightly parted, this required some adjustments to the corners of the lips called "sticky lips" to represent the natural adhesion between lips.

When posing the eyes, an unexpected discovery was that it was useful to have eyelashes in order to create a better representation of the true shapes of the eyes. Without the eyelashes, it was difficult to tell if the bald shapes of the eyelids were correct. This was because the curl of the eyelashes in the video occludes the underlying true shape of the eyelids. However, the standard method for creating and animating eyelashes (involving geometry and a series of blend shapes to follow the blend shaped-based movements of the eyes) was insufficient for my rig because the eyes are also controlled by the SWB-controls. In order to accommodate both blend shape and SWB adjustments, I had to create a much more complicated setup that involved a series of joints around the inner eyelids that are parent-constrained to nHair follicles in Maya. The joints inherit their motion from these follicles, which are attached to the eyelids via a wrap deformer. The eyelashes are attached to the joints via skin weights, and therefore inherit the motion of both the blend shapes and SWB-controls.[87] It is worth noting how the addition of the eyelashes improves the overall readability of expressions, as well as

[87] A full overview of this technical process is available here: https://www.youtube.com/watch?v=SeM792n6TpE

making the expressions feel more believable. Eyelashes were not used in the original BOE performance, so it is possible that had they been, I may have adjusted the shapes of the eyelids differently for various poses.

When converting the face performance data from 60fps to 24fps, the keys adjusted to non-integer frames. This is problematic from an animation perspective and also means that there were 2.5 keys per frame on average. In order to reduce the keys to discrete integer frame intervals, I needed a solution that would preserve the underlying shape of the animation curves. Maya's default solution for this is the Simplify Curve command, but this does not provide intuitive controls or predictable results, nor does it necessarily create keys on integer frames. Likewise, the resulting values of the simplified curve did not closely follow the original values to a point where I could trust the automated solution to accurately preserve the performance data. As an alternative to Simplify Curve, I found a third-party plugin called the Red9 Studio Pack for Maya by Mark Jackson[88]. The Red9 Interactive Simplify tool provides an intuitive and interactive workflow for altering the spread of keys along a curve (including snapping to integer frame numbers) while preserving the shape of the original animation curve. It proved to be vastly better than Maya's native solution.

I initially speculated that the number of keys created in Analyzer may have been counterproductive and contributed to the pronounced jitter after retargeting. In order to address the possibility of the Analyzer data being poorly affected by too many keys, I went back to the original data and reduced the density of keyframes in Analyzer. This was reinforced when I looked at the example Analyzer files available on the Faceware website, which contain far fewer keyframes than I was using. I reasoned that these demo files should represent good quality tracking from fewer keys. By reducing the number of keyframes, this reduced the precision of my tracking *within Analyzer*, but my hope was that it would improve the overall quality of the retargeting *within Retargeter*. However, the retargeted result from fewer Analyzer keyframes was virtually identical to the denser data – meaning

[88] The Red9 Studio Pack for Maya is available from:
http://red9consultancy.com/production-tools/red9-studio-pack/

that the eye data remained effectively useless whenever the eyes looked down. One solution was to delete the data and keyframe-animate the eyelids, but that would undermine the entire point of PeCap[89]. I noted that there was overall quite consistent lighting in the Faceware demo performances, whereas the lighting in the AUT MoCap studio varied; it is possible this has an effect on the quality of the tracking and retargeting. One other issue I considered was whether there were inconsistencies in the poses as the character looks down, which could potentially confuse the system. I removed four poses that were nearly duplicates of other poses – for instance, the amount that the eyeballs moved between two similar poses may have been so insignificant that the small variations "confused" the software. By limiting the number of training poses where the eyeballs look down, and especially limiting the number of similar poses in this regard, I hoped to produce retargeting with less noise. However, this resulted in only slightly better results without providing enough of a fix to feel like a good solution was achieved.

Another potential reason for why the data was problematic could be that I tracked sections that were too long, and it would be better to divide up the tracking into short intervals. This is something worth exploring in the future.

III.ii Retargeter Technological Overview

The amount of noise was significantly less in the mouth region. However, the shape of the mouth in the training frames (e.g. the original poses) was more accurate to the source video than the subsequent poses that resulted from the retargeting, even on the same frames as the training frames. This reminded me of Fincher's quote that performance capture "sandblasts the edges off a performance" (subsection 6.3.1). While the individual pose controls are *near* the correct values, at any given frame they are not necessarily at the correct values. For instance, sometime the jaw doesn't stay open long enough, or even open wide enough, or

[89] Other Faceware animators who I consulted for technical feedback during this project indicated they have run into similar problems and had to delete portions of their data, achieving better results from keyframe animation instead.

it may open a few frames too late. I suspect that this is as a result of there being a limited number of training frames from which the system attempts to create a full performance model. I also noticed that any small "bias" I inadvertently introduced into the training frames would propagate through the whole of the animation. For instance, the left side of the mouth may have been slightly more open in the training frames due to the selected facial shapes, and this propagated through the entire animation such that the left side of the mouth was more open than the right side of the mouth, even when the lips were meant to be fully closed.

The base model did not contain any eyebrow details and instead relied on a colour texture to suggest these. However, the final animation does not use the colour texture and I decided that it would be better to include physical eyebrows. I had originally conceived of adding eyebrows via a hair system, but I elected to simply model them directly into the mesh in ZBrush. This provides enough geometry detail to produce a suggestion of eyebrows in the viewport, the detail of which can be increased when normal/displacement maps are applied at render time. The addition of the eyebrows helps to better read and evaluate the expressions in the forehead region.

III.iii PeCap Artist Assistant Reflection

My PeCap assistant believed the body clean-up involved a less demanding thought process compared to the facial process. She felt she could work for longer periods of time with less distraction when working on the body PeCap, especially when listening to music or playing a video in the background. It also helped that the body PeCap was able to play back in real-time. She expressed anxiety over deleting data and the effect it could have on altering the positions of body parts. For instance, she noted that the positions of the arms were different after clean-up than in the original data. (This may also account for why the arms interpenetrated the body so much in the cleaned data.)

III.iv PeCap Body Clean-Up

I used animation layers to reposition the forearms to correctly rest against the legs and belly. The left arm needed to fully rest on the left leg over the majority of the animation, and it needed to inherit motion from the leg as the leg moved. In order to make it feel like the leg was driving the motion of the arm, I set up a system that parent-constrained the wrist control of the left arm to the left leg joint at the location where the hand rests on the leg.

I added a 3D model of a stool to the scene to accommodate the sitting position of the character. A core difference between the 3D model and the living actor is that the real-life body is naturally fleshy. That is, when the legs sit on the stool, they compress and squash outwards. Likewise, when the hand rests on the leg and against the belly, it creates a sense of indention into these parts of the body. Such deformations do not naturally occur on a 3D model without complex simulations so I created blend shapes to produce these deformations.

The Nansense suit used to record the body performance resulted in several errors common to inertial MoCap systems. For instance, sliding motions occurred in the hips and left foot, and some rocking motions in the hips were over-emphasised by comparison to the original performance. Any issues toward the centre of the character (especially at the hips) become much more noticeable at the extreme ends of the character, including the feet, hands, and head.

Appendix IV. Approach to Modelling the 3D Character

IV.i Introduction

During the second semester of the 2016 school year, I taught the "Motion Capture Live" paper at Auckland University of Technology for the first time. This paper is a part of the motion capture minor at AUT, and this paper had been taught only once before with different lecturers. I was brought on to teach the paper with greater regard for facial performance and animation, specifically. This would be achieved through the use of the Faceware performance capture system. I created a brief for students to model a photorealistic version of their individual heads using Pixologic ZBrush as the 3D sculpting software. Students were to also model a series of blend shapes based on FACS, and to assemble the model and blend shapes into a rig. They would then use their facial performance capture data to drive the animation of the rig. In order to assist the students in achieving such an ambitious project, I also participated in all aspects of the project. I recorded video tutorials of my work process and made these tutorials available on YouTube for my students to follow along.

By the completion of the paper in November 2016, I had a photorealistic model of my head, including high-resolution details, textures, and the minimum number of blend shapes that could support full-face animation. I assembled all of these facets into a functioning rig and applied some test facial performance capture data to it. Throughout this section, I will explain how I approached the modelling for my digital counterpart, including my design and performance considerations along the way.

IV.ii Reference Gathering

Prior to modelling, it is necessary to gather high-quality photographic reference. I used the Digital Design Department's green screen room as the basis for the photoshoot. Instead of shooting against the green back drop, I used one of the room's black walls, as this produced good contrast between me and the background, clearly defining the shape of my head without introducing any colour spill onto my skin (as can be the case when shooting against a coloured back drop).

The lighting was achieved with just two Mettle VL-650 LED panel lights, which produce diffuse lighting. The intensity of the light can be adjusted through an onboard dimmer. These lights were placed at approximately 45-degree angles to the subject and 1.5m away from my head. There were no other sources of light in this room, as it is windowless.

Photos were composed to include the base of my neck to the top of my hair. The first photo was taken with me directly facing the camera with a neutral expression. For each subsequent photo, the I turned my whole body by 45 degrees while maintaining the same position on the floor. This continued until a 360-degree circuit was completed. Closer photos were also taken of my eyes looking up, down, left, and right, as well as my mouth open wide, lips curled up, and my head rotated back and down.

In 2018, after it was decided to extend my head model into a full-body model, this process was repeated to obtain photos of my entire body. I was photographed participating in T-poses and A-poses by 45-degree intervals, as in the previous photo series.

IV.iii Head Modelling Process

The head model was created in ZBrush, a digital 3D sculpting software. I began with a low-resolution model of the head (4,388 polygons) with good animation topology. In ZBrush, I gradually refined the details of the model through seven subdivision layers to a total of 18 million polygons. At this resolution, I created high-frequency details such as skin pores and wrinkles. I imported the Level 3 subdivision (70,208 polygons) of the head model into Maya and used it as the basis for rigging. I saved all of the differences in the details between Level 3 and Level 7 subdivisions in normal maps and displacement maps.

Since the head was originally created without a body, later incorporating the head into a full body model created some difficulty. This was due to the fact that all of the facial blend shapes were linked to the existing topology of the head model alone and would not function properly when the topology was changed by adding the body. This is an issue I could have avoided if I had known from the start that I would create a full body character of myself, but this was a decision that was made in the course of the research. In order to maintain the existing head model and associated blend shapes, I create a version of the full body with a head identical to the solo head model. In Maya, I applied a wrap deformer between the solo head model and the body so that any deformations in the solo head resulted in identical deformations on the full body model. This was a workable solution but it is expensive to calculate and the chief reason why the animation rig does not produce real-time feedback.

All blend shapes for the head were created in ZBrush as separate layers. It was possible to create left and right versions of each blend shape by masking the head down the midline and applying a morph target that was based on the neutral expression. The mask could then be inverted and this process repeated to produce the equivalent blend shape for the opposite side.

IV.iv Facial Scan

While I had originally modelled my 3D head from photo-based references, I was aware that inevitably mistakes and approximations result from this method. In February 2018 I became aware of a 3D scanner at the BioDesign Lab at AUT. David White, the director of the lab and a senior lecturer in the Department of Mechanical Engineering, invited me to test the Creaform GoSCAN 50 3D scanner for my project. Together we used the optical scanner to create a 3D model of my head. This process took about 15 seconds to complete and involved me sitting still in a chair while the handheld scanner orbited around my head. The result of this scan was a triangulated 3D model in OBJ format, including colour textures of my head. I briefly considered why, if this technology existed at AUT, I did not just scan my head in the first place instead of modelling it from scratch. However, a quick look at the detail of the scan revealed that the geometry was too low-resolution to be useful without significant extra work. While the scan provides an overall accurate representation of the size and proportions of my head features, the geometry is replete with holes and poor topology. The utility of such a scan is in helping to guide the shapes and proportions of a 3D model created from scratch.

I imported the 3D scan geometry into ZBrush and matched it alongside the head geometry I modelled. It was immediately obvious that there were significant differences between the two models. I attempted to adapt my sculpted model to fit the shapes of the scan, but the resulting form of the sculpted model looked fat and bloated compared to the scan data. I attributed this to the fact that there were many minor details in the scanned model that I did not precisely translate into my updated sculpt. However, I generated a result that was very close to the shapes and proportions of the scan data if I reduced the influence of my updated sculpt by 50-percent.

IV.v Photogrammetry

Later in 2018, I learned of a suite of photogrammetry cameras that AUT had acquired and started testing. I worked with Lee Jackson in the AUT Motion Capture Lab to produce a series of images of my head that could be interpreted by a photogrammetry software (RealityCapture) and constructed into an accurate 3D representation with greater fidelity and detail than the 3D scanning technology. Photogrammetry is normally conducted from an array of cameras positioned around a subject. The cameras are at different heights and angles in order to produce a full 360-degree capture of all of the surface details. The cameras are all connected to hardware that triggers each camera to take a photo in rapid succession – the result being a nearly instantaneous capture from every angle. However, generally at least 20 cameras are required to perform this operation, whereas AUT possessed only ten Canon 700D DSLR cameras. This meant that the cameras could not be arranged to fully encircle the subject, and instead were positioned solely from a front angle. In order to produce full 360-degree coverage, the subject is seated on a rotatable chair. After the front angle is recorded, the subject rotates 45-degrees. A makeshift plumb-bob is used to ensure the centre of the subject's head remains consistent after each rotation. Photos are taken from each of these angles until a full circle is completed, resulting in a total of 80 photos. The subject is photographed in a neutral expression, although I found it useful from modelling perspective for the eyelids to be closed. Subsequently, expressions can also be recorded via photogrammetry; however, using this method, we decided to only record the front angle for expressions. This is because it is impossible for the subject to maintain a consistent expression throughout the amount of time required to rotate the subject through each 45-degree angle. The subject's expression must be consistent in order for the photogrammetry to accurately function. I operated under the assumption that once I constructed a good neutral expression model from the full 360-degree photogrammetry, I only needed to focus on the major details in the front of the face for expression detail.

Despite our efforts, the chair rotation method for capturing the full head produced numerous problems when RealityCapture attempted to compute a 3D model from the photos. However, the model it produced was of sufficient quality and detail to provide meaningful reference for my existing 3D head model. By importing the scan data into ZBrush, I was able to use it as a guide for how to adjust the shape and proportions of my head sculpt. The result was a much more accurate 3D representation of myself.

Appendix V. Approach to Rigging the 3D Character

V.i Overview of Rigging Process

The rigging process for this project was complicated by similar issues that I faced when modelling: namely, I originally planned for the model and rig to encompass just my head, and as the project evolved, I needed to account for the entire body. The original rig that I created for the head proved insufficient because of the rigging method I required for the body, as explained below.

I adapted the original head rig after the Victor rig, which is provided for free by Faceware Technologies and is used as an example in many of their tutorials. My choice to develop a similar rig was based on the content that I wanted to teach in the Motion Capture Live paper, for which my head model and rig was used as a guiding example. I felt that the Victor rig provided a wide range of facial controls and blend shapes, thereby providing a solid template onto which performance capture data could be accurately retargeted. I wanted my students to learn both modelling and rigging fundamentals from the perspective of animation performance. Since Faceware Technologies relied on the Victor rig, I took it as a cue for what would work well in terms of rigging for performance capture – a specific line of inquiry with which I had no previous experience.

From a practical perspective, I reverse-engineered the joint structure and animation controls of the Victor rig and adapted them to fit my head model. A single animation control at the top of the head provided the attributes for every blend shape applied to Victor, which I used to guide me in all of the blend shapes I would need to create, including those in addition to FACS shapes. With this, I produced a robust rig for my own head with blend shapes as the primary means of deformation. However, previous animation experience taught me that while blend shapes can be powerful and precise, relying on blend shapes for the sole means of deformation can be limiting. This is because the only deformations possible are those which can be created through the activation of one or more pre-determined shapes. This is as opposed to using a system of joints to deform the geometry based on skin weights. Skin weights assign influence to the geometry's vertices closest to a particular joint. As a result, if the joint moves, those vertices will follow it. This provides a great deal of flexibility in a rig, but is less precise in terms of generating exact deformations.

In order to maximise the amount of animation and deformation control I could achieve with my head rig, I created a hybrid system. I used blend shapes to create the foundational controls and created a second layer of control through a series of joints I assigned to the face using skin weights. These joints can be repositioned by spherical markers I assigned to the location of each joint along the surface of the face. The resulting system allows me to first animate the blend shapes to approximate a given expression, and then to move the localised skin weight controls to further adjust the shape.

This produced a worthwhile rig for the head alone, but when I added the body to the head geometry, I ran into several issues. The primary problem was that while I produced the original head rig from scratch, I used Rapid Rig: Modular, a rigging script for Maya, to automatically generate a rig for the entire body, including the head. After several experiments, I determined that I got a better result by doing this rather than attaching the original head rig to the Rapid Rig body joints. However, I did retain several of the animation controls from the original head rig, as these are not generated by the Rapid Rig script, including the animation control with the attributes for every blend shape.

While I thought I had completed the rigging process by February 2018, the only way to meaningfully determine what did and did not work with the rig was to perform animation tests with it. This included testing how the rig deformed the mesh in various poses, especially extreme poses. From the results of these tests I produced various "clean-ups" for the rig, including blend shapes that fixed deformation issues encountered when posing different parts of the body (known as *pose-space deformers*, or PSDs). A number of PSDs were required throughout the body, each of which necessitated additional time for modelling. This was also the first time in my experience that I needed to take a character this complex through to this level of completion.

I produced two versions of the full body rig, one at a low-resolution (13,336 polygons) and one at a high-resolution (213,208 polygons). The low-resolution model was useful for providing real-time playback when animating the body, but it did not provide facial animation. The high-resolution model was used for the purposes of facial animation and producing the final output, as well as for sending to Marvelous Designer, a cloth simulation software, to produce and animate the character's clothing.

V.ii Sculpting Deformations

A major challenge I faced when creating the PSD blend shapes involved my use of ZBrush as the software to sculpt the modified poses. While Maya has its own system for interactively modifying PSDs, this system relies on Maya's sculpting brushes, which lack the flexibility and speed of a dedicated sculpting software like ZBrush. The PSD blend shapes needed to be accurately sculpted and I therefore turned to ZBrush for this. However, in order to create the blend shapes in ZBrush, I needed to first pose the rig in Maya and send the problematic pose to ZBrush as an OBJ file. Once I completed the blend shape sculpt in ZBrush, I sent a copy of it back to Maya. However, a blend shape deformation is calculated based on the default pose, not the pose that I sent to ZBrush. I needed a means to calculate the difference between the ZBrush blend shape and the mesh's default pose rather than the ZBrush blend shape and the pose sent to ZBrush. I spent a number of days trying to work up a solution for this problem, considering different ways in which blend shapes are calculated inside of Maya, as well as how morph targets are calculated inside of ZBrush. Finally, I came across a plugin for Maya called extractDeltas, originally developed by James Jacobs. This plugin enabled me to create a new target shape that is based on the difference between the ZBrush blend shape and the default pose in Maya, thus bypassing the original pose sent to ZBrush.

I created a series of blend shapes from the high-resolution mesh for deformations in the neck, clavicles, shoulders, elbow, torso, legs, and feet, each with low-resolution and high-resolution versions. Each blend shape is based on the rotation of a single animation control in Maya. In most cases this presents no problem, but for more complex movements like raising the arm next to the head, more than a single animation control is required to produce this pose (e.g.: the clavicle and the shoulder controls). This is a complex modelling problem and I originally decided to sidestep it by creating a muscle system for the head, neck, torso, and shoulders. The muscle system produces realistic, physics-based compressions and extensions of muscles that drive the movement of the corresponding geometry surrounding those muscles. However, I soon realised that the muscle system produced its own complexities in terms of assigning the influence of the muscles to the geometry (a process known as painting muscle weights). This was a slow and tedious process, and I did not feel experienced enough to produce an accomplished result. Instead, I reverted to the modelling solution where, despite the tedious challenge, at least I knew I could achieve a useable product. I created a series of blend shapes that are activated based on the amount of rotation of a given animation control involved in the movement of the arm next to the head.

V.iii Resolving Incompatible Geometries

Producing facial animation required some inventiveness because the facial blend shapes were created for the solo head model. In order for a blend shape to deform a mesh, the blend shape has to have the same number of polygons and the same vertex order as the mesh it is meant to deform. As a result, the blend shapes created for the solo head model could not be used to deform the full body model. I had two options: recreate all the blend shapes and high-resolution facial sculpting from scratch based on the full body model; or, create a workaround. I produced the majority of the blend shapes for the solo head model when I originally taught the motion capture paper for which the model served as an example to the students. The creation of the blend shapes is labour- and time-intensive, and I did not want to have to redo all of that work. As a result, I opted for the workaround option. Since the solo head model and the full body model make use of the same head shape, their respective polygons overlap perfectly from the clavicle region and upward.

When there are two very similar overlapping geometries, the movement of one of those geometries can be used to deform the other geometry using a wrap deformer in Maya. I applied a *wrap deformer* to both geometries to let the use of the blend shapes on the solo head model deform the head of the full body model. In effect, I was able to use the solo head's blend shapes to deform the full body's head without having to create a new series of facial blend shapes for the full body geometry – thereby ignoring the issues of conflicting polygon counts and incompatible vertex orders.

One other issue I encountered with this process is that due to the different vertex orders between the meshes, I was unable to transfer the UV layout from the solo head model onto the full body model. The UV layout is a two-dimensional mapping of a three-dimensional object – essentially, the head and body laid out flat. This two-dimensional mapping is used to generate textures for the model. When I taught the original motion capture paper in 2016, I created a full set of textures for my head model from photographic references. Unfortunately, I was unable to transfer those textures to my full body model because they both use different UV mappings. I tried to replicate the UV map of the solo head model, but it is not possible to do this with precision. I made the decision to create new textures for the full body model. I created textures for only the parts of the model that would be unclothed: the head, neck, upper chest, forearms, and hands.

V.iv Rigging Deformations

Blend shapes can be applied directly to the base mesh using a blend shape node in Maya. However, instead of this direct application, I duplicated the base mesh of the body (without any rigging applied to it) and connected all of the blend shapes to that mesh. I call this duplicated mesh the "pass-through mesh". I then created a blend shape node between the pass-through mesh and the original. (I did this process for the head, high-res body, and low-res body geometries, respectively.) This setup enables the base mesh to receive the deformations from the pass-through mesh, which is itself deformed by the blend shapes. There are technical reasons for creating a pass-through mesh. All of the blend shapes are applied to a single blend shape node on a given mesh. This process, including assigning attributes to control the amount of deformation provided by a given blend shape, takes time to setup. It is conceivable that throughout the rigging process there will be a reason to delete the construction history on a mesh. Deleting the history of a mesh erases specific connections to that mesh, including blend shape nodes. In the event that this should happen, it is easier to re-establish a single blend shape connection from the pass-through node to the base mesh than it would be to reconnect all of the blend shapes, including any additional connections created for those blend shapes. (This is especially relevant for the head, which has far more blend shapes and complex connections.)

I created 104 facial blend shapes and 36 body blend shapes (in the form of PSDs). While the blend shapes provide a good range of facial movement, the possible expressions that can be achieved with these shapes alone is still too limited for the degree of realistic movement needed to faithfully replicate human facial performance. In order to achieve meaningful performance fidelity, I created a series of additional facial controls that are not based on blend shapes, but which deform the mesh in concert with the blend shape deformations. These on-face controls were based on skin weights, which mapped a degree of influence of a joint onto a given mesh vertex. I created a series of joints to match the location of markers I placed at key regions around the head to provide optimal deformation. I created these markers in the form of mesh spheres and parent-constrained each joint to its corresponding marker. As a result, by moving a marker, the corresponding joint moved along with it. Each of these markers was a SWB control at a region of the face that required greater flexibility than what could be achieved with blend shapes alone. These areas included the eyebrows, eyelids, cheeks, nasolabial folds, and lips. The joints associated with each of these markers were added to the existing skin weights for the head mesh. However, in order to prevent problems resulting from an automatic redistribution of skin weights to accommodate the new joints, I made sure that the joints' influence values were set to zero by default. This allowed me to "paint in" the amount and area of influence any of these given joints had on the head mesh. This operation was performed using post-normalisation of the weight values, which meant that I could paint influence onto a vertex without reducing an equivalent amount of influence from other joints on that same vertex (as is with the classic method for painting skin weights). Post-normalisation calculated the total amount of contribution of all joints on each vertex and scaled their contribution to a total value of 1.0.

I also created a series of controls to facilitate the activation of any given blend shape. By default, when a blend shape is connected to a mesh, in order to activate the blend shape, its weight factor (its relative contribution) must be increased or decreased by manual data entry or by using a slider associated with the weight factor in the Shape Editor. This was a cumbersome, non-intuitive process. In order to make it easier and more intuitive, it made sense to activate the blend shape by simply moving a control near the region of the face associated with the blend shape's deformation. For instance, by selecting a control for the left eyebrow and moving the control up, this could be used to trigger the Eyebrow_Raiser_L blend shape.

The association between moving a given control and triggering a given blend shape is created using driven keys in Maya. Driven keys are different from standard animation keys. With keyframe animation, a given attribute changes value over time. By contrast, driven keys do not produce animation based on time; instead, the change in value of one object or attribute results in the change of another attribute. That is, the first object or attribute is the *driver*, while the resultant attribute is *driven*. In this example, the eyebrow control moving up *drives* the Eyebrow_Raiser_L blend shape's weight factor to increase in value, activating the visual change of the blend shape in the mesh. The result is that when the eyebrow control moves up, the mesh's left eyebrow region also moves up. This is a much more intuitive workflow that makes the animation easier to edit. I created similar facial controls and driven relationships for the upper and lower eyelids, eyeballs, cheeks, nostrils, mouth, upper and lower lips, lips corners, chin/jaw, throat, and neck. Some controls are given more flexibility of movement than others, and may contain custom attributes to drive additional blend shapes.

Appendix VI. Overview of "The Seagull"

Treplev's love interest is Nina, a young actress who admires Arkadina. This scene does not normally include Nina; however, based on the input of Dawn Glover, my acting coach, she felt the monologue as a standalone scene would be more active if Treplev interacted with the person who matters most to him – namely, Nina. Sorin has a single but inconsequential line within Treplev's monologue, and we removed that when we replaced Sorin with Nina. By using Nina as Treplev's scene partner, this changed the intention and mood of the scene. Nina's goal is to be just like Arkadina, who represents the old forms of theatre that Treplev rails against. Therefore, due to Treplev's love for Nina, her inclusion in this scene presents a distinct challenge when Treplev equates her ambitions to worthless pursuits.

The script mentions other characters, most notably Trigorin, the writer who Arkadina dotes upon. In the Chubbuck Technique, whenever speaking of a character who is not directly in a scene, the actor requires an "inner object". Chubbuck describes inner objects as "[t]he images and pictures you see in your mind when speaking or hearing about a person, place, thing, or event." (2005, p. 77). An inner object should affect a particular emotional response from the actor that is consistent with the character's emotional response in the scene. As an inner object for Trigorin, I delved into a past relationship that ended when an ex-partner became involved with another person; because there is a strong connection to betrayal and contempt, this "other person" became a useful inner object for Trigorin. As an inner object for the theatre that Treplev disparages, I used homeopathy (a pseudoscience I despise), and for the playwright who creates such theatre, I conceived of him as a homeopath.

A notable point of difference between enacting this scene on its own versus within the context of the larger play is that a scene requires a stronger emotional build when it stands on its own. Therefore, the dynamic of the drama in my take is perhaps greater than what would normally be expected of the scene when it is buttressed by other moments that normally precede and follow it. The actor would normally have the rest of the play to get to a similar level of emotional intensity.

There is a moment in the source performance where I start to say one line and immediately shift to another. The line begins at 2:57 with a drawn out "I" and becomes "I… my passport says I'm a bourgeois of Kiev." The "I" is not part of the script and it is likely that in the moment, my instinct was to incorrectly say the line as "I'm a bourgeois of Kiev". Given the opportunity, I would edit out that mistake and attempt to create a seamless performance that connects the previous line as follows: "I don't have a kopek to my name. My passport says I'm a bourgeois of Kiev." This is not just an audio adjustment and would require blending and animating the flow of expressions between these lines after I remove the erroneous moment. This would be more difficult to achieve in a continuous shot and would be easier if I cut between the two shots.

Appendix VII. Instructions to PeCap Artist Assistant

The following are the exact instructions I provided to my PeCap artist assistant prior to her cleaning facial data for the dramatic performance.

1. The primary goal of this exercise is to complete the clean-up process of the performance capture retargeting. You are to do this by creating as close to a one-to-one correspondence between the facial expressions/motions in the source video and the 3D character onto which the performance data is retargeted.

2. This is not an animation exercise. I don't want you to interpret the resulting motion. That is, don't try to animate the result in order to make it "feel" more like the original performance. I only want you to use your best judgment to create a one-to-one correspondence between the source video and the motion of each facial component.

3. These are the following controls on the face rig that still need to be cleaned up:
    a. L_Lower_Lip_CTRL
    b. R_Upper_Lip_CTRL
    c. R_Lower_Lip_CTRL
    d. Nose_L_CTRL
    e. Nose_R_CTRL
    f. L_Eyebrow_CTRL
    g. R_Eyebrow_CTRL
    h. L_Eyebrow_Inner_CTRL
    i. L_Eyebrow_Outer_CTRL
    j. R_Eyebrow_Inner_CTRL
    k. R_Eyebrow_Outer_CTRL

4. Make sure to clean up all of the parameters on all of the controls listed above. Do

this by using the Red9 plugin to access the Red9: Interactive Simplify function. First select the animation control's parameters in the Graph Editor, then click the checkbox in the Red9 window to enable Snap to Frame. Then, under "Curve Resampler", set Resample to 4. Leave the values of Curve Simplifier at their defaults. Ensure that the resulting new keyframes are spaced three frames apart and on integer frames.

5. The frame range is 626-1583.

6. Work from the controls that affect the largest regions of the face first to the controls that affect the smaller regions in a progressive order. For instance, with the brows, start with L_Eyebrow_CTRL before moving onto L_Eyebrow_Inner_CTRL or L_Eyebrow_Outer_CTRL.

7. Once you complete all of the main controls, then move onto a clean-up pass for the Head_Markers (the localised skinweight-based controls on the Jason_Head_Markers display layer).

8. Complete all of the frames on a control before moving onto the next control.

9. Before starting a new control, always append the file name by one interval. For instance, when you start on you first control, the file name should go from Jason_DP1_Faceware_18_626_1583 to Jason_DP1_Faceware_19_626_1583.

10. Always ensure you have incremental saves enabled and frequently back up your work to another storage disk or cloud backup.

11. You may find you need to make adjustments to controls I have already cleaned up in order to create the best one-to-one correspondence for the control you are adjusting.

12. As you work, please make note of any time that you run into a moment that you consider noteworthy. I leave the interpretation of this term up to you. For instance, you may consider a moment when you are struggling to resolve a particular facial shape to the source video to be noteworthy, or when you have to adjust a control I've already animated in order to get the right shape on your present control. Don't be afraid to point anything out – it is better to err on the side of over-documenting than under-documenting. Even just voicing your frustrations or accomplishments may be potentially insightful for me.

13. When you are documenting your experience, use a capture software that records your work screen and your voice. I recommend using Open Broadcast Studio (OBS). If you decide to record the entire session, then use Kuadro to summon a red square to overlay on the screen whenever you get to a noteworthy moment, then make sure to remove the red square from the screen as soon as you finish with that moment.

The most recent version of the Maya scene file is available here:
[Link removed]

The red square for use with Kuadro is available here:
[Link removed]

This is the original link that I sent you, which contains the rig file, textures, and video frames. Make sure to replace this scene file in this package with the more recent one above:
[Link removed]

VII.i Assistant's Feedback

In her feedback, my assistant noted that when starting the work, she found it difficult to get her workflow correct. This was a matter of her adapting from her experience of using the similar but less complex Victor rig, and learning the relationship between the controls on my rig and how they influence the localised movements of the face. She also helped note moments in the 3D character's performance where there were any inconsistencies between what I had produced and the original performance, and adjusted them accordingly.

In one of her reflections, she noted that working with facial performance data was quite different to the workflow she had been taught at Weta Digital, which involved a strictly keyframe-based approach to creating facial performance. She observed that she felt she was making so many changes to the original data that the original data was almost overwritten. In such instances, she questioned whether the animation should be considered "keyframe animation" rather than "retargeted data". This is an important observation that is consistent with my own conclusions within this project: the final PeCap facial animation ontologically becomes keyframe animation that is heavily influenced by (e.g. uses as a basis) the original retargeted data. My assistant also noted that when looking at some of the data, especially when it demonstrated a lot of noise or unwanted variation, her desire was to simply delete all of the data and keyframe it from scratch. She attributes this to the keyframe workflow she learned at Weta. It is worth speculating how frequently PeCap animators actually delete PeCap data and keyframe animate in its place, while still working to honour the original performances.

It was interesting to note how my assistant became frustrated by having to adapt to a different workflow/pipeline from what she has encountered in the past. This is an element that I did not account for in my research so far: the struggle of the animator to embrace different workflows and methodologies, especially when these are dictated by technologically determined factors. She noted that at times, the new workflow "sort of overwhelm[ed her] and [her] focus". She noted that when she shifted her focus from trying to complete the full animation to “focusing bit-by-bit and making sure [the animation is] really polished”, it helped her to feel more confident and less overwhelmed. This shift in approach came as a result of my video feedback where I demonstrated side-by-side comparisons of the original performance and the animated performance; this helped her to see how many changes actually needed to be made over small intervals. It was in noting how much I had to adjust the underlying data that helped her to feel more confident about doing the same. She noted that her updated workflow felt tedious and slower, but it made her feel more confident with the final results. Instead of trying to race through the animation to get it done as quickly as possible (consistent with certain industrial imperatives), she forced herself to examine the minutia of every frame; this was the only way to achieve the one-to-one result I required. However, she noted that the degree of focus and repetition required was taxing (“unstimulating for my brain”). In order to maintain her focus, she needed to ensure her other senses were engaged, such as listening to music, audiobooks, or podcasts. She also noted that not being able to interact with the rig in real-time took a toll on her focus. She indicated that as a visual person, she found her “mind getting frustrated that it can’t see results in real-time.” This is especially true when dealing with highly nuanced facial adjustments, which are difficult to accurately evaluate without real-time interaction and playback.

When it became evident that she would not be able to finish all of the facial clean-up within her allotted time, I decided to switch her to cleaning up the body performance data (I later completed the remaining facial clean-up). I provided her with instructions for cleaning the body data similar to the instructions for the facial clean-up. She reflected that between the two clean-up processes, the process for facial data was more about adding to the data and changing the animation curves via keyframe animation, whereas cleaning the body performance capture was more about deleting extraneous data and smoothing out the animation curves.

Appendix VIII. Interview with Andy Serkis

JK: Are there any kinds of adjustments you make to your acting practice when you work with motion capture versus live action?

AS: Absolutely none. I think that's one of the greatest of misconceptions about performance capture is that it requires you to be anything other than an actor if you're playing a role using performance capture technology. Obviously it helps if you have an access to your physicality, but not necessarily. I mean, it depends on the role, of course. Every actor, of course, as you would well know, has their own methodology of work, whether they work from the inside out, whether they work using emotional recall, whether they observe or research. There are millions of different ways of approaching a character, and all of those ways apply when playing a role using performance capture technology. I've always maintained that there is absolutely no difference in the acting process between playing a role where you're authoring a role wearing a costume and makeup, and when you're authoring a role wearing a suit with facial markers and a head-mounted camera. Of course, depending on what sort of character you're playing, there's periods of research, there's periods of methodology in terms of how you're going to physically pull that role off, but that's the same in live-action and the motion capture space.

JK: Great, thank you for that. What do you feel are common stumbling blocks for actors new to working with motion capture?

AS: There is, again, a perception that you have to pantomime, or over-perform, or demonstrate the character. It's actually the opposite of that. It's very much about being the character and allowing the subtlety of your movement, of your physiology, of your physiognomy to play as you would if you were giving any live-action performance. So, there's not a necessity to project the character in any sort of theatrical way. Again, if you're playing a physical character such as Caesar in the Apes movies, then there's a physical language which will have to be built. But actually, in the case of all the characters I've played using performance capture technology, really the stillness you have on screen as a live-action actor still pertains. The stillness you need to you need to hold a close-up still pertains in the performance capture realm.

JK: Now speaking of that, there is an awful lot of stillness in Caesar, which is great. I watched Dawn again last night just to kind of refamiliarize myself, and the thing that I saw that was really what I thought was quite amazing – I haven't seen it exactly this way before – there was so much of you in the eyes and this sort of ocular region. You could really see you in that performance. I don't want to call it uncanny because it wasn't uncanny – it was actually perfectly canny. It's getting to that level where your performance is so strong and coming through and being felt in that natural stillness. And I think like you were saying with motion capture and performance capture there's this notion that it's big and it's action-packed, and so often I think that misses out on those subtleties that a more realistic approach to acting will give you. Do you think that animation in any way changed our understanding of what is acting?

AS: I think if you look at performance over the centuries, the human condition is being portrayed using many different artforms, in terms of performance. So, if you go back to Indonesian shadow puppetry, or Greek mask work, or *Comedia del Arte* in Renaissance Italy, and so on, so forth, the use of mask work, dance, interpretive, dance drama, physicality, physical theatre – there are so many different ways that the authoring of a role and the imbuing of a character with life. It's like, throughout the centuries there have been these different forms that have allowed the actor to take an abstract character – a character that is not photoreal, a character that is not themselves – and to transform, transmute themselves into those characters. And so animation, for me, falls into that category. In recent times we've been challenged to ask where does the role of the animator and the actor cross over? Where does the role of the actor... how far is the authorship of the role – and that's one of the grey areas in the big debate about performance capture – where does the authorship of the role begin and end? Does it begin and end with the actor? How much is handed over to an animator for interpretation? And I think that very much comes down to individual projects, it comes down to the will of the director and what the director wants to see out of that. So I would say that on a whole that animators, like all human beings, have the potential to act, have the ability to act, to place themselves mentally in the mind of another being – another person, another creature, another character – as all human beings do, if they allow their imagination to go in that way. Obviously, the skill of the animator is through modern-day technology using the computer to transmit that. Previously, obviously, it was through drawing and through stop-motion, etcetera, etcetera. But I think the difference between animators and actors is that actors have the ability to completely bury themselves in a character over a long period of time – i.e.: if they have any theatre experience, they're able to play a role for three hours straight and go into improvisations that last for four or five hours, or a day, or however many... I mean, to have extended periods of time getting under the skin of the character emotionally, and truthfully allowing themselves to lose themselves within that role. And I think that's the differentiation with animators, because I believe animators do that, but it's not required of them to do that – to actually take a script and to work on a character from page one to 120, and create the emotional arc and journey, and so on. Because that's not part of, generally, what animators are asked to do. They're asked to work on short segments of a character. Or if they're working on one character, then they're often not aware of the entire story. As an actor, you have to have a full understanding of the function of your role within the whole, and obviously the emotional journey and the emotional arc. So, I think that's one of the key differences. But it's interesting because in recent times, I think more animators now are – whereas they were not so open to the idea of performance capture and saw it as a threat – they are now embracing that as a form of trying to understand character.

JK: Sure, sure. I mean, I think that's... one of the things I'm trying to propose with the PhD is the possibility of using performance capture as a way of creating better animation reference for animators, and to encourage animators to find ways to make better animation reference, understand what good animation reference is versus bad, and how what they do can result in good versus bad. Good in quotation marks, of course. You started touching on some of my other questions here...

AS: Just to add to that, actually... For instance, I think generally animators are asked to choreograph a scene physically, and to block a scene physically, and there are many instances where performance capture – if a director so requires – allows the happy actor to rule over the staged moment. What feels real and emotionally truthful for that moment for an actor is something that I don't think can be... That's why they're qualitatively different things when you watch – because the process by which we arrived at that moment, you're open as an actor to receiving what another actor's doing. And often in animation you're working in isolation. The chemical exchange, if you like, between actors is generally what informs the freshness of that moment, and finding in each different take and series of takes – and if you're working for a director who's looking for the heart of a scene, who is looking for different ways of interpreting the moment – the arrival at those moments that feel real and engaging are because of that.

JK: I mean, I've always found it difficult myself when I try to create animation references – lacking a distinct scene partner. And unless you have somebody there with you who can at least function as a scene partner, it's very isolated. In the end, what I see so much of in strict animation is what I think are very surface level performances. And that works, I think, for certain kinds of animation, and I don't suggest that the entire animated medium needs to all of a sudden go to hyperrealism. But I think that there are a lot of possibilities that could exist within animation – if there were alternative approaches to the way that we create our acting process around our animated characters, and therefore, the way that we ourselves do an acting process as animators. I look at... are you familiar with the animation "Ryan" by Chris Landreth? It's about Ryan Larkin, Canadian animator. It's a brilliant animation and it features some of the best acting in animation I've seen. You have a completely different feel when you come out of watching that animation than you do when you watch more entertainment-based style animations.

AS: What's his process, out of interest?

JK: I think he, himself, also has a much more tuned-in approach to creating his reference.

I'm not sure that he's necessarily an actor but he's certainly somebody who is very clever at picking up subtlety and how that affects the entire face. Everything that he does, like Ryan, there was no performance capture for that at all. It was too early. So, everything was key-animated. It was quite a stunning result really. This is one of the questions that sort of sits as the crux of the thesis. What does acting become when the product of acting starts as data and finishes as a digital performer whose performance resembles, but is also different from, the actor's "original" performance?

AS: Again, there are different answers to this. Depending on what the director is wanting from this performance, I think it can vary. The actor's role with any form of screen work… the role of an actor is to serve the character and to serve the story. Now, you author the role – by your choice of ways of responding to the script, by the choice of costume that you choose to put on, by the choice of makeup, portray the character, arrive that by internal decisions about who the character is.

JK: Up until now, you are an actor on stage or on film. The performance you give is effectively the performance you get back. With film, of course there's editing involved and that can change how things were put together. But it's always you that we're seeing in the end. With performance capture and animation technologies – there's the you that starts there, we digitise that, it goes through a process whereby animators and other people are applying lots of changes – and what we get out is inspired by you, certainly. It has all of its origins in you but it's not that performance that you originally gave. Does that in anyway change what our understanding of what acting is?

AS: I don't think it ought to because you are authoring a role in the same way. As screen actors, we don't control, ultimately, the emotional arc of the character because of the edit, because of the choice of music, because of the choice of shot, because of the choice of so many different things. So, we are augmented. Every single actor's performance is augmented from a minor way – from an edit point, to having a blink removed, or a tear, and so on and so forth. So, unless you're talking about stage acting where you purely see the actor on stage, then I would say all forms of screen acting are augmented. I would say that, again, it's the approach to how you see performance capture. In playing the roles that I have played, it's always the process of arriving at that point, as I have said, is exactly the same in terms of emotionally engaging with the character. I would say the way that you transform an audience, the way you connect with an audience, the way you touch an audience emotionally or whether you make them think and feel is no different at all. Because I don't believe that the animator actually changes anything to do with the authorship of the performance emotionally. What the actor gives off emotionally cuts through and touches the audience. So, really what I believe the animators do is, depending again on the view of the director. Because you have certain directors who work with performance capture who will send the shots back time and time again until they can absolutely read the actor in that performance. Then again, you have directors who might want to, for a specific project, have that performance enhanced in some way by the animator. If it's a slightly more cartoon-like, slightly more heightened or cartoon physical rendering of the character.

JK: Thirty percent more angry.

AS: Exactly! Or not thirty percent more angry, but thirty percent more let's have the jaw open thirty percent more. More in the technical level. I don't believe the actual emotional intensity or authorship is altered at all.

JK: And would that for you be the essence of where the acting is, is more in that original intensity rather than the actual final product necessarily.

AS: Yeah, because... I mean, I'll give you an example. Matt Reeves who directed Dawn of the Planet of the Apes. He directs the scenes with the actors. We are shot, we are filmed. He then takes the performances and cuts them. And, if there's anything that varies from absolutely, quintessentially what he is given by the performance of the action on the day – and then he lives with that cut by the way, and screens that cut to many, many people [sic]. He creates the arc of the story, he creates the journey, understands the dramatic principles of the film, of the movie, from watching those performances. Nine months later on the visual effects shots come in and the actors are gradually replaced by those shots. But he's already made the film, he's already told the story with those performances. So, he is looking to basically have them clothed, or make-upped if you like. Hence this term, which really angers animators: "digital makeup."

JK: I'm not too fussed either way of the term to be honest.

AS: No, it's interesting because it's caused all manner of debates. All sorts of rancour, and anger.

JK: Oh, I know. Well, I mean, I've read things where you have been quoted about a people than angry about it. I think probably..

AS: And it certainly, by the way, did not come from me.

JK: And as these things usually are, there's a lot more to the story than what you read in, wherever. Okay, cool. I think I can start focusing on a few different questions here now because you've touched on some of them. You said that basically your acting process is pretty much the same if we're going into performance capture as it is for any other thing. Are there any things that you developed new for dealing with performance capture, or is it strictly…

AS: Well it, again it depends on… I mean, what I'm interested more now in the use of performance capture is expanding the types of roles, or the experiences, or using The Imaginarium, which is our performance capture studio. We're looking beyond just playing actors playing roles. It's performance capture with dance, using performance capture in conjunction with live theatre, it's using it… It's such a broad term. It's ridiculous sort of, that, it's not a genre of acting, it's a technology. So, that's why I say in terms of the process of playing a role is entirely the same, in terms of building blocks of making a character.

JK: When you watch a movie with a digital character, do you ever feel anything is missing in that performance? Especially if it was something that you had created yourself, as well.

AS: Sometimes a performance can be slightly misread and the intentions slightly misread. But I have to say, that Weta have established a pipeline over the years which is, increasingly being able to understand and interpolate the performance to such a high level now – and "Dawn" is a great example of that. With every iteration of technology, with the pipeline, with the workflow and all the people that – and not to say, it's not just about technology, it's about people. It's about people having an understanding what is required, and not seeking to augment unnecessarily or to add their flourish. I think if you're an animator working with performance capture, you have to for these types of films – you've got to want to honour the performance and that's what it comes down to. Because that is why all of the trouble has been gone to in the first place: of shooting on set with the director who wants to have that relationship with actors. They want to have that sense of the drama they create in the moment on the day, on the film set, honoured.

JK: It seems that if the clean-up artists and the motion capture artists have a good understanding of what good acting is themselves, then that would benefit their ability to preserve those performances. I mean, I have a number of students now who graduated from our program who work at Weta – especially Weta Mocap, because we're one of the only schools that teaches motion capture as a large area. So, people you're talking about are them, as well, which is quite cool. Let's see, just going through some of the more important ones here… Have you ever noticed whether certain emotions consistently read better than others when performance captured?

AS: I mean, the ones that are probably easiest to read are the most obvious ones like anger, and rage, sadness. I think the ones where the physiognomy is very obviously dictating or showing, it's much easier to read. I think it's much harder on a more enigmatic or a character that you can't read. Or actually, characters who, for instance I would say that a character like… if you take Tintin as an example. If you take Tintin and then you take a character like Captain Haddock, there are so many more visual cues in Captain Haddock's physiognomy than on a character like Tintin who has a smooth face. I think characters which have less defined in their physiognomy and have smooth, younger faces, sort of unwrinkled skin, simple facial structure – I think actually that's some of the hardest stuff to pull off. Whereas if you take an old, battered, weather beaten, wrinkled face. It's much, much easier.

JK: Do you consider animators, people who just do complete key-based animation to be actors?

AS: I think some of them are. Absolutely, and I think some of them are definitely not.

JK: What would be the distinguishing feature?

AS: I think, as I say as we've sort of discussed really, but a full understanding of character… I mean, when you play someone as an actor, you don't look into a mirror. Because I think a lot of animators use a mirror to actually create the character. But actors don't stand in front of mirrors on a whole. You're finding… you're using emotional recall or you're using parts of yourself – using a lot of parts of yourself – I'm putting that under the microscope. And I'm sure animators do do that, too. But I think, again, I go back to this notion of – if you asked an animator to play a role over the course of the length of the play of an evening – they might find that challenging. Whereas, what they are brilliant at doing is taking a moment and making it crystalline. Sometimes the good ones, the good actors, make that moment crystalline but also subtle and emotionally engaging, and the ones that are not so good make it look flashy and it's performed. You could say the physical blocking to be exactly the same but you can tell, you can always tell… I mean I know from the work that I've done – the ones who want to show their wares, to show their skills.

JK: But the actual animated characters themselves, in the end. Are they actors?

AS: I don't think so. No, they are a skin. Or that's how one should approach them as an actor. Because if you're playing a motion capture role or performance captured role, you have to believe that you are the driver of that. It is like being a driver of a performance racing car. The analogies that have been drawn in that kind of… the conductor of the orchestra – there are of course lots of other moving parts and fundamental, crucial moving parts to making a concerto scene – really play beautifully. Or a Ferrari racing around a track. But, the actor is the driver or the conductor.

JK: How do you understand "emotional authenticity" in acting?

AS: Lots of different ways. I'm not one actor who sticks to one particular method so I don't just use the Strasburg's method, or Sanford Meisner, or Stanislavsky. I'm very much a… a lot of actors are actually magpies and pick and choose and gather experience. If you work with Mike Lee, or whether you work with Peter Jackson, all of the different directors that you worked over the years, you accumulate your own catalogue of ways of accessing characters I suppose. Also, as I say, that the actor that I am today at my age is a very different person to obviously, what I was when I started out at the age of twenty. So, of course if you're open to, and I think most actors are open to learning at every single job that they do because you have to because every single job you do as an actor, you put yourself through a chemical change. You cannot go back from that experience. So, working with a director like Mike Lee for a character for six months to a year where you're completely losing yourself in that character – you're only listening to the music that character listens to, you only associate with likeminded people, or you don't see your own friends, you don't communicate with the outside world – you're living in the realm of that character. You actually are forcing a chemical change to yourself as a person, and you can't go back from that. You just can't. You're opened up. You've opened up a conduit to another part of your personality. I find it easier to cry now because of the experiences of my life to this point, for instance. Certain points of your life you feel more full of rage, at certain points of your life you're more inquisitive. And those things are accessed and put under the microscope in every single character you play.

Appendix IX. Interview with Dan Lemmon

JK: Randy Cook says that "Andy [Serkis] really should be considered the principal author of Gollum's performance but there is a hell of a difference between principal author and sole author. I can't speak for the recent performances in Andy's performance capture career, but the animators on the Lord of the Rings were most certainly not 'digital make-up artists' and nobody has any business saying that they were". And that's the end of Randy Cook's quotation there. Given that you, Dan, have had first-hand experience of working with Andy's recent performances, do you feel that the notion of "digital make-up" is more appropriate to the ways in which the synthespian performances in the "Apes" trilogy were created? Should Andy be seen as the sole author of Caesar's performance for instance?

DL: So, well as you pointed out, it's a little bit of a contentious topic. And there's certainly a lot of different people's perspectives on what, you know, what's involved from where they're sitting, creating a character, and all the different pieces that go into making these creatures live and breathe on the silver screen. And I think that's kind of the, the key thing is, you know, a successful execution and continued maintenance of the suspension of disbelief. You want your characters to appear in the movie and for everybody to just accept that they're real, that they're participating in an environment that you believe, that their emotions you believe, that they seem motivated, that they seem, you know, that their physical appearance appears real but also emotionally you buy them as well and this is, you know, this is a comment that's been, or I guess criticism that's been labelled against each of these movies, I guess in particular the first two. But, you know, when the apes go toe-to-toe with the humans, in a lot of cases, the critics in the audience believed the ape characters more than they believed the human characters.

JK: Interesting.

DL: So, the thing that we're, you know, creating with pixels and computers ends up being more lifelike, more believable. Not necessarily because of its appearance, but emotionally the performance seems more authentic than the performance that the human actor is putting in, and I think this is an astute observation. I think it comes down to the quality of a number of things, but one is specifically the quality of the actors that are playing different roles. How many movies have we seen where somebody is fantastic in a scene and then you cut away to the actor who is maybe less talented or a little bit less engaged, and the whole scene falls over? And this is something that the performance capture process is really good at, is really useful in giving filmmakers and actors the tools with which to craft scenes, to block out in stage and act in scenes in a way that's more comfortable, more familiar than the animation room, than the way we would've been making digital characters before the invention, adaptation of performance capture. When you've got a grip waving a stick around with a tennis ball at the end of it – and, you know, saying that this is a scary monster, react to it – you get a totally different kind of performance from the human than if you actually had something or somebody in there playing that character. Actors do better when they have real emotional things happening in front of them to react to. And directors do better as well, you know, everybody does better. And this is something, it's not just about making the apes look better, it's also about giving the humans something to react to and to play off of as well.

The other thing about the process has to do with how you handle things as you get into post, long after the actors have gone home and moved on to other projects. You know we go back with the dailies, and this case Matt Reeves, our director, has taken all those performances that he shot on set and he's cutting together with Bill [Hoy] and Stan [Salfas], his editors, they're putting together these scenes and all of the storytelling decisions that he's making and cutting that scene together, are based on what Andy was doing on the set. And so, he's making specific choices from cut to cut, when to cut, all the pacing, all the mood, all of the tone is based on what Andy did. And so, there's a version of the movie that goes together where there's no apes at all: it's just the dailies with Andy, and Kieran and Terry – all the ape actors wearing their funny grey suit with the little dots on them and the scenes play out with them in those funny grey pyjamas. But you watch the scene and as happens in sort of every kind of aspect of storytelling from the dawn of time, as soon as you kick into that story, your suspension of disbelief engages and you forget that you're looking at actors wearing funny grey pyjamas, you just accept that that's an ape now, that's the big gorilla. The cues that tell you that's true has all to do with what the actors are doing, their face, their voice, their posture, and you either buy it or you don't based on what they're doing. On set this is something that Matt spends take after take trying to get right, both with the ape actors and also with the human actors and his approach to directing scenes and my approach to doing digital characters, digital effects – we're very similar. We look at what's in front of us and we ask ourselves the question: "Do I believe it?", and if the answer is yes then we can move on to the next scene; if the answer is no, we've got to pick that apart, we got to figure out what is it about this that doesn't feel right, that I'm not believing. Ok sure he doesn't look like an ape but that's going to be okay later, but emotionally is he doing the right thing, does he seem motivated, is the reaction authentic that he's giving to the other actor? You know he'll, after every take, he'll run out there, he'll talk to the actors, he'll try to pry some things out, he might change the blocking. Sometimes one of the important things that happens on a set is the actors will have objections, or they'll have suggestions, they'll say "well look this word doesn't feel right in my mouth, like I don't understand why I'm saying this at this point. I thought I had this information but I'm acting like I got this extra piece of information that I don't think I've caught onto yet." Something like that.

There's this brilliant scene in the second movie in Dawn of the Planet of the Apes, where Toby Kebbell, who played Koba – it's the scene where he goes into the human's ammo depot and there's some soldiers there drinking whiskey. In the storyboards we originally planned to shoot this on this kind of navy ship that we thought we had access to, and we lost access to it and we ended up just shooting it in an alley behind another set that we had, it was an improvised, last minute kind of thing. The set really wasn't a whole lot to speak of, the actors that came on board to play the role of the humans, they were only there for a couple of days, I think all of their scenes were shot in two days. Nobody knew them. They showed up, they blocked the scene out, they left. But Toby came on the set and they started blocking what they were going to do in the scene and he said – Toby's English, he grew up in the UK and spent time in pubs in London – he said "I wonder if I played this like, I feel like I should play this like these English toughs. That they act like they're your best friend, but as soon as you drop your guard they come around and stick a knife in you." And he recounted a story where something very much like that happened to a friend of his. That's where he was bringing his own background, his own story into that scene. That scene played out, when we were watching it on the set on the day just from the video monitor, it was chilling – it was amazing how well it was and how scary that character became because you trusted him for a minute. You're laughing along with him, he's clowning around, and all of a sudden, he murders these two guys.

JK: A defining moment for Koba's character.

DL: Yes, yes it really is and he's so cold about it and it was something that was a singular... It came from Toby, from a specific suggestion, and then a series of decisions that he made in terms of how to play that character and in then negotiating with the other actors and Matt. All of that kind of hesitancy from the soldiers and warming up to him, that just unfolded as actors worked together in that scene. It was really something kind of amazing to watch, and of course we went back in and we scrubbed every bit of Toby out of that scene, at least in terms of his physical appearance and we replaced it with Koba. So, I guess it's kind of a long way of saying that the whole point of the performance capture process is to allow actors to work with each other, as they would in any other movie, and to allow the director and the other people that are on the set making the movie to react to them, to put them in an environment, the environment as well. And have all of that authentic interaction that you would normally have when you're shooting a live action movie, but then capture every bit of information you can with that, what the actor is doing, so you can at the best of your ability, you can replace their appearance with that of a character that looks totally different from them.

I guess this is the thing when you're talking about authorship, when you talk about what goes into making a character. We're really talking about very different parts of the process, I'm trying to think of an appropriate metaphor, but basically you got somebody that's basically laying down all of the instruction, all of the DNA of what's going into this character emotionally and then you got another set of artists and technicians and crafts people that go to a tremendous amount of work to service that performance, to stay as true as possible to those decisions – to that intentionality – to keep everything firing in the same way it was fired before. You know, this is one of the things because Matt built his cut around Toby or Andy playing these scenes. The cut doesn't work anymore if we start mudding it up if we start making different decisions about taking the face to a different place emotionally – changing the blinks. The cut functions on very delicate, specific things that happen on the actor's face, on the actor's body – and if you alter those things even in a small amount, it won't have the same… Well, in the best case, the emotional impact is diluted and in the worst case, the whole cut and the scene fall over because it just doesn't work anymore.

JK: So, what you're talking about there is a lot of what Andy actually said to me as well, so it's quite cool to see that you're both on a very similar page. Now when Randy was talking about Gollum back in 2003, he was talking about how some of the Gollum performances that were actually captured were completely changed, or sort of redone, by animators for various scenes. Were there any scenes in the Apes trilogy, for instance, that were constructed by animators as opposed to the performance captured data?

DL: Yes sure, for background characters. In cases where physically the actors couldn't do what is impossible for a human being to do what an ape could do. In those cases, we would revert to reference. Probably less so in "War", but in "Dawn" for example there were cases where apes were running along wires or swinging along the Golden Gate bridge. These are things where we actually had an Olympic class, maybe not an Olympic-class but a champion college gymnast. One of the, like, power to weight ratio, the strongest, lightest guy I've ever met who was an expert gymnast-slash-stunt guy, who was trying to do the what we call the brachiation, which is basically when an ape or monkey, swings from one arm to the next. And, you know, you see this classic movement – he did it really well. Basically, they release and they swing from one arm, and they'll be airborne for a moment, and then catch with the next arm, and without double pumping or anything they'd continue the same forward movement in their body, while they catch and then swing through to the next arm. We had guys that could release and then they'd pump. Basically, when they caught with the next hand they'd have to pump before they could go again. And this guy, Jason Chu, he could do it, and he did it three times and then he tore his chest muscle. He was no good to us for the rest of the shoot – it was a real shame. But even those three times, we used them over and over again as good solid motion capture data we could build cycles from. But anyway, I guess the long way of saying that there are some things just physically that humans can't do, and in those cases the animators sort of takeover, either 90% or 100%. In some cases we'd look at video references, apes doing the thing we want, we would do all sorts of things. But in that case, when we're creating motion – a lot of the background action, we'd pick up with either our local stunt player teams to kind of fill out that action. We call it a sort of un-directed background action. As far as the hero characters go, the only things that they can do, is pretty much what they've done. There's a scene in Apes, in War, where they arrive at the oyster shack just before they find the little girl, and they pull up and they hop off their horses, and because Karin's character Maurice always sort of walks on all fours, never really walks bipedally, we had her crutches, her arm extensions, sitting off to the sides. So, she rides up on a horse, she hops down from her horse and then she's got to put her arm extensions on, but Terry Notary ended up crossing off in front of her so she couldn't get around to get to her arms, so she's kind of doing this shuffling back and forth thing. So, there's a few technical things like that that we would smooth out or we'd tidy up. But for the most part, mechanically, it's almost 100% what the actors are doing.

JK: So, it sounds like what Randy was originally talking about in terms of digital make-up is more in line with the way that is actually produced now.

DL: Well I wouldn't say that, there's a couple of big caveats coming through here. We would take 100% what the actor did, to the extent that the actors anatomy lines up with the ape anatomy.

JK: I was hoping we could get to that.

DL: As you know, ape anatomy doesn't actually line up that closely to human anatomy. They're a little smaller than us, but their legs are a lot shorter than ours, so the way they would ride horses – well most apes don't ride horses – but the way we made them ride horses is a bit different from humans. Their arms are much longer, much stronger than ours and, in some cases, they've got a lot more mass. The orangutan in particular, Maurice, has a lot more mass. There would be some times where we would sort of dampen her movement a little bit, although she was pretty good about feeling that weight although it wasn't physically on her. You know there were a few little things like that, but the face is the big place where the anatomy is just very different. The brows in particular. We would do things where we would try to sneak in as much of the actor's faces onto their faces, as much as we could get away with, but there are certain things that just don't quite work. You put too much in there and it starts to look like this sort of weird "human-zee". We did a similar thing on Jungle Book with Christopher Walken. We probably pushed it a little further actually with Christopher rather than what we've done with Caesar and Andy. We took signature folds from Christopher's face and his particular lip shape, and we added them pretty strongly onto King Louie. But with Andy we took a lot of the eyebrows, the kind of fold in his upper eyelids and that was just about as much as we could get. You know, we paid attention to some of the lip shapes he would make in particular face expressions, but the ape muzzle is so much bigger and just anatomically different than the human lip and kind of cheek area. You just can't do a direct one-to-one map, it would just look weird. So that kind of figuring out how to compromise and how, you know, figure out when Andy makes this facial expression, how can we figure out emotionally that same reaction off of Caesar's face, what would we need to do, and War was the third film of the trilogy where we had Andy Serkis playing Caesar, and we had a lot of time to work on these face rigs. Well we ended up rebuilding fairly substantial sections of the facial rig just because what Andy did emotionally in the scene, like where he took the kind of level of conflict and extremity with what he was doing with his face was so different from what he'd ever done before, so we ended up rebuilding substantial portions of it. But, yes – so those are the things, the muzzle in particular, the nose, the what we call the naso-labial fold that goes down from the nose down to the corner of the mouth. That's really important for facial expressions like the sneer, even a smile, and, you know, Caesar does a lot of sneering in the movie and that's something we kind of had to find a way around. We cheated a little more naso-labial fold than an ape probably should have, and then we just figured out other ways to carry that same read-through like nose-wrinkling, furling through the nose and cheek area.

JK: Ultimately that is a very important part of the performance though. Because it's taking, as far as I can understand it, it's taking the really strong acting template that the actor provides, but still has to be passed through the animator's eye in order to be able to make sure that it reads well on the ape character.

DL: That is right, that's exactly right.

JK: Because authorship is a really difficult word to define sometimes. It's a contentious word. But what level of authorship do you think is at play there from the animator's perspective?

DL: It's a good question, I guess it depends on how you define authorship. To me I would say the authorship is the actor's, in the same way that you have muralists that have a concept for the mural, might do some sketches, they lay out the template for what this picture needs to look like, and then they hire sometimes dozens of worker painters, who go out, work on the scaffolding and execute the mural.

JK: Somebody like Chuck Close, who, you know, gives an idea of what he wants and then he hires twenty people to help fill it in.

DL: I don't think Chuck works that way.

JK: Oh, maybe I'm thinking of a different person then.

DL: That's one of the things that's kind of amazing about Chuck, is that he actually executes even though he's quadriplegic, has limited functions, he has mechanical systems in his studio to allow him… In fact we were having an interview where he was talking about how – specifically about this topic, where he said he has friends in the art world that ask him "why do you bother doing the execution yourself?". In art right now, for better or for worse, craftmanship is sort of a dirty word. There's certainly more value put into the ideas and who actually bothers with the execution is sort of, in a lot of cases, considered irrelevant. In the interview anyway, his position was: "This is the fun part. Like, I actually enjoy laying down the strokes, making the shapes and the colour, and I don't know why anybody would want to give that up." So, he was actually defensive of that, the execution of things. But anyway, I'm trying to think of who – maybe Christo would be a better example.

JK: Actually, that is a good example.

DL: You've got a guy that comes up with this audacious idea, and then he hires a bunch of engineers, labourer people, and everybody else to execute…

JK: I take it you have a fine arts background?

DL: I have a bachelor in fine arts, in product design and industrial design.

JK: Ok, I have a bachelor in fine arts as well so, when you're talking about these things, I'm like, "Oh, okay, we can talk the same language", which is good.

DL: Yeah! It's a slightly different thing, you know, Andy doesn't – this is part of film-making. This is what, when I take about collaborative art, this is part of what it is. Andy doesn't stick around for the year and a half after we wrap principal photography, to supervise and make sure that his performance is executed in exactly the way he imagined. Nor would you want him to. You know, this is part of what I think allows actors to do what they do. They're not too – if they're in the zone, from what I've seen anyway, the ones who are really good, they don't go back and scrutinize themselves at the monitor after every take. They don't try to, you know, direct the mechanics of their performance. They try to engage emotionally, let it happen, and trust that the director will give them the feedback that they need, and that other actors will give them the feedback that they need to put in a good performance. This is one of those things – it was really interesting watching Gary Oldman on Dawn set versus some of the other actors he was working with. He – it almost seemed effortless: he showed up, he knew every single one of his lines, backwards and forwards, and he would turn on his American accent and you believed every moment of it. You absolutely believed that this character, in this moment, this guy standing here in front of me is coming up with these words and he means them, and there they are.

JK: I felt, when I was interviewing Andy, I thought to myself this guy must be a very generous actor because the reception he was giving me was just phenomenal. As I was talking to him, he was giving back to me. It was just this wonderful back and forth and I said this guy must be really easy to act with, and you know you could just feel it, you know. The charm comes into play as well.

DL: Yeah, that sounds like the Andy I know. Which is one of the reasons why it's unfortunate that he's got, I think, unfairly sucked into a position in this whole debate as, like, cast as the anti-animator. Because, like, I don't think that's his stance at all and in fact, you know the animators that actually work on the show, have enjoyed his performance quite a lot, have enjoyed working with his material, and enjoyed him as a person, you know, those that had the opportunity to be on set.

JK: I got a number of students that work down at Weta right now – former students who one of their biggest heroes was Andy. Because we teach motion capture here, a lot of our guys go down and work in the motion capture department and they got into it in part because of that, And so, yeah, I agree, there is in my opinion, quite a – he's unfairly maligned, but at the same time, I think there have been some things that he said that he didn't mean to be taken the way that they were but have obviously, you know, they – you could be critical of what he said to certain extents. But when I talked to him, I think he was quite contrite and willing to admit what he really meant and what people thought about it.

You've been excellent, Dan. You've pretty much answered all the questions I didn't have to ask you. Because I wrote them down but I didn't have to ask them. There are – one main question and one other follow-up if you want. Something that you've been quoted on saying I believe in the Cinefex article was "A lot of times it's more about the change than it is the rest pose. If you can get the change to read, the character's performance tends to feel like the actor". Which I thought was interesting, because we spent so much time creating meaningful blend shapes and trying to make sure everything looks good, but you're talking about the actual transition between those big poses. I'm wondering if you could expand on that with any examples?

DL: Yeah, so here's a funny thing. So, there's this Fast & Furious 7, right? Where through a tragedy one of the actors passed away while they were filming the movie and the movie went on hiatus for a number of months while the insurance company and everybody else figured out how they were going to finish the movie. And this, of course, was Paul Walker, and the filmmakers consulted with his family and everybody was adamant that Paul would want the movie completed, would want his legacy kind of preserved. And they came up with a plan where his brothers would stand in for him and then they would replace their faces with Paul's face. One of his brothers in particular, Caleb, was a reasonable lookalike and we built, as part of the process, a digital Paul and we also built a Caleb to kind of help validate, match in those performance ticks and just to make sure that lighting and everything was totally matching. The funny thing that happened was we would take a scene where maybe Caleb was sitting on the driver's seat, driving and looking out the window, reacting to something – and we would replace Caleb's face with Paul's face, doing the same thing that Caleb was doing. And what was really interesting was that it didn't look like Paul anymore. Like, a still frame still looked like Paul, but when you played it back, it looked like Caleb. Because they looked close enough to one another that suddenly it wasn't about the actual still appearance of their face – it was the performance that was driving it, and the performance itself bore the signature of Caleb, not Paul. So, what we ended up having to do was go back through a bunch of Paul's old movies and do the Paul-version like the way Paul would've moved his face. These sort of hesitancies, moves and ticks, that kind of movie star swagger that Paul had developed over years of being a movie star and inhabiting that skin – that role was totally different from Caleb who was stepping into his brother's shoes to try and preserve a legacy and, you know, emotionally, you have a totally different – but also a totally different person, so you just behave in a different way. So I think that was part of what I was getting at in that conversation with Joe Fordham for the Cinefex article, was that actually – it's easy to point at a photograph of a person and say "that's the person," and when you look at a video of a person, often you think of the collection of images as being the person. But, in fact, it's the movement between the images – that going from one spot to the next spot and the timing of those movements – that actually is the performance. And the performance is as much the signature of the actor as the facsimile of the actor – the static visual appearance of the actor. And what we're doing with performance capture – we're completely erasing the visual facsimile and preserving the performance. And when you watch Andy Serkis side by side with the final Caesar as appeared in the movie, I think you can see what exactly that process is: all those details of Andy, the feel of something that only Andy could do are carried through to Caesar, so Caesar is doing something that only Andy could do, as well.

JK: You know it's interesting that the notion of the performance actually being between the facial shapes – between the main key poses – is so similar to Norman McLaren's comment about animation not being the key poses but being between them as well. And so, it's kind of a very interesting relationship there. I'll have to go back and find the exact quote because it sorts of escapes me at the moment, but when you're saying this, actually that sounds a lot like McLaren. So ok, I mean that's great.

First off, I wanted to say congratulations to you on the Academy Award, that's more than impressive.

DL: Oh thanks!

JK: That's outstanding. And obviously before that, even the previous nominations for that as well. Do you feel that synthespian performances should be recognized within categories such as Best Actor in a Lead or Supporting Role? Or as part of Best Visual Effects, as it's currently considered? Or should a new category be created? And if so, who should receive it?

DL: My personal opinion is that, you know, if you're going to give an award to somebody like, if you feel like somebody like Gary Oldman playing Churchill in The Darkest Hour is the kind of role that's worth rewarding as a Best Actor kind of performance, even though it's hard to tell that actually Gary Oldman there, then I think that the digital characters can and should be put into the same kind of category. I think it's just a matter of having a role and having an actor play that role come through. The voters can't help but stand up and take notice and recognize it as one of the outstanding performances of the year. But it's one of the challenges, you know. There's so much good acting every year that to break into that, to be one of the five is hard. I absolutely believe that Andy Serkis is one of the finest actors out there working right now, even though until recently we haven't seen much of his own face on the screen. His performances are phenomenal. Just the reactions I've seen, you know… I've had the opportunity of being around when both Rupert Wyatt and Matt Reeves saw Andy acting for the first time, and they were both floored. Both of them were just gobsmacked by how good he is. And you can't underscore enough how critical how the success of these films – how they've hinged on Andy's performance. I think without Andy Serkis, I don't honestly think that we would have had a second or third movie. He's amazing to watch as an actor, so I think it just comes down to somebody like him landing in a role that people can't ignore. Traditionally the kind of acting roles that get recognized with Oscars, nominations, and that sort of thing tend to be not so much the ones that are in blockbusters. The bias of the voting audience – they tend to favour various more independent, smaller, more "housey", roles that are more kind of unique and less engaged with the mass market. But I think visual effects are becoming ubiquitous enough, the techniques are becoming more understood and more widely used. I don't think it'll be long until before you find roles like that, the performance capture roles that make their way into those kinds of movies that have more appeal to those voters.

JK: I think so, definitely, because right now I've been tracking the number of synthespian roles in movies – and that's a fairly broad category of what I'm using in terms of synthespian – since 2010. And as I look at the number of films that have synthespian roles, and the number of synthespian roles within a single movie, it's increasing at a very fast rate. You know, we're getting a lot of movies coming out right now where we're just accepting synthespians as part of the normal cast and I think that's going to be really influential in trying to change the opinion of the awards systems. Eventually, I think you're right, we're going to see people say, "Hey, that's a worthy performance" regardless of how it's constructed.

DL: It's definitely a thorny topic, because I'm speaking strictly from my own perspective on the Apes movies. If we were talking about a Guardians of the Galaxy kind of thing, where you've got a character like Rocket who actually is much more of an amalgamation of a lot of people's input, then suddenly that to me is a totally different thing. I don't know if that's an appropriate kind of Best Actor situation. Would you give a Best Actor credit to somebody that provided the voice for a traditional cell-animated movie? I don't know, maybe if it was really, really good and that's – there's a certain point where you got to hand the decision-making over to the body that's running those awards and those decisions.

JK: As far as this interview goes, I'll make sure that any comments that you make will be encapsulated specifically as referring to the Apes movies, and not in a broader sense of trying to mean that every synthespian should be considered in a particular way.

DL: Oh yeah, sure, yes, that sounds good. Just going back to the Rocket example for a second. This is something where, when they went from Guardians of the Galaxy to Guardians of the Galaxy Vol. 2, it was a different visual effects supervisor. Stéphane Ceretti did the first one, and Chris Townsend did the second one. They had the same director, but the VFX team changed and the animation team, as well. I could see who worked on the first one didn't do a lot of work on the second one. The second one, Weta did work and Framestore had worked on the first one. We had some animators that had come from

Framestore, had come from MPC, that were now working at Weta and we relied on them heavily to help us to kind of crack the code of, like, what is it that James Gunn likes about Rocket from the last movie? What is it that he doesn't even know that he likes that we need to know in order to get these shots working and be true to this kind of Rocket character? Because Chris Townsend was still coming up to speed, as well, with what it was that made that character *that* character. Meanwhile you had, you know, James Gunn's brother, whose first name I've forgotten, but he was used on set to play Rocket – you know he'd have him as a stand-in, act out parts of the scene, sitting on the floor or whatever. You had Bradley Cooper creating the voice and also, we had video reference of him in the sound recording area reading his lines but he was still on book – he was reading from a page in front of him. There was one point where James Gunn asked the animators: "why whenever I see these shots for the first time is Rocket waving his arms around so much?". So, they started seeing the shots with Bradley Cooper with his video reference as picture-in-picture in the submission, and he said, "Oh, okay, that's why." Because Bradley keeps throwing his arms around as he's reading off the page. So, it was definitely on that movie – it was a more traditional animation process where you have a team of dozens of animators that are kind of together trying to figure out what this character – you know what decisions he is going to make for this shot and that shot. And you get the personality of each animator coming through pretty strongly. We joked around here at Weta, like, a lot of times just by looking at two great pieces of finished animation work, again you'll see the signature of the performance, of the animator, live in the scene and be, like, "Oh, yeah, that was a Jack Tema facial shot. Oh, yeah, that was – whomever." There's a lot of specific personalities that can't help but make its way through into those performances. And it's crazy you know, you're watching the same silly racoon bouncing around the screen, and you can tell who set those keyframes.

JK: Does that come across more in those types of films where you're not relying so heavily on something like the facial performance capture that you did for Apes for example?

DL: Yeah, I mean yes and no. I mean the performance capture for Apes is probably worth noting that what we are talking about there is mostly excellent reference, excellent video reference, that then ends up having a team of animators coming behind and set specific key frames for every tick and nuance of those faces. And so, on these movies, absolutely, like, one of our lead facial animators is this guy Jack Tema, and he animates in a particular way and it's a wonder to behold. You see his shots out of the box and they're amazing: they're expressive and they always have this sort of slightly bouncy, slightly exaggerated – he goes big first and then he dials back, and you can pick his shots out of a line-up, you know, without any other clues. So yes, absolutely on Apes we still see that, we still see that in the shots, and you see that in – and, you know, that's not to take anything away from the actor again. But it's just that process when we're talking about how you're going from what the actor's face is doing to trying to get what the ape is doing, the same emotional reaction, beat for beat, tick for tick. It requires specific input from animators who are also making specific decisions through the process. What's interesting is that they are – the whole thing, like what we said, hinges on them – like matching that performance or else the storytelling and the cut falls apart, to the point where, like, at least on this movie, every tear that Caesar sheds – like, Caesar cries a lot, both in this movie and the last movie and the movie before that. Every tear that Caesar sheds, Andy Serkis cried on set – we didn't add any embellishment, tear-wise, to what Andy did. And that's not necessarily true of the human actors, you know. Without divulging any details, we absolutely went in and added a tear or two here and there where we had to help enhance the human actor's performances.

Appendix X. Treplev's Monologue from "The Seagull"

The following is the monologue used in the dramatic performance of this research, excerpted from Anton Chekhov's "The Seagull" and translated by J-C van Itallie (1997).

TREPLEV. (*Pulling petals from a flower.*) She loves me, she loves me not, she loves me, she loves me not, she loves me, she loves me not. (*Laughs.*) You see, my mother doesn't love me. Why should she? She wants excitement, romance and younger men, and here I am already 25. When I'm not around she's only 32 and when I am she's 43 – she hates me for that. Also she knows I don't believe in The Theatre. She adored The Theatre; she believes she's serving both mankind and the Sacred Cause of Art. As far as I'm concerned, Theatre today is boring and trite. The curtain goes up and inside three walls, lit by artificial light, our famous stars, those high priests of art, parade about, showing us how to eat, drink, love, walk and wear our clothes. And then the playwright tries to squeeze a moral out of it, some smug little moral, all cozy, perfectly fit for home consumption, and each play repeats this formula with only infinitesimal variation. That's why I have to run, like Maupassant who ran from the Eiffel Tower because he knew it would cheapen his mind.

SORIN [Line omitted]. But we must have a theatre.

TREPLEV. What we need is a new kind of theatre, Nina [Uncle]. We need new forms, and if we can't have them, let's have no theatre at all. (*Looks at his watch.*) I love my mother, but she leads such a stupid life, always fussing about that writer of her, always getting her name in the paper. It's too boring. Maybe it's just selfishness but I wish I didn't have a famous actress for a mother. I'd be much happier if she were just an ordinary mortal. Can you imagine what it's like, life at home with her, Nina [Uncle], how I feel? The living room is always full of people, all of them famous, writers or actors, and the only nobody is me, and who am I? What am I? I am her son, and that's the only reason they put up with me. I left the University my third year owing to circumstances, as they say, beyond our control, I have no talent, not a kopeck to my name, and according to my passport, I am a "bourgeois of Kiev," that's my social position. My father was a "bourgeois of Kiev" too, only he also happened to be a famous actor. So whenever the famous actors and writers and musicians in my mother's living room noticed me, I felt as if they were measuring my insignificance. I imagined I could read their thoughts, and I was going through agonies.

Appendix XI. Links to Thesis Animations

Basic Object Exercise:

https://vimeo.com/440903340

Basic Object Exercise – Side-by-Side Live Action and Animation:

https://vimeo.com/440906414

Dramatic Performance:

https://vimeo.com/416243575

Dramatic Performance Side-by-Side Live Action and Animation:

https://vimeo.com/440904210

Dramatic Performance Side-by-Side Face Camera and Animation:

https://vimeo.com/440904552

Appendix XII. List of Published Outcomes

Kennedy, J. (2020). Acting and its double: framing acting within a motion capture context. In press.

Kennedy, J. (2020). Cataloguing vactors by performance style in films from 2010. Proceedings of the 9th International Conference on Illustration & Animation (CONFIA), 147-158.

Kennedy, J. (2019). Acting-centred definitions of vactors, synthespians, and digital doubles. Proceedings of the 25th International Symposium on Electronic Art (ISEA), 25, 297-305.

Kennedy, J. (2019). Critiquing the screen presence of synthespian counterparts. Proceedings of the 9th International Conference on Illustration & Animation (CONFIA), 190-202.

Kennedy, J. (2017). The animator's (missing) hand: how practice informs seeing in 3d animation. Proceedings of the 7th International Conference on Illustration & Animation (CONFIA), 298-309.

Kennedy, J. (2015). Beyond the mirror: producing emotionally-authentic facial performance for animation reference. Proceedings of the 7th International Conference on Illustration & Animation (CONFIA), 101-114.