The Virtual Worlds of Cinema

Visual Effects, Simulation, and the Aesthetics of Cinematic Immersion

Paul Sunderland

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy.
Faculty of Arts and Social Sciences.
University of Sydney.
2019
I hereby declare that, except where indicated in the notes, this thesis contains only my own original work.

Abstract

This thesis develops a phenomenology of immersive cinematic spectatorship. During an immersive experience in the cinema, the images, sounds, events, emotions, and characters that form a fictional diegesis become so compelling that our conscious experience of the real world is displaced by a virtual world. Theorists and audiences have long recognized cinema’s ability to momentarily substitute for the lived experience of reality, but it remains an under-theorized aspect of cinematic spectatorship. The first aim of this thesis is therefore to examine these immersive responses to cinema from three perspectives – the formal, the technological, and the neuroscientific – to describe the exact mechanisms through which a spectator’s immersion in a cinematic world is achieved. A second aim is to examine the historical development of the technologies of visual simulation that are used to create these immersive diegetic worlds. My analysis shows a consistent increase in the vividness and transparency of simulative technologies, two factors that are crucial determinants in a spectator’s immersion. In contrast to the cultural anxiety that often surrounds immersive responses to simulative technologies, I examine immersive spectatorship as an aesthetic phenomenon that is central to our engagement with cinema. The ubiquity of narrative – written, verbal, cinematic – shows that the ability to achieve immersion is a fundamental property of the human mind found in cultures diverse in both time and place. This thesis is thus an attempt to illuminate this unique human ability and examine the technologies that allow it to flourish.
Acknowledgements

Thanks firstly to Bruce. You’re an incredible mentor and have been such an amazing support over the past few years. I cannot thank you enough for all your advice and encouragement, and I’ll always be grateful for your enthusiasm about my early writing on Stanley Kubrick.

Thanks also to my family. Your academic and professional achievements have always inspired me, and I’m so grateful for your constant support and encouragement. There are too many of you to mention individually, but you know who you are. I love you all.

And finally, thanks to Monique. You’re amazing, and I love you. And I can’t wait for December!
# Contents

Abstract 3  
Acknowledgements 4  
Contents 5  
Introduction 6  
Visual Effects, Diegetic Immersion, and Hollywood Cinema 9  
The Phenomenology of Immersion 12  
Immersion Through History 17  
Perceptual Immersion 27  
Narrative Immersion 32  
Visual Effects: Technologies of Illusion 37  
Cinema’s Virtual Spaces 42  

Chapter 1 – Spatial Immersion and Hollywood Storyworlds 47  
Medium Transparency 51  
Visual Effects Compositing 53  
Synthetic Images and Realism 57  
Compositing in Early Cinema 68  
Rear Projection and Hitchcockian Modernism 75  
Compositing in the Digital Era 85  

Chapter 2 – Camera Movement and the Immersive Long Take 91  
Spatio-Temporal Immersion 95  
Proprioceptive Aesthetics 100  
Aesthetic Immersion and Distanciation 106  
*The Shining* and the Aesthetics of the Steadicam 112  
The Steadicam and Diegetic Presence 117  
Moving Through Virtual Spaces 126  
Digital Movement 130  
*Gravity*: Virtual Reality Cinema 137  

Chapter 3 – Emotional Immersion: The Response to Character 144  
The Varieties of Aesthetic Immersion 150  
Other-Directed Emotions 153  
The Neurophenomenology of Emotional Immersion 156  
Simulation and its Discontents 167  
The Uncanny Valley 174  

Conclusion – The Future of Digital Immersion 183  
References 188
Introduction

All of us have experienced the unique pleasure of being immersed in a fictional world. During such experiences, the images, sounds, events, emotions, and characters that form a fictional diegesis become so compelling that our immediate surroundings disappear, and the drama of the diegetic world becomes more immediate to us than the real world of our everyday lives. Such responses to fictional worlds can be found in a variety of media, including novels, cinema, virtual reality, poetry, music, and even non-fictional accounts of real events, such as immersive journalism. These immersed responses to media can be found in a variety of historical eras, and in diverse cultural traditions. My concern in this thesis is with the immersive worlds of popular cinema, with a focus on the form of cinema outlined by Bordwell, Thompson and Staiger in *The Classical Hollywood Cinema*.¹ This is a form that uses the elements of cinematic style to maximize the credibility and coherence of the virtual world that exists beyond the screen. It is a form that uses the technologies of visual simulation – the camera, and increasingly, visual effects – to construct a tangible space in order to elicit a specific spectator response. It is a response characterized by the spectator’s relocation away from their immediate environment, and into the diegetic world of the cinema.

This thesis has two main aims. The first is to describe the exact mechanisms through which a spectator’s immersion in a cinematic world is achieved. This aim can be further broken down into three broad approaches: a formalist analysis of the cinematic techniques used by filmmakers to create diegetic worlds; a technological analysis of both the apparatus and the visual effects techniques used to capture and edit the images that make up such worlds; and finally, the application of research from the sciences of the human mind and brain to cinematic phenomena to help explain the cognitive and neuro-physiological responses that occur during immersive experiences. The hope is that by examining immersion from these three perspectives – the formal, the technological, and the neuroscientific – a more comprehensive understanding of the phenomenon can be achieved.

A second aim of this thesis is to examine the historical development of the technologies of visual simulation that are used to create these immersive diegetic worlds. My analysis shows a consistent increase in the vividness and transparency of simulative technologies, two factors that are crucial determinants in a spectator’s immersion. This increase in the power of simulative technologies has been the source of much recent cultural anxiety, much of which focuses on the relationship between representation and reality and the threat posed to this relationship by these technologies. This anxiety generally aligns with particular technological developments, but this concern with the relationship between a representation and the reality it ostensibly represents goes back at least as far as Plato, for whom all representations were acts of deception:

The art of representation is therefore a long way removed from truth, and it is able to reproduce everything because it has little grasp of anything, and that little is of a mere phenomenal appearance. For example, a painter can paint a portrait of a shoemaker or a carpenter or any other craftsman without understanding any of their crafts; yet, if he is skilled enough, his portrait of a carpenter may, at a distance, deceive children or simple people into thinking it is a real carpenter.²

The deceptive power of simulative technologies has been the focus of more recent fictional explorations. For example, the dangers of a perfectly immersive technology appear in Brave New World (1932), Aldous Huxley’s depiction of a dystopian society in which an individual’s every want and desire is immediately satisfied. In this novel, the citizens of the futuristic World State are amused by a sophisticated form of entertainment that has added a range of sensory modalities to cinema’s more conventional visual and auditory aspects. This latest iteration in cinema’s technological evolution is so sensorially engaging that it dulls the ability for critical engagement and rational thought, which, Huxley concludes, helps deprive the citizens of the World State of any kind of meaning or agency in their real lives.³

Real-world immersive technologies have also been the focus of more recent attack. For example, in an article titled Virtual Reality is Addictive and Unhealthy, William H. Davidow

³ A detailed discussion of this aspect of Huxley’s novel can be found in Neil Postman, Amusing Ourselves to Death (London: Penguin, 2005).
warns that ‘as we take bodies and brains adapted to physical space and immerse them in virtual worlds, they are not only unable to cope, they respond in unanticipated ways.’ Among the dangers that Davidow associates with virtual reality are ‘a persistent and worrying increase in levels of cortisol,’ an unnatural release of the neurotransmitter dopamine that ‘drives us to crave more such stimulation,’ and elevations in heart rate and blood pressure that resemble the ‘elevations in cortisol-triggered responses seen in combat veterans.’ Similarly, in *Simulation and Its Discontents*, Sherry Turkle draws attention to some of the problems associated with immersive technologies, focusing in particular on the introduction of simulative technologies to a variety of fields, including architecture and the social and natural sciences:

But immersed in simulation, we are also vulnerable. Sometimes it can be hard to remember all that lies beyond it, or even acknowledge that everything is not captured in it. An older generation fears that young scientists, engineers, and designers are “drunk with code.” A younger generation scrambles to capture their mentors’ tacit knowledge of buildings, bodies, and bombs. From both sides of a generational divide, there is anxiety that in simulation, something important slips away.

But of course, immersive technologies also offer a wide range of practical applications and benefits. Again, Turkle explains:

Immersion has proved its benefits. Architects create buildings that would not have been imagined before they were designed on screens; scientists determine the structure of molecules by manipulating them in virtual space; nuclear explosions are simulated in 3D immersive realities; physicians practice anatomy on digitized humans.

Without denying the importance of the concerns expressed above, I offer in this thesis an alternative view of immersive technologies. I hold immersion in virtual worlds as an aesthetic

---

5 Davidow, ‘Virtual Reality,’
phenomenon that is central to our engagement with art. In the chapters to follow I’ll attempt to show how filmmakers have used immersive ideals to affect us in challenging and profound ways. Perhaps the most surprising aspect of the phenomenon is how easily it can occur. Richard Gerrig, in one of the first detailed examinations of the psychology of reading, notes that even simply reading the word ‘Texas’ can conjure up a range of thoughts, feelings, sensations, and memories, all contributing in some way to ‘transporting’ us, however fleetingly, to an imagined Texas.\footnote{Richard Gerrig, 	extit{Experiencing Narrative Worlds} (Boca Raton, FL: Routledge, 2018), 4, Taylor & Francis eBooks.} The ubiquity of narrative – written, verbal, cinematic – shows that the ability to achieve immersion is a fundamental property of the human mind, a property that evolutionary psychology has recently shown is universal and found in cultures diverse in both time and place.\footnote{See for example Brian Boyd, 	extit{On the Origin of Stories: Evolution, Cognition, and Fiction} (Cambridge, Mass: Belknap Press of Harvard University Press, 2009) and Denis Dutton, 	extit{The Art Instinct: Beauty, Pleasure, & Human Evolution} (USA: Oxford University Press, 2009).} This thesis is thus an attempt to illuminate this unique human ability, and examine the technologies that allow it to happen.

**Visual Effects, Diegetic Immersion, and Hollywood Cinema**

On the popular website the Internet Movie DataBase (IMDB), there is a list of films that claims to represent the ‘25 most immersive worlds in cinema.’\footnote{‘The 25 Most Immersive Worlds in Cinema,’ IMDB, last updated February 26, 2016, http://www.imdb.com/list/ls079103330/#1.} James Cameron’s 	extit{Avatar} (2009) opens the list, followed by the 	extit{Harry Potter} series (2001-2011) and 	extit{The Lord of the Rings} (2001-2003) at numbers 2 and 3, with 	extit{The Hunger Games} (2012-2015), the 	extit{Mad Max} series (1979-2015), and the 	extit{Jurassic Park} franchise (1993-2018) also making the top 10. A short introductory paragraph describes the films as ‘cinematic fantasy worlds that are so completely realized and captivating that they ... make us want to go there.’ The list does not specify the exact formal characteristics of the films that make them immersive, nor does it elaborate on the phenomenology that is at the center of such experiences, but several common characteristics are immediately apparent, helping to form a working definition of the general experience of immersion as understood by the authors. The films conform closely to classical narrative form, with sympathetic goal-oriented characters, psychologically plausible character motivation, and a
sequence of events determined by the laws of cause and effect.\textsuperscript{11} These clearly articulated characters inhabit richly detailed narrative worlds, large parts of which are created and visualized using complex and expensive visual effects technology. These worlds are so believable and captivating that they, according to the authors, compel us to enjoy them. The films were also significant commercial and popular successes, often originating in source material that experienced a similar level of popular appeal, demonstrating a close relationship between immersive cinematic worlds and commercial and popular success. From this brief overview, we can begin to understand immersion as a pleasurable psychological phenomenon characterized by increasing affective, cognitive, and phenomenological engagement with a narrative world. It underlines the enormous box-office success of major Hollywood franchises, and is fundamental to popular engagement with mainstream cinema.\textsuperscript{12}

The terms ‘immersion’ and ‘immersive’ are common signifiers of approval among popular film critics, but as shown in the example above, the phenomenon being referred to, as well as the cinematic and formal characteristics that contribute towards this experience, remain under-theorized. This thesis claims that the experience of immersion is not only fundamental to popular Hollywood cinema, but that it also characterizes many other spectator-media relationships, and enormous value therefore exists in more closely examining the precise mechanisms involved in an immersive experience. To begin to more fully examine the aesthetics of immersion, I’ll examine in detail an excerpt from a review focusing on the opening shot of Alfonso Cuarón’s 2013 film \textit{Gravity}, an enormous commercial success, and a film frequently praised for its ability to produce an immersive response. The scene being described is an elaborate 13-minute sequence, the central narrative line of which focuses on Dr. Stone, played by Sandra Bullock, as she is detached from her spacecraft after it is destroyed by debris:

\begin{quote}
In one continuous shot, the film has not only introduced its central crisis - will Stone survive? - but also completely immersed us in the beauty and
\end{quote}


\textsuperscript{12} Empirical research into the psychology of narrative engagement supports this claim. Green, Brock and Kaufman for example, use empirical research to argue that a key element of an enjoyable media experience is this ability to be transported away from ‘mundane reality and into a story world.’ Melanie C. Green, Timothy C. Brock, and Geoff F. Kaufman, ‘Understanding Media Enjoyment: The Role of Transportation into Narrative Worlds,’ \textit{Communication Theory} 14, no. 4 (2004): 311.
majesty of a dark, pitiless universe. While “Gravity” is hardly the first film to send characters into orbit, few have so powerfully and subjectively evoked the sensation of floating right there with them. As it glides nimbly around the action, the camera induces a deeply pleasurable feeling of weightlessness … that can suddenly turn from exhilarating to terrifying, leaving us gasping for oxygen alongside the characters.¹³

The excerpt highlights two dominant understandings of immersion as understood in both popular and scholarly discourse. The first of these is a physical, sensory involvement with the cinematic image. When the reviewer refers to a ‘sensation of floating,’ he is describing a physiological response resulting directly from the spectator’s relationship with the technology of cinema. Cinema’s formal elements, including the movement of the camera and the visual effects and cinematography used in the construction of an illusory 3-dimensional image, all contribute to an intense and tangible sensation of moving in space alongside the characters. This is an embodied response that is experienced on an immediate level by the spectator. It is primary and non-conceptual, and can be understood as the subjective experience of spatial presence.

The second conception of immersion implicit in the review is an immersion deriving from an intense engagement with the scene’s narrative elements. When the reviewer mentions the ‘central crisis’ of the sequence, which involves the survival prospects of Dr. Stone after her shuttle is destroyed, he is invoking a complex relationship between spectator and character that is central to classical Hollywood storytelling. This relationship is determined by the spectator’s empathy and identification with the character, and hope for a positive outcome to the crisis in which she is involved. The spectator’s immersion in the scene is therefore activated by two immersive narrative responses - empathy and suspense - which stimulate the spectator’s involvement with the drama unfolding on screen. Where perceptual immersion is primary and immediate, narrative immersion is secondary and conceptual, operating largely through cognitive processes resulting from the temporal development of narrative.

This division of immersion into two general categories – an immediate and low-level perceptual response, and a secondary, higher-level cognitive response – reflects broader trends in both popular and scholarly discourse that tend to conceptualize cinematic reception in either one of these two ways. As an audio-visual medium, popular cinema appeals to the spectator’s spatial and sensory responses, but it also requires the more cognitively-demanding processes involved in following a developing narrative. A comprehensive theory of cinematic immersion must therefore account for the full range of immersive responses, including both perceptual and cognitive registers. An important implication for the study of diegetic immersion emerges from this division of cinematic reception into perceptual and cognitive responses: while classical narrative forms are relatively consistent and unchanging, the technology of cinema is dynamic, showing a broad trend towards an increased ability to recreate photographic imagery synthetically using digital technology. As Stephen Prince succinctly puts it: ‘[t]he more perceptually convincing these imaginary worlds can be made to seem, the more virtual and immersive the spaces of story and image become.’ Understanding an immersive experience as a dynamic process involving a complex synthesis of perceptual and cognitive responses as outlined above can provide a framework that may go some way towards conceptualizing the immersive effect of which new cinematic technologies are capable. This is a point to which I will return in each of the chapters to follow. In the remainder of this introduction, I will provide an overview of the different understandings of immersion in the literature, which will provide the conceptual basis for the case studies in the chapters to follow.

The Phenomenology of Immersion

There is a vast literature on the phenomenon of immersion, encompassing research fields as diverse as literary theory, neuroscience, and consumer behavior. This has led to a

---


‘terminological babel’ due to lexical differences between theorists and across disciplines, many of whom use different terms to explore the same phenomenon as it relates to their unique areas of research. In addition to the multiple understandings of ‘immersion’, there is ‘diegetic absorption,’ ‘immediacy,’ ‘re-centering,’ and ‘transportation,’ among others, all of which refer to the same basic experience of an intense engagement with an aesthetic object. And as noted by Marie-Laure Ryan in *Narrative as Virtual Reality*, the application of the term extends beyond aesthetics:

The term immersion has become so popular in contemporary culture that people tend to use it to describe any kind of intensely pleasurable artistic experience or any absorbing activity. In this usage, we can be immersed in a crossword puzzle as well as in a novel, in the writing of a computer program as well as in playing the violin.

The very flexibility and imprecision of the term has almost rendered it useless as a descriptor of the response to cinema I want to describe. I retain it here however, as it probably enjoys the most transdisciplinary consensus, though I acknowledge that the term is not unproblematic. Furthermore, clearly identifying the type of immersive experience I am attempting to describe can help mitigate some of the difficulties represented by the diverse uses of the term. Mapping out the exact structural components of this type of immersive experience highlights its specific characteristics, and helps to clarify the areas of difference from other activities or experiences that are often also described as ‘immersive.’ The kind of diegetic immersion I’m focusing on describes a response to a narrative world, and thus the most basic structural component of this kind of immersion is the existence of a narrative world for the spectator to be immersed in. Activities such as solving a problem or playing a musical instrument

---

20 Gerrig, *Experiencing*.
clearly lack this component, instead involving a degree of interactivity, with the intense focus and attention inherent in such experiences directed towards achieving a specific goal. Such experiences closely resemble what Mihaly Csikszentmihalyi calls ‘flow’, which he defines as a subjective state of being ‘completely involved in something to the point of forgetting time, fatigue, and everything else but the activity itself.’ Flow no doubt shares certain psychological aspects with my notion of diegetic immersion, but the interactivity and goal-directed behavior involved in flow experiences suggests they may be distinct. Flow experiences also generally lack the fundamental component of diegetic immersion - a sense of involvement in a diegetic world. One possibility for the shared psychological characteristics of the two experiences is that diegetic immersion may be a particular form of flow experience. The cognitive activity involved in following a narrative may be sufficient to activate the processes that lead to a flow experience.

The importance of a narrative world in forming diegetic immersion is highlighted by Janet Murray, whose *Hamlet on the Holodeck* includes one of the most widely cited descriptions of this kind of immersive experience. She offers a useful starting point in exploring its complex phenomenology:

> We seek the same feeling from a psychologically immersive experience that we do from a plunge in the ocean or swimming pool: the sensation of being surrounded by a completely other reality, as different as water is from air, that takes over all of our attention, our whole perceptual apparatus. We enjoy the movement out of our familiar world, the feeling of alertness that comes from being in this new place, and the delight that comes from learning to move within it.

This passage is helpful in pointing to the pleasures inherent in immersion in an alternate world, and it introduces important experiential aspects of immersion such as the complete occupation of attention, and the sensation of a movement away from the familiar. But a more comprehensive exploration of immersion as a response to a diegetic world is found in Werner

---


Wolf’s introduction to *Immersion and Distance*, an edited collection exploring ‘aesthetic illusion’, the term Wolf uses to describe the phenomenon.\(^{24}\) He writes that aesthetic illusion consists ‘primarily of a feeling, of variable intensity, of being imaginatively and emotionally immersed in a represented world and of experiencing this world in a way similar (but not identical) to real life.’\(^{25}\) For Wolf, the process is inherently pleasurable, and is marked by a powerful shift away from everyday experience, and into a represented world. It is primarily an imaginative act, and is ‘first and foremost the inner, mental experiencing of (elements of) an imaginative world which is elicited and shaped by an artefact.’\(^{26}\) In aesthetic illusion the narrative world is distinct from the real world, and there is the sense that the latter has been replaced by the former. As Wolf writes:

> Rather we seem to forget about the pastness (where applicable) or about the fact that the artefact once was made by an artist or author, and re-experience the illusionist world as re-presented, as something rendered present to us like real life, and consequently we feel to be in an ongoing presence and present.\(^ {27}\)

This ‘movement into a story world’ metaphor is central to most understandings of immersion.\(^ {28}\) Its importance can also be seen, for example, in the papers presented at *Immersion and the Storyworld*, a symposium held at Oxford University in June 2012. The conference proceedings define immersion as ‘the intricate phenomenon of getting 'lost', 'involved' or 'drawn into' storyworlds created by literature, film and other media.’\(^ {29}\) The movement into a storyworld metaphor is similar to the transportation metaphor developed by Gerrig, which he uses to describe the cognitive shift that occurs during immersion. Gerrig points out that readers become “lost in a book”, moviegoers are surprised when the lights turn on at the end of a film, and

\(^{24}\) Wolf, *Immersion*.


\(^{27}\) Wolf, *Immersion*, 11.

\(^{28}\) Gerrig also notes that the use of metaphor here enables him to ‘refer concretely to otherwise elusive aspects of readers’ experiences’, and that conceptual metaphors are often used in this way ‘to structure domains of experience that cannot be accessed through literal language.’ Gerrig, *Experiencing*, 2.

viewers become caught up in the fates of soap opera characters, and in all these cases, ‘narrative serves to transport an experiencer away from the here and now.’ The metaphor also clarifies that for immersion or transportation to occur, there must be a ‘place’ or ‘world’ to be immersed in or transported to. Gerrig argues that the transportation metaphor goes ‘a long way toward capturing one of the most prominent phenomenological aspects of the experience of narrative worlds.’

To ensure immersion, a narrative world must maintain certain crucial structural aspects, including internal coherence and logical consistency. Any lapse in either of these domains can reveal the constructed nature of the narrative world, threatening the integrity of the immersive moment. This suggests that immersive experiences are inherently unstable: an immersive moment is always under threat from distraction. The immersive moment’s tenuous hold on the spectator’s attention is a result of the spectator’s recognition of the constructed nature of the medium responsible for the immersion. Wolf writes: ‘As opposed to delusions and hallucinations, this constitutive immersion is … counterbalanced by a latent rational distance, which operates owing to the culturally acquired awareness of the difference between representations and reality.’ Wolf here avoids a common problem in conventional use of the term, which often sees immersion as consisting of a totality of experience. Wolf’s framework suggests that it may be more accurate to describe an illusory artefact as ‘more or less likely to induce immersion’, rather than the more absolute ‘immersive.’ An immersive moment therefore involves a complex relationship between spectator and artefact that relies on a fragile and momentary – and entirely voluntary – lapse in the spectator’s critical faculties.

Understanding immersion as a gradable phenomenon in this way also avoids a problem sometimes characterizing popular film criticism, which implies an either-or relationship between immersion and distance, with ‘immersion’ reserved for films that are seen as artistic or aesthetic successes. This dichotomy of course fails to account for films that are deliberately distancing, such as avant garde works, or much (post)modernist cinema, the aim of which is often to expose

31 Ryan also points out that for immersion to take place, ‘the text must offer an expanse to be immersed within, and this expanse, in a blatantly mixed metaphor, is not an ocean but a textual world.’ Ryan, *Virtual Reality*, 69.
illusion, rather than facilitate it. To help disrupt this dichotomy, Ryan proposes four degrees of absorption involved in immersive experiences. The first of these, *concentration*, applies to non-immersive works that offer so much resistance to recentering that the observer fails to be transported to the textual world. This includes the avant-garde works and the modernist and postmodernist cinema mentioned above. The second, *imaginative involvement*, represents the experience of a reader or viewer that is transported into the textual world, but remains able to contemplate it with ‘aesthetic or epistemological detachment.’ The third, *entrancement*, refers to the enthralled and non-reflexive experience of a reader or viewer so caught up in the textual world that ‘language truly disappears.’ This is the aim of the ‘invisible style’ of classical cinema, and the aim of the majority of the case studies I discuss in the following chapters. The fourth, *addiction*, consists of an inability to distinguish the textual world from the real world. Wolf calls this ‘delusion’, and it is marked by the absence of the ‘latent rational distance’ that defines aesthetic illusion.

**Immersion Through History**

Immersion is not specific to Hollywood cinema, and numerous studies have demonstrated that it is in fact a transhistorical and transmedial aesthetic ideal. Perhaps most comprehensive in this regard is Oliver Grau’s book *Virtual Art*, which challenges the common understanding of immersive virtual reality as a modern phenomenon deriving from relatively recent technological developments, and subsequently theorized by figures such as Jaron Lanier who is generally credited with coining the term in the late 1980s. Grau argues instead for a transhistorical understanding of virtual reality as an aesthetic impulse independent of the specific technologies with which it is involved:

---

36 Ryan also calls this level of absorption the ‘Don Quixote syndrome’, in reference to the episode in Cervantes’ novel in which Don Quixote becomes so convinced that a puppet show is real that in trying to help the show’s heroes destroy their enemies he ends up destroying the performers’ set.
The idea of installing an observer in a hermetically closed-off image space of illusion did not make its first appearance with the technical invention of computer-aided virtual realities. On the contrary, virtual reality forms part of the core of the relationship of humans to images… 

[T]he idea goes back at least as far as the classical world, and it now reappears in the immersion strategies of present-day virtual art.\(^{38}\)

Grau traces immersive virtual realities back to classical wall paintings such as Pompeii’s *Great Frieze in the Villa dei Misteri*, created around 60BC. Such paintings surrounded the observer with images, producing a 360-degree image space and creating ‘an illusion of being *in the picture*, inside an image space and its illusionary events.’\(^{39}\) (italics original). The medieval cathedral has often been identified as an important development in the history of immersive aesthetics. In *Shivers Down Your Spine*, Alison Griffiths describes the medieval cathedral as ‘a hugely significant pre-cinematic site of immersive viewing experiences.’\(^{40}\) By placing the spectator at the center of a spectacular and all-encompassing image space, with spectacular imagery lining the walls, stained-glass windows, and frescoes that blurred the distinction between pictorial and physical space, the cathedral privileged a mode of vision characterized by wonder and awe, and tied to a heightened religious experience for the onlooker.\(^{41}\) The upward gaze, spectacular imagery, and a heightened sense of spiritual communion with the divine, are, Griffith writes, ‘vital elements in a revered gaze that is transhistorical and that has transmuted today into theme park rides, IMAX films … and other jaw-dropping spectacular experiences.’\(^{42}\)

Like their modern counterparts, such historical virtual realities were heavily dependent on available technology as well as contemporary knowledge of the laws of visual representation. One of the most significant developments in the history of this knowledge was the discovery of the laws of perspective during the Renaissance. According to Grau, the principles of perspective gave ‘strategies of immersion … a tremendous boost, for they allowed artists to portray

\(^{38}\) Grau, *Virtual*, 4-5.
\(^{39}\) Grau, *Virtual*, 25.
\(^{42}\) Griffiths, *Shivers*, 35.
convincingly much that formerly could only be alluded to."43 Prior to the discovery, visual representation was symbolic, conveying meaning indirectly through suggestion and symbol, rather than iconically. This visual strategy relied on a spectator’s knowledge of the contents of the representation for full understanding of its intended meaning. The laws of perspective were such a significant development in the history of visual representation because they allowed for the objective recreation of a scene ‘as it might appear to the eye.’44 Furthermore, by enabling a 2-dimensional medium to represent a third dimension, perspective provided the depth that is essential for the physical projection of the spectator into the virtual image space.45 Ryan elaborates on the importance of depth perception in the context of immersive aesthetics:

This projection opens up a depth that assigns spatial coordinates - the center of projection, or physical point of view - to the body of the spectator. Perspective painting immerses a virtual body in an environment that stretches in imagination far beyond the confines of the canvas. From its spatial point of view, the embodied gaze of the spectator experiences the depicted objects as virtually present, though the flat surface of the painting erects an invisible wall that prevents physical interaction.46

The 19th century was significant for two immersive aesthetic modes that neatly align with the two forms of immersion - perceptual and narrative - outlined above. The panorama, a popular 19th Century entertainment, aimed at maximizing perceptual immersion by surrounding spectators with a totality of images, with the represented world completely replacing the real world. The panorama placed the spectator at the center of a 360-degree image space, with a 3-

43 Grau, Virtual, 37.
44 Grau, Virtual, 37.
45 The laws of perspective are also fundamental to contemporary 3D animation computer software programs. William Mitchell points out that as early as the 1960s, pioneering computer-graphics researchers developed perspective-construction algorithms executable by computer, an achievement ‘as momentous, in its way, as Brunelleschi’s perspective demonstration.’ William Mitchell, The Reconfigured Eye: Visual Truth in the Post-Photographic Era, (MIT Press, 1994), 118. Mitchell also suggests that this enabled two functions fundamental to the 3D animation and digital compositing that are central to modern visual effects: firstly, the creation of perspective images that are often indistinguishable from high-quality color photographs, and secondly, to allow synthesized objects to be convincingly inserted into real scenes.
46 Ryan, Virtual Reality, p. 3.
dimensional foreground extending out towards a 2-dimensional canvas that wrapped around the interior circumference of the space (see Figure 1). An imperceptible connection between the foreground action and the background image ensured that the illusion of a coherent and consistent space remained intact. Grau writes that the panorama ‘creates a sense of journeying through space and time - a complete universe of illusion.’

It was an enormously popular form of entertainment during the second half of the 19th Century, with 100 million people attending European and American panoramas between 1870 and 1900. It was an alignment of technology and art that, according to Grau, ‘achieved hitherto unknown dimensions of illusionary effect.’

The realist novel, another popular 19th Century entertainment epitomized by the work of writers such as Charles Dickens, Gustave Flaubert, and George Eliot, perfected classical narrative form, foregrounding character, setting, and drama, and removing technique and style from the reader’s attention. In this literary tradition, the structure of the work is subordinate to its contents, with style and technique aimed at maximizing an immersive relationship with the

---

48 Grau, Historical, 367.
49 Grau, Historical, 366.
50 Ryan suggests that eighteenth century narrative style was more flexible in its adherence to an immersive ideal. ‘[O]n the one hand,’ she writes, ‘it cultivated illusionist effects by simulating nonfictional narrative modes (memoirs, letters, autobiographies); on the other hand, it held immersion in check through a playful, intrusive narrative style that directed attention back and forth from the story told to the storytelling act.’ She argues that ‘the aesthetics of the nineteenth-century novel tipped this balance in favour of the storyworld.’ Ryan, Virtual Reality, 14-15.
narrative world. Werner Wolf describes this form as ‘comparatively inconspicuous, serving mainly to transmit the storyworlds and to support their consistency and lifelikeness.’ The realist novel also focused on characterization, emphasizing psychological depth and credibility, as well as causally motivated action. As Ryan writes, the realist novel ‘penetrated the mind of the characters, transported the reader into a virtual body located on the scene of the action, and turned her into the direct witness of events, both mental and physical, that seemed to be telling themselves.’ By relegating style and technique to the background of the reader’s attention, and by maximizing the psychological depth and credibility of the characters, the realist novel has, Wolf points out, ‘always been accorded a particularly high potential for eliciting illusionist immersion.’

This brief historical overview of some of the major forms of immersive art shows that immersion is a transhistorical phenomenon not limited to a particular period nor specific to a particular technological apparatus. Immersion has rather been an aesthetic ideal that has existed throughout the history of visual and literary representation. During significant art-historical moments, Grau writes, ‘extraordinary efforts were made to produce maximum illusion with the technical means at hand.’ Contemporary immersive cinema can thus be understood in the context of a transhistorical goal of maximizing illusion and more fully immersing the spectator in a virtual space. As Grau suggests, major cinematic immersive technologies including Cineorama, Sensorama, Expanded Cinema, 3-D, Omnimax, and IMAX, are just more recent manifestations of this long history of artistic attempts to ‘integrate the image and the observer.’

This evidence for immersion as a transhistorical aesthetic ideal points to the value in including a cognitive and neuro-physiological explanation of immersion, in addition to

---

51 This does not suggest there was not significant diversity in the approaches to representing the real world among the writers of this period. Alison Byerly, for example, points out that ‘the novels of Charlotte Bronte, William Thackeray, George Eliot, and Thomas Hardy are filled with both explicit references to artworks that have a function within the narrative – portraits, caricatures, charades, musical performances – and metaphors that implicitly compare the novelists’ own representation to specific forms of art. Insistent reminders of the disjunction between art and life, these artistic references threaten to sabotage the realist claim to unmediated representation.’ Alison Byerly, Realism, Representation, and the Arts in Nineteenth-Century Literature, (Cambridge University Press, 1997), 2.


53 Ryan, Virtual Reality, 4.


55 Grau, Virtual, 5.

56 Grau, Virtual, 5.
accounting for its technological, formal, and historical aspects. Identifying the phenomenology of an immersive experience is clearly a challenge, as the phenomenon takes place entirely inside of the spectator’s subjective experience. However, an increasingly valuable body of research from the human sciences is starting to provide answers to this difficult question. Research from several fields including evolutionary psychology, ecological film theory, embodied cognition, and neuroscience, is proving particularly helpful in uncovering the precise cognitive and neuro-physiological shifts that occur during an immersive experience.\textsuperscript{57} Most of these approaches adopt an evolutionary account of aesthetics, which uses our evolutionary history to explain the mechanisms involved in the appreciation of art.\textsuperscript{58} According to this account, the processes of natural selection have determined the basic functioning of the human brain, a fact that can be used to answer questions about the art that audiences enjoy, as well as the experiences that art produces, including immersed responses to fictional representations of diegetic worlds. In \textit{The Art Instinct}, Denis Dutton explains the human ability to understand fictional worlds from an evolutionary perspective:

> Far from being derived from sets of cultural conventions, the enjoyment of fiction shows clear evidence of Darwinian adaptation, for instance, in how even very young children can rationally deal with the make-believe aspects of stories, distinguishing story-worlds from each other and from reality with a high degree of innate sophistication.\textsuperscript{59}

According to this account, the ability to appreciate fictional worlds is an innate ability of the human brain, deriving from its unique architecture and developed over hundreds of thousands of years of evolutionary history. This idea is central to ecological film theory, according to which the same perceptual apparatus is used to understand and navigate both the real world and


\textsuperscript{59} Dutton, \textit{Instinct}, 5.
cinematic worlds. According to Joseph Anderson and Barbara Anderson, the two theorists most responsible for introducing ecological theory into film studies, cinema is essentially an illusion that the human perceptual system has not evolved to accommodate:

For example, there was no way that the forces of evolutionary development could have anticipated that humans sometime in the future would go to all the trouble to … construct an entire fictional world composed only of patterns of light as in a motion picture.\(^6^0\)

When faced with such constructions, they write, ‘the visual system simply processes them as it would the natural world.’\(^6^1\) Similarly, Torben Grodal argues that as a result of our evolutionary history we have a natural tendency to believe in the reality-status of cinematic representations. As the fundamental architecture of the brain developed at a time when fictional representations did not exist, fictional representation has not been a factor in determining how the brain processes visual and auditory perceptions. This means that lower-level cognitive processes receive cinematic images as though they are real images of an environment, rather than created fictions. The evaluation of the reality status of representations - our recognition that a representation is in fact a representation – is, according to Grodal, a result of higher order cognitive processes.\(^6^2\) Lower-level cognitive processes receive cinematic images as though they derive from the real-world, which accounts for cinema’s powerful reality effect, and contributes to the sense of immersion in the diegetic world.

A correspondence between real-world perception and cinematic perception is also found in neuro-scientific accounts of spectatorship. In an article proposing the application of cognitive neuroscience to the study of film reception and aesthetics, cognitive neuroscientist Vittorio Gallese and film theorist Michele Guerra write that ‘[r]ecent studies within cognitive film theory, visual psychology and neuroscience bring out strong evidence of a continuity between perceiving scenes in movies and in the world, as the dynamics of attention, spatial cognition and action are


\(^{62}\) Grodal, Visions, 185.
very similar in direct experience and mediated experience.\textsuperscript{63} They argue that this correspondence is largely responsible for cinema’s powerful ‘reality effect.’\textsuperscript{64} These accounts are of course heavily dependent on the representations in question, and a significant difference exists between the cognitive processes involved in the evaluation of the reality-status of a photorealistic depiction and a stylized animation, for example. But these theories do help to explain the powerful sense of realism that so often accompanies cinematic representation, and helps to illuminate the sense of immersion that accompanies photorealistic depictions of virtual spaces in films such as \textit{Gravity}.

A further application of neuroscience to the study of cinematic spectatorship involves the use of functional magnetic resonance imaging (fMRI) to observe and record the brain activity of viewers as they view a film. In his introduction to \textit{Psychocinematics}, an edited collection of articles exploring such an approach, Arthur Shimamura writes that ‘the advent of fMRI has allowed scientists to assess on a moment-to-moment basis brain activations in response to sensory stimuli.’\textsuperscript{65} By observing and recording the brain activity of spectators, fMRI allows researchers to ‘identify brain regions that are specifically active while viewing moving pictures.’\textsuperscript{66} In one study, researchers used fMRI in combination with inter-subject correlation analysis (ISC) to assess the similarities across multiple spectators as they viewed select scenes.\textsuperscript{67} Participants were shown filmed scenes from four different sources, including an episode of \textit{Alfred Hitchcock Presents}, Sergio Leone’s \textit{The Good, The Bad, and The Ugly}, an episode from the TV show \textit{Curb Your Enthusiasm}, and an unstructured, unedited shot of Washington Square Park. The authors of the study found low ISCs for both the \textit{Curb Your Enthusiasm} segment and the unedited shot of Washington Square Park, while the ISC for the \textit{The Good, The Bad, and The

\textsuperscript{64} Gallese and Guerra, ‘Embodying,’ 183.
\textsuperscript{66} Shimamura, \textit{Psychocinematics}, 15.
\textsuperscript{67} ISC is a means of measuring and comparing brain activity between multiple participants. The authors of the study explain its application to cinema spectatorship: ‘In cinema, some films (or films’ segments) lead most viewers through a similar sequence of perceptual, emotional, and cognitive states. Such a tight grip on viewers’ minds will be reflected in the similarity of the brain activity (high ISC) across most viewers. By contrast, other films exert (either intentionally or unintentionally) less control over viewers’ responses during movie watching (e.g., less control of viewers’ emotions or thoughts). In such cases we expect that there will be less control over viewer’s brain activity; that is, more variability across viewers (low ISC).’ Hasson, ‘Neurocinematics,’ 3.
Ugly segment was extensive. Perhaps most significantly, the highest recorded ISC was for the segment from Alfred Hitchcock Presents, which, the authors write, ‘evoked similar responses across all viewers in over 65 percent of the cortex, indicating a high level of control of this particular episode on viewers’ minds.’ Though indicating little in relation to aesthetics, such findings may have important implications for the study of immersion and the kind of attention and engagement that is central to it. These findings may, the authors of the study write, ‘serve as an objective scientific measurement for assessing the effect of distinctive styles of filmmaking upon the brain, and therefore substantiate theoretical claims made in relation to them.’ The authors conclude that ‘we may speculate that part of the mesmerizing power of movies stems from their ability to take control of viewers’ minds, and that viewers often seek and enjoy such control because it allows them to become deeply absorbed (and mentally engaged) in the movie.’

These approaches to immersion offer compelling and encouraging hopes for uncovering the cognitive and neurophysiological aspects of immersion, and for this reason I will make use of them where appropriate in the chapters to follow. My focus however, remains for the most part on the technological and historical aspects of immersion, an approach that must also consider the individual and unique responses to immersive media, while describing more general aesthetic tendencies. Immersion is a transhistorical phenomenon, as shown above, but there of course exists a degree of variation in the reception of immersive media by different individuals from different historical and cultural contexts. A variety of sociocultural factors could render an aesthetic object highly immersive for one individual or group, while distancing others. Some of the case studies in the chapters that follow attempt to address this issue by considering the cultural contexts of the various eras in which the specific films were produced and released. In other cases, I make use of the ‘average recipient’, a theoretical construct Wolf outlines in Immersion and Distance in an attempt to overcome the methodologically problematic factor of the individual recipient ‘in whom the aesthetic illusion is supposed to take place.’

70 Hasson, ‘Neurocinematics,’ 16.
71 Hasson, ‘Neurocinematics,’ 17.
72 Wolf, Immersion, 27.
outlines the ‘anthropological constants’ that play a role in immersive responses, which include our propensity for role-playing and other games, our curiosity for new information, and our capacity for empathy. These factors are clearly transcultural and transhistorical, and they all enable us to ‘immerse ourselves in others’ and our own (past) experiences, be they actual, represented or merely imagined.\textsuperscript{73} However, Wolf acknowledges that even if one accounts for these anthropological constants, immersion is ‘still heavily dependent on elusive individual factors which have to be added to the anthropological groundwork and may substantially vary from recipient to recipient.’\textsuperscript{74} (italics original). Among these factors are the recipient’s range of experience, age, gender, interests, cultural background, and the ‘ability to ‘read’ works of art aesthetically.’\textsuperscript{75} These individual differences no doubt vary substantially from culture to culture, from historical era to historical era, and from individual to individual, problematizing any attempt to describe a generalized immersive response to a particular artefact.

As a way of overcoming the problems presented by these individual differences, Wolf introduces the concept of an ‘average recipient’, which allows for the examination of the immersive qualities of artworks while attempting to account for the more diverse responses from varied audiences. For Wolf, the ‘average recipient’ is someone who is ‘not only able but also prepared to ‘willingly suspend disbelief’ when confronted with illusionistic artifacts.’\textsuperscript{76} This recipient must be capable of understanding the ‘denotational as well as connotational information stored in the respective work,’ and must be free from any concerns that may divert attention away from the representational contents and towards technical details of form, ‘as can be the case with researchers in film theory watching a film for professional purposes.’\textsuperscript{77} This average recipient allows for the examination of general immersive tendencies of various artefacts, without denying the possibility of unique and varied individual responses. As Wolf writes, ‘[g]iven similar recipients and similar reception contexts, representations will appear more or less illusionist according to ‘intracompositional’ factors.’\textsuperscript{78} Where examining the unique and varied responses of different audiences to immersive technologies is impossible, the ‘average

\textsuperscript{73} Wolf, \textit{Immersion}, 29.
\textsuperscript{74} Wolf, \textit{Immersion}, 29.
\textsuperscript{75} Wolf, \textit{Immersion}, 29.
\textsuperscript{76} Wolf, \textit{Immersion}, 31.
\textsuperscript{77} Wolf, \textit{Immersion}, 31.
\textsuperscript{78} Wolf, \textit{Immersion}, 30-31.
recipient’ offers one way of examining the more general immersive tendencies, or the intracompositional factors, that contribute towards an immersive experience.

**Perceptual Immersion**

One of the more general immersive tendencies in cinema is the use of cinematic technology to stimulate the senses. This form of perceptual immersion comes in a variety of forms, but these variations are all derived from the immediate sensory reception of the images and sounds of cinema. As an example, Michael Snow’s *Wavelength* (1967), an avant-garde work that avoids any attempt to re-center the viewer in a storyworld, achieves an immersive response through a powerful multi-sensory address. The film consists of a 45-minute zoom, starting at one end of an inner-city apartment, and moves slowly towards a photograph on a wall at the other end of the room. A minimalist score of shrill tones becomes increasingly intense as the zoom slowly traverses the space. The film completely rejects realism in its depiction of space, and the image abruptly shifts appearance, moving from a naturalistic color scheme to various shades of green, violet, blue, and an eerie red glow. Narrative concerns are marginalized in favor of an intense sensory experience, which Snow claims is a ‘summation of [his] nervous system, religious inklings and aesthetic ideas.’\(^79\) Despite the film’s neglect of narrative concerns such as plot and character identification, and its rejection of a realist aesthetic, its powerful sensory-motor register makes it a good example of cinema’s affective potential.

A form of direct sensory immersion relies on the technology of cinema to maximize the intensity with which a spectator experiences cinema’s sounds and images. Tim Recuber describes this form of ‘immersion cinema’ as cinema that employs ‘a host of technological and architectural devices, such as wide screens, large-format projection, digital surround sound, and stadium-style seating, to create a more absorbing and intense sensory experience for viewers.’\(^80\) Like the example from *Wavelength* described above, this cinema involves a focus on the

---


spectator’s sensory reception over the more cognitively-demanding processes involved in understanding a narrative.\textsuperscript{81} Cinematic technology is fundamental to this kind of immersion cinema, and the drive to more intensely involve the spectator in the cinematic image often drives the development and introduction of innovative cinematic technologies. For example, in \textit{Cinematic Appeals: The Experience of New Movie Technologies}, Ariel Rogers details the introduction of major cinematic technologies such as the various widescreen processes introduced in the 1950s, and notes that they were often associated with an aesthetics of immersion. ‘The widescreen frame,’ she writes, ‘which lies outside the viewer’s central field of vision, presents an image that seems to constitute the viewer’s own physical space.’\textsuperscript{82} By attempting to blur the boundaries between image and theatre space, ‘widescreen rendered screen space as the viewer’s environment.’\textsuperscript{83} In this case, innovative technology was central to an immersion that was promoted as a novel form of cinematic experience.

Blurring the boundary separating image space and theatre space in this way was an attempt to maximize the spectator’s sense of relocation away from the immediate world and into the narrative world depicted on screen. This sense of relocation is what the media and communications literature refers to as ‘presence.’ Matthew Lombard and Theresa Ditton, in attempting to arrive at a coherent definition of the concept of presence, conducted an extensive overview of its various understandings in the media and communications literature, ultimately defining it as the ‘perceptual illusion of nonmediation.’\textsuperscript{84} An illusion of nonmediation occurs, they write, ‘when a person fails to perceive or acknowledge the existence of a medium in his/her communication environment and responds as he/she would if the medium were not there.’\textsuperscript{85} In a conception more directly related to representational works, they write that the medium ‘can appear to be invisible or transparent and function as would a large open window, with the

\textsuperscript{81} Recuber positions his study of ‘immersion cinema’ in opposition to psychoanalytic accounts of film spectatorship, which he argues ‘fail to consider the implications of a different, more material factor - the lived and experienced spaces of cinema and the ways that changes in those spaces over time have affected the functions of the apparatus.’ Recuber, ‘Immersion,’ 316.

\textsuperscript{82} Ariel Rogers, \textit{Cinematic Appeals: The Experience of New Movie Technologies} (Columbia University Press, 2013), 93.

\textsuperscript{83} Rogers, \textit{Cinematic}, 92.


\textsuperscript{85} Lombard and Ditton, ‘Presence’.
medium user and the medium content (objects and entities) sharing the same physical environment.' They argue that a number of emerging technologies, including virtual reality, simulation rides, video conferencing, home theater and high definition television, ‘are designed to provide media users with an illusion that a mediated experience is not mediated.’

Presence clearly has a role in the response to many forms of cinema, particularly those that rely on innovative cinematic technology to provide novel visual experiences, such as the widescreen processes mentioned above. As Rogers notes, novel forms of cinematic technology have often been associated with an increase in immersion, and the digital visual effects of contemporary Hollywood cinema are another later technology often directed towards the creation of a sense of presence. One of the first key papers to examine the concept in detail was Jonathon Steuer’s *Defining Virtual Reality: Dimensions Determining Telepresence*, originally published 1992, a time when virtual reality was enjoying an unprecedented level of public attention. In this paper, Steuer identifies presence as a key component of virtual reality, and suggests shifting the focus from the technological hardware involved in virtual reality, to the psychological processes. Doing so can provide ‘a means for examining VR in relation to other types of mediated experience.’ Steuer’s focus on the psychological aspects of virtual reality provides a valuable set of concepts that can help to elucidate important aspects of cinematic spectatorship and the role of presence in forming an immersive cinematic experience. One of the important concepts he identifies is *vividness*, a key factor in determining the degree of presence experienced by a spectator undergoing an immersive virtual reality experience. Vividness refers to ‘the ability of a technology to produce a sensorially rich mediated environment.’ It is a measure of the ‘representational richness of a mediated environment as defined by its formal features, that is, the way an environment presents information to the senses.’ Steuer furthermore identifies two important variables that determine the degree of vividness. The first of these, *sensory breadth*, refers to the number of perceptual dimensions simultaneously addressed.

---

86 Lombard and Ditton, ‘Presence’.
87 Lombard and Ditton, ‘Presence’.
89 Steuer, ‘Defining,’ 80.
90 Steuer, ‘Defining,’ 81.
The greater the number of perceptual channels addressed the more presence experienced, and traditional media, including cinema, are relatively low in breadth, ‘relying primarily on the visual and auditory channels.’\textsuperscript{91} Several attempts have therefore been made to address cinema’s perceived deficiency in this area, perhaps most notoriously in the 1950s when “AromaRama” and “Smell-o-Vision” attempted to incorporate an olfactory component into the cinematic experience.\textsuperscript{92} Other attempts from this period include \textit{The Tingler} (1959), a film which included a gimmick whereby some members of the audience were shocked by devices attached to their seats, adding a haptic element to the cinematic experience. Similarly, several 1970s films including \textit{Earthquake} (1974) and \textit{Midway} (1976) used “Sensurround”, a system developed by Universal, which, according to Barry Salt, ‘relayed very low frequency sound, mostly below audibility.’\textsuperscript{93} The “Sensurround” effect was intended to be felt, rather than heard, and was triggered at “appropriate scenes in the film.”\textsuperscript{94} Salt notes that the process failed to catch on, and was soon replaced by developments in the Dolby system of sound recording.

My primary focus in this thesis is on the visual aspects of immersion, but sound design undoubtably plays a crucial role in a fictional world’s impression of reality and is thus an essential component in facilitating spectator immersion. Cinema has of course always included an auditory component, with screenings of silent cinema accompanied by live orchestras and pre-recorded soundtracks. However, the addition of synchronized sound in \textit{The Jazz Singer} (1927) represents a significant development in the spectator’s relationship to the diegesis. Bazin would later describe this move from silent to sound cinema as less an ‘aesthetic revolution’\textsuperscript{95} than the fulfillment of cinema’s natural tendency towards realism in the representation of a diegetic world. For Bazin, the addition of sound to the image meant that a crucial element of everyday perception that had so far been denied from cinema was now accounted for. Using Steuer’s taxonomy of presence, we can understand the alignment of sound and image as a significant

\textsuperscript{91} Steuer, ‘Defining,’ 82.


\textsuperscript{94} Salt, \textit{Film Style}, 319.

\textsuperscript{95} Bazin, ‘Evolution,’ 23.
technological milestone because it provided an additional perceptual channel. Adding an auditory channel that corresponds to the visual channel entailed an increase in the spectator’s presence, and it thus became a central component of classical style.

Steuer’s second important determinant of vividness, *sensory depth*, refers to the resolution of each of the perceptual channels. Though classical forms of cinema are confined to two perceptual channels – the visual and the auditory – no such formal restraint limits the depth of these channels. There are however technological limitations on depth, with a diminished resolution being a necessary fact of technological mediation when compared with natural real-world perception. As Steuer pointed out in 1992, ‘No currently available auditory or visual recording systems match the capabilities of the human auditory or visual system.’96 Advances in digital technology have succeeded in minimizing this gap, as subsequent chapters will show, but limitations on resolution remain a challenge for the technicians and engineers involved in the development of cinematic technology.

The resolution available in each of the perceptual channels has important implications for an immersive aesthetics. An image that more closely resembles a real-world counterpart can more readily produce the ‘perceptual illusion of nonmediation’ that Lombard and Ditton identify as essential to presence. Conversely, an image lacking in resolution can draw attention to the medium, distracting the viewer away from the image’s contents and presenting a barrier to the complete immersion in the intended illusion. Because of this, classical forms of cinema have tended towards the adoption of technologies that increase resolution. The chapters to follow will explore this trajectory in relation to visual effects, but in an article on the immersive implications of sound design, Mark S. Ward has argued that sound technologies have undergone a similar trajectory:

Similarly, the impulse toward immersion is also traceable through successive generations of sound technologies, from optical to magnetic to digital. Sound design that was once crushingly limited to the bandwidth

---

96 Steuer, ‘Defining,’ 83.
of a single monaural channel may now be presented as stereo, 5.1 or some other surround sound format.  

He concludes that the development of sound design praxis ‘exhibits a tendency toward techniques and technologies with greater and greater immersive characteristics in the production of an aesthetic experience.’ Because of the implications of resolution limitations on presence, research and development in both visual effects and sound technology continues to be directed towards minimizing the difference between the resolution achievable with imaging technologies, photographic or otherwise, and the resolution of the basic human perceptual apparatus. These technological developments will continue to have a significant role in generating the sense of presence that is conducive to immersion in cinematic storyworlds.

Narrative Immersion

Despite wide public attention and significant corporate interest, virtual reality has yet to achieve a popular or commercially viable storytelling ability. Current attempts at achieving ‘cinematic virtual reality’ generally use a head-mounted display to place the spectator at the centre of a 360-degree video, surrounding them in a totality of images. This is an expansion of the frame of traditional cinema reminiscent of the panorama or the 360-degree image spaces.

99 Steuer also mentions the fascinating but terrifying scenario involving technological advances in vividness that eventually result in systems capable of passing a ‘perceptual Turing test.’ He writes: ‘The ramifications of media systems whose representations are perceptually indistinguishable from their real-world counterparts is both exciting and terrifying – exciting because of the possibilities afforded by such systems to experience distant and nonexistent worlds, yet terrifying because of the blurring of distinction between representation and reality.’ Steuer, ‘Defining,’ 84.
100 In a report published in 2016, Goldman Sachs predicted that revenue from VR / AR technology over the next ten years could range from $23bn in a ‘delayed uptake’ scenario, to $182bn in an ‘accelerated uptake’ scenario. The areas most expected to drive the market, according to the authors, include videogames, live events, video entertainment, healthcare, real estate, retail, education, engineering, and the military. Heather Bellini, Virtual & Augmented Reality: Understanding the Race for the Next Computing Platform, (Goldman Sachs, January 13, 2016). https://www.goldmansachs.com/insights/pages/technology-driving-innovation-folder/virtual-and-augmented-reality/report.pdf.
discussed above. The spectator’s freedom to direct their gaze anywhere in the space comes, however, at the cost of the sequential ordering of images upon which cinematic storytelling relies. Virtual reality is a medium of presence, with the spectator’s interaction with the contents of the medium occurring in a spatial domain. Cinema, in contrast, is primarily a temporal medium, and early in its history innovative filmmakers learned that placing images in sequence is an effective means of telling a story. Today it remains a storytelling medium *par excellence*. A central claim of this thesis, however, is that despite cinema’s basis as a temporal medium, its history shows a gradual increase in its ability to simulate a diegetic space. The technological developments that have increased the vividness of these simulations have allowed increasingly powerful illusions of spatial presence. Cinema has thus increasingly incorporated the aesthetics of VR into its storytelling, as subsequent chapters will show.

Cinema does, however, remain at essence a temporal medium, and the primary source of immersion in popular forms of cinema is the temporal development of narrative. This section moves away from the immediacy of perceptual immersion to examine the ongoing construction of a narrative world in the mind of the spectator. The construction of a diegetic world through narrative, though not as perceptually immediate as the form of immersion outlined in the previous section, is still a powerful stimulus for an immersive cinematic experience. As Janet Murray writes: ‘[a] stirring narrative in any medium can be experienced as a virtual reality because our brains are programmed to tune into stories with an intensity that can obliterate the world around us.’\(^{102}\) In *Narrative as Virtual Reality*, Ryan offers a comprehensive analysis of the role of narrative in facilitating an immersive experience. Focusing on literary fiction, she outlines three key immersive responses to fictional worlds: ‘spatial immersion, the response to setting; temporal immersion, the response to story; and emotional immersion, the response to characters.’\(^{103}\) The first of these, spatial immersion, involves the reader’s construction of the spatial aspects of a coherent and internally consistent diegetic world. Ryan sees it as fundamental to the pleasures of narrative fiction, and argues as an example that ‘[t]he profusion of details with which J. R. R. Tolkien imagined Middle Earth is a strong factor in the cult that his fantastic world generated.’\(^{104}\) Ryan for the most part limits her discussion to spatial immersion in literary

---

102 Murray, *Hamlet*, 98.
103 Ryan, *Virtual Reality*, 90.
fiction, so the phenomenology she describes has only minimal application to more perceptually immersive media such as cinema. As she herself acknowledges, ‘a picture transports the spectator almost instantly into a landscape’, whereas language ‘can only describe it detail by detail, bringing it slowly into the reader’s mind.’\textsuperscript{105} She adds that for all ‘the affinities between storytelling and sense of place, language is not an ideal medium to convey the presence of a place.’\textsuperscript{106}

Temporal immersion consists of the reader’s interest and involvement in the progression of a narrative scenario from many potential developments into one possible outcome. Ryan points out that ‘[i]t is because past events cast a shadow on the future and restrict the range of what can happen next that we perceive narrative lines and experience temporal immersion.’\textsuperscript{107} She divides temporal immersion into three main viewer responses: suspense, curiosity, and surprise, all of which depend upon the timely release of information. She notes that while suspense and curiosity are similar as they both rely on the desire to know, surprise is clearly different. To highlight this distinction, she refers to Hitchcock’s well-known account of cinematic suspense:

\begin{quote}
We are now having a very innocent little chat. Let us suppose that there is a bomb underneath this table between us. Nothing happens, and then all of a sudden, “Boom!”, there is an explosion. The public is surprised, but prior to surprise it has seen an absolutely ordinary scene, of no special consequences. Now, let us take a suspense situation. The bomb is underneath the table and the audience knows it, probably because they have seen some anarchist put it there. … The public is longing to warn the characters on screen: “You shouldn’t be talking about such trivial matters. There’s a bomb beneath you and it’s about to explode!” In the first case we have given the public fifteen seconds of surprise at the
\end{quote}

\textsuperscript{105} Ryan, \textit{Virtual Reality}, 92.
\textsuperscript{106} Ryan, \textit{Virtual Reality}, 91.
\textsuperscript{107} Ryan, \textit{Virtual Reality}, 104.
moment of the explosion. In the second we have provided them with fifteen minutes of suspense.\textsuperscript{108}

Hitchcock’s conclusion is that surprise is too short-lived to facilitate immersion, and that cinematic storytelling should always prioritize suspense. Ryan notes however that surprise can have an immersive effect when it is involved in the formation and violation of narrative projections. An example of this might be the ending of The Planet of the Apes (1968), which reveals to the crew (and to the viewer) that the mysterious planet on which they have landed is actually earth, many years into the future. This surprise ending encourages the viewer to re-experience the film’s previous events, re-examining them in light of this new information. Ryan argues that though short-lived, a surprise development such as this may in fact facilitate immersion because the ‘interpreter will be tempted to revisit mentally the entire sequence of events, in order to locate the hidden clues to the truth.’\textsuperscript{109} Ryan considers this a form of temporal immersion, even though it is based on a momentary surprise, ‘since it turns the interpreter’s attention toward an extended period of time, rather than being limited to a single moment.’\textsuperscript{110}

Suspense is usually considered more effective than surprise in facilitating immersion, as Hitchcock’s example makes clear. The effectiveness of suspense as a storytelling technique has made it a predominant feature of storytelling across media, and the desire to know that is inherent to suspenseful situations can be seen as central to the enjoyment of much genre cinema. Consider for example, the curiosity involved in a classic whodunnit, the fear and concern for a protagonist’s fate in a typical horror film, or the desire to learn the outcome of a couple’s relationship that drives the narrative in a romantic comedy. In all these cases the viewer’s immersion is facilitated by a complex mix of curiosity, a consideration of past events, and a projected expectation about what may happen based on those past events.

Ryan also notes that the intensity of suspense is inversely proportional to the range of possible outcomes into which the narrative can develop. At the beginning of a story the possible developments are too numerous to create suspense, but as the story progresses towards the

\begin{footnotes}
\item[109] Ryan, \textit{Virtual Reality}, 106.
\item[110] Ryan, \textit{Virtual Reality}, 106.
\end{footnotes}
climax, ‘the spectrum of possible developments is reduced to the dichotomy of one branch leading to success and another in failure.’\textsuperscript{111} This is epitomized by the classic Western scene involving a heroine tied to the railroad tracks as a train approaches. Here the possible outcomes have been minimized - either the heroine will be rescued or she will not - and as Ryan points out, ‘the focus of attention in this type of suspense is the imminent resolution of a binary alternative: Will good or bad happen to the heroine?’\textsuperscript{112} Our concern with the outcome in this kind of scene is to some extent a result of vicariously experiencing the character’s situation. A typical response to a sequence like this may involve panic, fear, distress, and a racing heartbeat. Suspenseful cinema, however, remains an enjoyable experience, despite the unpleasantness of these particular responses. Green, Brock and Kaufman explain that the enjoyment of such immersion is not a result of experiencing the same emotions as the characters, but is rather due to ‘the process of temporarily leaving one’s reality behind and emerging from the experience somehow different from the person one was before entering the milieu of the narrative.’\textsuperscript{113} This goes some way to explaining the pleasurable immersion experienced during suspenseful scenes that would be unpleasant if actually experienced.

Ryan’s third key immersive response, emotional immersion, can be broadly understood as an emotional participation in the fate of fictional characters. Central to forming an emotional connection between the spectator and the fictional character is empathy, which Suzanne Keen defines in \textit{A Theory of Narrative Empathy} as a ‘vicarious, spontaneous sharing of affect’ that can come from ‘witnessing another’s emotional state’ or by ‘hearing about another’s condition.’\textsuperscript{114} Keen suggests that character identification is central to evoking an empathetic response, and that ‘[n]arratives in prose and film infamously manipulate our feelings and call upon our built-in capacity to feel with others.’\textsuperscript{115} Ryan adds that ‘an explanation of empathy based on a full emotional identification with the characters is not satisfactory … because it would in many cases make the reading of fiction into an unpleasant experience.’\textsuperscript{116} Ryan’s conclusion is that our

\textsuperscript{111} Ryan, \textit{Virtual Reality}, 107.
\textsuperscript{112} Ryan, \textit{Virtual Reality}, 107.
\textsuperscript{113} Green, Brock, and Kaufman, ‘Understanding,’ 315.
\textsuperscript{114} Keen, \textit{Empathy}, 208.
\textsuperscript{115} Keen, \textit{Empathy}, 209.
\textsuperscript{116} Ryan, \textit{Virtual Reality}, 113.
emotional involvement in a fictional narrative is limited to a pleasant immersion that falls short of a truly vicarious experience.

**Visual Effects: Technologies of Illusion**

Visual effects are increasingly important in constructing the storyworlds that are fundamental to immersive cinema. Digital visual effects technology has become so successful in creating cinematic illusions, and removing the material barriers separating spectators from them, that it has become ubiquitous in mainstream production. Decreasing costs have also widened access to the technology, with visual effects now increasingly employed on low-budget, independent, and other forms of cinema that exist outside mainstream Hollywood studio production. Due to the commercial advantage in supplying novel visual experiences, a healthy market exists in effects research and innovation. Significant technological breakthroughs such as the use of high dynamic range imaging to light scenes realistically or synthesizing the human form (see chapter 3), occur as a result of significant research and development, often from independent effects companies. These innovations are then used by the studios who rely on novel cinematic experiences to attract audiences in a competitive media-entertainment environment.

The importance of technological innovation to the economic success of Hollywood cinema can be seen by again examining *Gravity*. The film was an enormous commercial success,

---

117 There can be confusion around terminology when discussing cinematic effects. In this thesis, I use the various terms in accordance with the overview provided by Kristen Whissel and Charlie Keil in their introduction to a recent history of editing and special/visual effects. ‘Special effects’ refers to any effect that is produced on set, in front of the camera. ‘Optical effects’ are those that are achieved in the camera, through the use of masking and other matting techniques. ‘Visual effects’ refers to effects produced during post-production. I use the term ‘effects’ when referring to these processes in aggregate or in a more general way. See Charlie Keil and Kristen Whissel, eds., *Editing and Special/Visual Effects*, (London: I.B. Taurus, 2016), 12-21.

118 Despite the close relationship between visual effects innovation and commercial success, the VFX industry has suffered significant turmoil over the last decade, entailing labor issues and several high-profile bankruptcies. See Michael Curtin and John Vanderhoef, ‘A Vanishing Piece of the Pi: The Globalization of Visual Effects Labor,’ *Television & New Media* 16, no. 3 (2015): 219-239.

119 For a discussion of the immersive effect of high dynamic range imaging, see Prince, *Visual Effects*, 192-198. Bordwell, Staiger and Thompson identify ‘product differentiation’ as one of three components that determine the introduction and integration of technological developments in the cinema. The others are ‘production efficiency’ and ‘adherence to standards of quality’. See Bordwell, Staiger, and Thompson, *Classical*, 243-244. I cover this in more detail in chapter 1.
no doubt in part because of its status as a new form of cinematic experience. Contemporary reviews almost unanimously praised the film’s visual effects, describing them variously as ‘cutting-edge,’120 ‘realistic and breathtaking,’121 ‘impeccable,’122 and ‘flawless,’123 with a general agreement that they represented a significant step in the technology of cinematic representation. Many of the reviews make favorable comparisons to significant effects milestones of the past, including 2001: A Space Odyssey (1968) and Star Wars (1977), and contrast the film with other more recent effects-heavy films that are perceived to use CGI excessively, and at the expense of narrative. This latter point reveals an important aspect of the relationship between effects and narrative as understood by many in the critical community: computer-generated effects are acceptable, as long as they do not detract from the plot, characterization, and other elements fundamental to narrative. For many reviewers, Gravity was a success because it was able to integrate its visual effects achievements into a form that prioritizes narrative. As one reviewer put it, the visual effects work because they are ‘at once so novel, so technologically ahead-of-the-curve, and so essential to the underlying narrative.124 In this regard, the film succeeds in balancing the delicate interplay of spectacle and narrative often characterizing popular cinematic forms.

The contemporary ubiquity of spectacular effects in blockbuster cinema is often traced back to the 1970s with the release of films such as Star Wars and Close Encounters of the Third Kind (1977). Julie Turnock writes that these films moved beyond the studio era’s typical approach to effects, which almost exclusively limited them to a background supporting role for the narrative. She writes that the ‘studio’s aesthetic ideal was unobtrusiveness: all elements appeared in proper perspective in the frame and blended seamlessly with the live-action

124 Orr, ‘Spectacular.’
cinematography and mise-en-scène.¹²⁵ Live-action principal photography was the principal point of focus, with effects work subordinated to the background. Turnock argues however that *Star Wars* and *Close Encounters* represented a move away from the invisible effects of the studio era, to a ‘more visible and spectacular form of special effects.’¹²⁶ A typical effects sequence after these films often draws attention to itself, resulting in a response characterized by some degree of astonishment or awe at the technology used in the effects’ production. This tendency continued into the 1990s with what Michelle Pierson calls the ‘wonder years.’¹²⁷ Filmmakers of the era, including Steven Spielberg and James Cameron, began integrating computer-generated sequences into live-action photography, using updated and more sophisticated effects technology to achieve more efficiently what Spielberg and Lucas had previously done with non-digital methods. Pierson describes the mode of address characterizing a typical effects sequence of this time:

> Effects sequences featuring CGI commonly exhibit a mode of spectatorial address that - with its tableau-style framing, longer takes, and strategic intercutting between shots of the computer-generated object and reaction shots of characters - solicits a contemplative viewing of the computer-generated image.¹²⁸

Such sequences clearly encourage a direct physical immersion in a spectacle of sensory excess, but they are not, as Geoff King has pointed out, entirely divorced from the films’ narrative components.¹²⁹ These sequences usually involve a much more complex synthesis of the two forms of immersion - narrative and perceptual - outlined above. As Pierson describes it: these ‘temporal and narrative breaks might be thought of as helping to establish the conditions under which spectators’ willed immersion in the action … is suspended long enough to direct their attention to the display of the digital artefact.’¹³⁰ A spectator experiencing the typical

---

effects sequence of the early years of CGI is thus likely to vacillate between marvel at the technological achievements on display, and complete immersion in the narrative being portrayed.

Such spectacular sequences occupy the majority of both popular and scholarly discussion of effects, but a significant and perhaps underappreciated area of effects work lies in producing effects that are far removed from the spectator’s conscious awareness. By avoiding the viewer’s attention, invisible effects allow the contents of the storyworld - the characters, the plot, the themes - to remain the sole focus. This has clear implications for a consideration of immersive aesthetics: removed to the background, invisible effects allow the spectator to look past the act of mediation and become fully immersed in the diegesis. The popular and scholarly focus on spectacular effects has no doubt stemmed in part from the fact that they are simply more obvious than their invisible counterparts. Several excellent book-length treatments of effects have however highlighted the crucial role less obvious effects play in defining and upholding narrative worlds. One of the significant achievements of these studies has been to challenge the conventional view of effects as necessarily spectacular and therefore antithetical to classical cinema’s narrative focus. Shilo McClean, in Digital Storytelling: The Narrative Power of Visual Effects in Film, points out that popular film critics often blame effects for ruining films, where poor quality storytelling is more likely responsible for a particular film’s failure.131 Such views stem from an assumption that effects exist in opposition to narrative, a clear misunderstanding of the diversity of roles and functions performed by effects, and an incorrect equation of effects and spectacle. As McClean writes: ‘the assertions that all effects work always has these kinds of connotations and roles is limiting and, if accepted, precludes consideration of the other impacts and significances that might pertain to effects usages overall.’132

To broaden the general scholarly understanding of effects, McClean offers a helpful taxonomy of effects usages, which clarifies their diverse roles and functions and helps to move beyond an assumption that effects are necessarily spectacular and self-conscious in their mode of address. Most importantly for my discussion here, she distinguishes between two classes of narratively-motivated effects. Invisible effects are those that are completely undetectable by an

---

132 McClean, Storytelling, 102.
audience. The use of invisible effects is ‘neither open nor apparent’ but ‘deliberately concealed,’’ and any detection of the techniques ‘is considered by effects artists to be a failure to achieve the necessary standard.’ Invisible effects can best be illustrated by using an example of where the necessary standard was not achieved. During reshoots on the Warner Bros. superhero film *Justice League*, Henry Cavill, the actor cast to play superman, was simultaneously performing a role in Paramount’s *Mission Impossible 6*, a role which required him to have a moustache. Contractually forbidden to shave by Paramount, Cavill performed the scenes in *Justice League* with the moustache, and the effects team were given the task of removing it in post-production. The result fell far short of the invisible standard necessary for the work to pass unnoticed, and the effect became the subject of ridicule by the critical community. According to one reviewer, ‘the replacement is painfully obvious,’ and it ‘shatters the movie’s illusion whenever it happens.’ The effect clearly fell far short of the invisible standard required of such work, and the film suffered aesthetic consequences as a result.

*Seamless* effects are those that could be reasonably understood by an audience to be effects, though they are essentially undetectable through close scrutiny. Like invisible effects, seamless effects ‘seek to pass unnoticed,’ but given reasonable consideration of the production’s context, ‘their usage is detectable.’ They differ from invisible effects in that they ‘are discernible if subjected to scrutiny and consideration.’ McClean gives the example of a matte painting in *The Pianist* that constructs the World War 2-era setting of Warsaw. The effect is used to show the destruction caused to the city, but its seamless integration with the live-action footage means it is essentially undetectable upon viewing the sequence. However, given wider contextual knowledge of the film and its production, a viewer can reasonably assume that the setting has been re-created through set extensions. In this example, the effect is essential to the film’s overall narrative purpose, providing ‘both an accurate visual representation of the diegetic

---

133 McClean, *Storytelling*, 77.
135 McClean, *Storytelling*, 78.
world and the emotional impact of the extent of destruction.¹³⁷ A seamless effect such as this ‘does not seek to draw attention to itself for spectacular purposes,’ but rather ‘seeks to ensure the narrative coherence within classical Hollywood storycraft.’¹³⁸ Invisible and seamless effects make up an important if understudied part of the visual effects work done in contemporary cinema. The important role these less obvious effects play in upholding the illusion of the narrative world means they are crucial to any consideration of immersive cinematic aesthetics.

Cinema’s Virtual Spaces

In his wide-ranging archaeology of visual culture, Lev Manovich places cinema in a long history of media forms that create the illusion of a space that exists independently of the real world. He writes that ‘the visual culture of the modern period, from painting to cinema, is characterized by an intriguing phenomenon - the existence of another virtual space, another three-dimensional world enclosed by a frame and situated in our normal space.’¹³⁹ In pre-digital cinema, continuity editing was used to construct these illusionary spaces. As Manovich writes:

> Traditional fiction film transports us into a space - a room, a house, a city. Usually, none of these exists in reality. What exists are a few fragments carefully constructed in a studio. Out of these disjointed fragments, a film synthesizes the illusion of a coherent space.¹⁴⁰

Through cinema’s various formal elements, including set design, staging, lighting, and camera placement and movement, the viewer is positioned in the optimum viewpoint of each shot and is “present” inside a space that does not really exist.¹⁴¹ Continuity editing, or what Manovich calls ‘temporal montage,’ is the principal means of creating these virtual spaces. By

---

¹³⁷ McClean, *Storytelling*, 79.
¹³⁸ McClean, *Storytelling*, 82.
placing separate realities into consecutive moments of time, cinema constructs the illusion of spatial and temporal continuity, creating ‘a sense of presence in a virtual space.’

Continuity editing has important consequences for an immersive aesthetics, but more important to this discussion of the role of visual effects in forming these virtual spaces of illusion is Manovich’s second form of montage, ‘montage within a shot,’ which involves the combination of multiple elements into a single coherent image. Manovich writes that ‘montage within a shot’ involves separate realities forming ‘contingent parts of a single image.’ It is here that digital technology has its most radical potential in the synthesis of virtual spaces of illusion. Digital technology represents a significant increase in the ease and efficiency with which these spaces of illusion can be synthesized. Digital imaging tools can augment existing photographed footage, combine different pieces of footage, and create entirely synthetic forms that have no counterpart in the real world, integrating them with existing footage to create a seamless illusion of a virtual space. This then is one of the major achievements (or dangers) of digital visual effects technology: the synthesis of believable and photorealistic image spaces that have no extra-filmic referent.

Writing in 1994, early in the transition to digital technology that would soon entirely transform the media environment, William Mitchell highlights the immersive power of digital images:

They can take us beyond the boundaries of the real world and insert our disembodied viewing presences into modelled fictional worlds - three-dimensional worlds that once were, might have been, will be, or are projected forward by designers, imagined by film directors, sought by visionaries, proffered by dissimulators. In hyper-Cartesian fashion, our perceiving faculties are pried apart from our corporeal existence and sent to places where our bodies cannot follow.

142 Manovich, Language, 149.
143 Manovich, Language, 148.
144 Mitchell, Reconfigured, 134.
In the decades following Mitchell’s claim, the complexity and power of digital compositing technology has only increased, allowing the construction of synthesized image spaces that are increasingly believable and immersive. As Manovich points out, digital compositing represents a new step in the history of visual simulation because ‘it allows the creation of moving images of nonexistent worlds.’\(^{145}\) (italics original). Modern digital compositing is a significant technological development because it radically increases the ease and efficiency with which these moving virtual spaces of illusion can be synthesized.

To illustrate the role of compositing in creating and sustaining cinematic illusions, Manovich contrasts two different methods by which compositing creates a virtual world. As an example of the first method, he describes a scene from \textit{Cliffhanger}, in which Sylvester Stallone is shown ‘high in the mountains hanging over an abyss.’\(^ {146}\) This shot is a composite of Stallone filmed against a blue-screen with a mountain landscape background plate. In this example, Manovich explains, ‘it is immediately clear that the composited shot represents something that never took place in reality,’ and the result is therefore clearly identifiable as a ‘virtual space.’\(^ {147}\) As an example of the second method, Manovich describes the shots from \textit{Jurassic Park} that feature computer-generated dinosaurs added to a live-action shot of an exterior location. Manovich suggests that in this second example, ‘it may appear that the existing physical space is preserved’, but, he argues, ‘here, as well, the final result is a virtual world that does not really exist.’\(^ {148}\) In both cases, digital compositing is used to furnish the storyworlds of Hollywood cinema, synthesizing virtual spaces that have no real-world referent.\(^{149}\)

This use of digital compositing to create and expand Hollywood’s storyworlds is a theme I take up in chapter 1. I focus on a particular spectator response – spatial immersion – and examine the historical development of compositing as a means of eliciting spatial immersion as a

---

\(^{145}\) Manovich, \textit{Language}, 153.

\(^{146}\) Manovich, \textit{Language}, 137.

\(^{147}\) Manovich, \textit{Language}, 138.


\(^{149}\) Mike Jones sees the virtual camera, ‘a non-physical device possessing all the traits and characteristics of a physical camera’, as fundamental to digital cinematic forms that no longer stage the mise-en-scene for the camera, but rather ‘stage and compose the camera itself as a form of specific purpose scenic content.’ One implication of this change, along with the increasing use of constructed and composited 3D spaces, is that ‘live-action photography is increasingly seen as an option rather than a given’, (italics original). See Mike Jones, ‘Spatial Composition and the Virtual Camera,’ \textit{Animation} 2, no. 3 (2007): 226.
response to diegetic worlds, a primary component in Hollywood’s aesthetic paradigm. I argue that the industry-wide shift to digital compositing techniques and away from traditional methods of combing multiple image elements – multiple exposure, rear projection, and optical compositing – has occurred primarily because digital methods are more conducive to facilitating the spectator’s spatial immersion in the diegetic world. I examine Hitchcock’s *To Catch a Thief* (1955) and Ang Lee’s *Life of Pi* (2012), two films separated by more than 50 years of technological and stylistic development, but both of which seek the seamless combination of multiple image elements with the ultimate aim of upholding the illusion of an expansive diegetic world in which the spectator can become immersed. The rear projection used in *To Catch a Thief* draws attention to the process of its creation, while the digital compositing in *Life of Pi* is invisible, effectively hiding the artifacts that often result as a consequence of the process of joining multiple image elements. Digital compositing thus achieves the medium transparency that is essential for the spectator’s unhindered access to the contents of the diegetic world. Digital compositing’s radical expansion of the control filmmakers and effects artists have over the final image, and its ability to more effectively hide the means by which cinematic illusions are created, have made it the standard method of creating the virtual spaces that furnish Hollywood’s storyworlds.

A seamless virtual space can be a powerful stimulus to a sense of presence – the displacement of the real world by a virtual world in the spectator’s conscious experience. But for a cinematic world to function as part of an ongoing narrative, it must also include a temporal dimension. Adding duration to the spatial immersion in virtual world can produce a powerful sense of relocation to the time and place of the narrative event, an immersive response that Ryan describes as ‘spatio-temporal immersion.’ In chapter 2 I examine the long take, cinema’s celebrated shot of extended duration, as a particularly powerful means of eliciting such a response. I focus in particular on two functions of camera movement. Firstly, I examine the way camera movement functions to seamlessly join disparate parts of a diegetic space, eliminating the need to break up the space into distinct parts through editing, and allowing perceptual continuity in the spectator’s relationship to the space. And secondly, I examine the way a moving image can generate a visceral sense of movement through space that more closely resembles

phenomenological engagement with the real world. I’ll examine the 13-minute long take that opens Alfonso Cuaron’s *Gravity*, positioning the shot within a Bazinian framework and showing that despite its use of digital tools to facilitate the extended duration, the digital long-take can still be conceptualized within a Bazinian phenomenology. I’ll argue that contrary to those that describe a post-classical cinema of fragmentation and discontinuity, digital technology represents in many cases an intensification of classical continuity. A classically-inflected conception of space and time remains common in contemporary effects-heavy cinema, and digital technology and effects are often used to enable and extend continuity, rather than displace it.

In chapter 3 I examine the spectator’s relationship to the characters that populate diegetic worlds, focusing on the powerful emotional responses that emerge from identifying with the subjectivity of a character. To clarify the role of character relationships in facilitating diegetic immersion, I will develop a distinction between two types of affective response in cinema that are often conflated in critical discourse under the broad notion of ‘immersion.’ *Aesthetic immersion* is a response to cinema that can be experienced independently of narrative concerns such as plot and character. These powerful affective responses emerge not from a relocation to the time and place of a narrative event, but from moments in the cinema that exceed narrative. Some examples of aesthetic immersion involve a self-conscious gesture to the medium itself, while others, such as symbolism and metaphor, involve an extra-diegetic level of communication. But most forms of aesthetic immersion tolerate an intrusive authorial voice and some degree of medium awareness. *Emotional immersion*, on the other hand, is an affective response that emerges from a strong identification with a character and involves the powerful sense of relocation to the diegetic world. Emotional immersion relies on medium transparency, placing character and plot at the forefront of the spectator’s conscious awareness. It thus forms an important part - along with spatial immersion and spatio-temporal immersion - in the phenomenology of diegetic immersion that is my primary concern. I will discuss the prospect of emotional immersion as a response to digitally-augmented characters, challenging the view that the use of digital technology to simulate human-like virtual characters diminishes an essential part of the aesthetics of cinema. I’ll use *The Curious Case of Benjamin Button* (2008) as an

---

example to argue that the use of digital visual effects to augment human performances represents an extension of the aesthetics of the medium.

Chapter 1 – Spatial Immersion and Hollywood’s Storyworlds

This chapter examines changes in visual effects technology in relation to the conventions and practices of a more stable, or classical, cinematic style. Significant debate surrounds the extent to which the classical norms established early in cinema’s history, as set out in the most comprehensive analysis of the form, the Classical Hollywood Cinema, continue in the effects-heavy digital cinema of today. David Bordwell, for example, argues that there is a large degree of stylistic continuity between the classical cinema of the studio era and contemporary cinema. He writes that ‘nearly all scenes in nearly all contemporary mass-market movies (and in most “independent” films) are staged, shot, and cut according to principles which crystallized in the 1910s and 1920s.’ These principles include psychological causality, continuity in depicting space and time, and a level of narration that remains invisible, functioning transparently to mediate the diegetic world. Bordwell does admit that there have been stylistic changes, but argues that these amount to an ‘intensification of established techniques’, rather than a complete overturning of them. (original italics). He concludes that in ‘representing space, time, and narrative relations … today’s films generally adhere to the principles of classical filmmaking.’

Contrary to Bordwell, a number of scholars have suggested that since the 1970s there has been an important epochal shift toward a ‘post-classical’ cinema that is closely aligned with a more general cultural embrace of postmodernism. This post-classical cinema is characterized

152 Bordwell, Staiger, and Thompson, Classical.
154 Bordwell, Intensified, 16.
155 Bordwell, Intensified, 16.
156 See for example, Steve Neale and Murray Smith, eds., Contemporary Hollywood Cinema, (London: Routledge, 1998). Others, Elizabeth Cowie for example, argue that the formal and stylistic norms outlined in the Classical Hollywood Cinema detail a cinematic form so broad that it encompasses everything, without allowing for unique exceptions. ‘A dominant mode of narration,’ she writes, ‘is defined, but its very definition includes, it seems, virtually all possible deviations, so that every exception therefore proves the rule.’ Elizabeth Cowie, ‘Storytelling:
by parody, nostalgia, a mixing of popular culture and high art, a spectator relationship defined by distance and irony, and fractured narrative and stylistic disunity. I am generally sympathetic to Bordwell’s view, and agree that the large degree of stylistic continuity in Hollywood cinema makes the argument of a post-classical cinema that differs fundamentally from classical forms difficult to sustain.\textsuperscript{157} There are enough stylistic and formal continuities in the cinema of the studio era and in the cinema of today to suggest that rather than a fundamental break, there has been a process of gradual change in certain aspects of cinematic style, with the fundamental tenets of classical storytelling remaining intact. At the very least, I argue, there is sufficient continuity to justify the more general claim that there are certain aspects of classical form that transcend immediate historical moments.

The principle of medium transparency, for example, has remained fundamental to popular cinematic storytelling from the earliest days of cinema, when the art first began to take a narrative form. Bordwell describes this principle in the opening chapter of \textit{The Classical Hollywood Cinema}. ‘[T]he Hollywood film,’ he writes, ‘strives to conceal its artifice through techniques of continuity and ‘invisible’ storytelling.’\textsuperscript{158} One of the reasons for the continuity of this principle, I argue in this chapter, is that an immersive relationship with a diegetic world is fundamental to popular engagement with cinema, and removing the means by which such worlds are constructed is central to forming such a relationship. However, this continuity overlooks the significant technological changes that have occurred in the visual effects used to construct diegetic worlds. Digital compositing, the process of combining multiple image elements into a single, coherent, and seamless whole, is the basic function underlying most of the visual effects seen in contemporary popular cinema. Like invisible storytelling more generally, the combination of multiple image elements extends back to cinema’s beginning in the late 19th

\textsuperscript{157} Eleftheria Thanouli also agrees with Bordwell’s account of contemporary cinema, but suggests that it fails to fully account for important exceptions. Films such as \textit{Moulin Rouge} (2001), \textit{Magnolia} (1999), \textit{Fight Club} (1999), and \textit{Requiem for a Dream} (2000), Thanouli argues, ‘do not fit comfortably in the classical mould and go a step further in their narrative experimentation.’ Rather than fully embracing Bordwell’s rejection of a post-classical cinema, Thanouli advocates considering the ‘emergence of a new post-classical mode that can be defined in strictly poetic terms based on a close textual analysis devoid of the postmodern tropes that have been exhaustively recycled.’ Eleftheria Thanouli, ‘Post-Classical Narration: A new paradigm in contemporary cinema,’ \textit{New Review of Film and Television Studies} 4, no. 3 (2006): 184. DOI: 10.1080/17400300600981900.

\textsuperscript{158} Bordwell, Staiger, and Thompson, \textit{Classical}, 3.
Century, but the addition of digital techniques has seen radical changes to the efficiency with which these images are produced, and though performing essentially the same function as their pre-digital counterparts, digital methods result in significant improvements to the final images that appear on screen.\textsuperscript{159} This has important implications for the spectator’s spatial immersion in the diegetic worlds represented. According to Stephen Prince, visual effects produced through digital compositing ‘are able to compel belief in the fictional world of the film in ways that traditional special effects could not accomplish.’\textsuperscript{160} This chapter will show that digital compositing represents a significant gain in the ability for Hollywood cinema to ‘conceal its artifice’, entailing a qualitative change in the ability of filmmakers to elicit a spatial immersion in diegetic worlds.

To examine the historical development of compositing and the implications this has for the spectator’s spatial immersion, I will focus on rear-projection and chroma-keying, two specific techniques used to combine images to produce the virtual spaces fundamental to popular cinema. My examples are taken from Hitchcock’s \textit{To Catch a Thief} (1955) and Ang Lee’s \textit{Life of Pi} (2012), two films separated by more than 50 years of technological and stylistic development, but both popular mainstream entertainments that use visual effects to elicit the same response in the spectator. In the two scenes I examine in detail, the films seek to immerse the spectator in the diegetic worlds portrayed on screen, encouraging the ‘quasi-experience’ of being on a boat with the characters and surrounded by a body of water.\textsuperscript{161} However, it is only the digital compositing in \textit{Life of Pi} that achieves a level of transparency in its mediation of the diegetic world. The technological limitations of rear projection mean that the compositing images in \textit{To Catch a Thief} are clearly identifiable as such.

The difference between these two examples raises questions about the tension between classical cinema’s seamless aesthetic, and the limitations of the medium at certain points in its

\textsuperscript{159} The evaluative and aesthetic implications of using the word ‘improvements’ here are problematic. ‘Improvements’ in the way I’m using it should be understood in terms of changes to an image that reduce its generally undesirable elements such as grain, limited resolution, or anything else that interferes with the transparent mediation of the contents of a representation.

\textsuperscript{160} Prince, \textit{Digital}, 33.

\textsuperscript{161} Wolf suggests that when immersed in a represented world, spectators experience psychological responses that are similar to those that would be felt in real life. He thus argues that immersive media elicit ‘quasi-experiences.’ Wolf, \textit{Immersion}, 11-12.
technological history. A significant determinant of technological development in the context of visual effects has been the drive to eliminate perceived deficiencies in the image, which in many cases consist of artifacts resulting directly from the image manipulation inherent to compositing. Artifacts generally violate the ideal of transparency central to the classical style, and therefore significant engineering and technical effort has been directed towards their removal. One of the significant achievements of digital compositing is increased control over the removal of such artifacts, and therefore greater transparency in the mediation of diegetic worlds. Digital compositing thus closely aligns with the industry’s ideal seamless aesthetic.

This chapter will also show that while digital compositing represents, as Lev Manovich puts it, ‘a qualitatively new step in the history of visual simulation,’ there are important historical and contextual factors that must be taken into account when considering the audience response to such images. Rear projection was a common technique during the classical era, but it is problematic for contemporary audiences due to its obtrusive violation of classical cinema’s transparent aesthetic. Its use by Hitchcock in To Catch a Thief and numerous other films shows that there is a degree of historical contingency in the ability for certain visual effects technologies to remove themselves from the viewer's conscious awareness. Production and practical techniques in specific historical contexts can be shown to compel belief in a fictional diegesis that later technologies and perceptual itineraries render obtrusive. These changing standards are to some degree caused by developments in visual effects technologies that ‘denaturalize’ previously accepted standards of medium transparency. As Werner Wolf writes, ‘recipients simply get used to certain technical standards, and representations falling short of them may trigger more distance than thorough immersion would tolerate.’

162 An artifact is the industry term used to describe any aspects of a composite that reveal that the image has been manipulated. Brinkman defines them as the ‘undesirable items in an image that are the result of the process used to create the image.’ Ron Brinkmann, The Art and Science of Digital Compositing, 2nd edn., (Burlington, MA: Morgan Kaufmann, 2008): 13.

163 Manovich, Language, 153.

164 Hitchcock used the technique consistently throughout his career. Some notable examples include the location shots of Rio de Janeiro in Notorious (1946), the crop-duster scene in North by Northwest (1959), and the scenes of Scottie following Madeleine-Judy in Vertigo (1958).

165 Wolf, Immersion, 26.
audience standards regarding transparent mediation means that rear projection has been almost completely replaced by other, more transparent methods of combining multiple image elements.

Medium Transparency

Numerous theorists have identified medium transparency as central to not only popular narrative cinema, but also to a much wider range of media relationships. Jay David Bolter and Richard Grusin, for example, describe transparency as a key aspect of new media generally, arguing that it constitutes a fundamental part of the ‘double logic of remediation’. Remediation refers to the process by which new media refashion earlier media, a process characterized by the conflicting demands of immediacy and hypermediacy. Immediacy involves placing the spectator in an immediate relationship with the contents of a medium, erasing any traces of the mediation. Hypermediacy describes the tendency of new media objects to draw attention to themselves, emphasizing the process of construction rather than the finished product. The ‘double logic of remediation’ describes the interrelationship between immediacy and hypermediacy, with new media seeking to erase any traces of themselves, while at the same time remaining conspicuous as media objects.

Bolter and Grusin argue that medium transparency is a key factor in immediacy. Transparency involves the removal of the medium itself, ‘so that the user is no longer aware of confronting a medium, but instead stands in an immediate relationship to the contents of that medium.’ By disappearing, transparent media ‘put the viewer in the same space as the objects viewed.’ Bolter and Grusin describe transparency in relation to various examples of new media including virtual reality, three-dimensional graphics, and user interface design, and also ‘two- and three-dimensional images projected on to traditional computer, film, or television screens.’ They argue, however, that remediation is not specific to new digital media, but is rather an historical tendency characterizing the development of Western visual representation. For example, Western painting that used oil paint in an attempt to ‘erase’ the paint from the

---

surface of the painting ‘concealed and denied the process of painting in favor of the painted product.’ The ideal of transparency was also central to the development of linear perspective. ‘By using projective geometry to represent the space beyond the canvas,’ they write, ‘linear perspective could be regarded as the technique that effaced itself as technique.’ They further note that although effacement was by no means universal in Western painting, ‘it was one important technique for making the space of the picture continuous with the viewer's space.’

Bolter and Grusin describe contemporary conceptions of virtual reality as the clearest examples of the logic of transparent immediacy. In its ideal incarnation, virtual reality produces in the user the experience of completely crossing the threshold separating the space of reality from the virtual space, where they are free to explore and interact with the simulated environment. For Bolter and Grusin, virtual reality has become ‘a cultural metaphor for the ideal of perfect mediation.’ It is seen as the ultimate representational medium, representing the fulfillment of the ideals of transparent mediation promised by earlier media such as the telephone, cinema, and television. They argue however, that this attempt to elevate virtual reality above all other forms of mediation ‘seems instead to anchor virtual reality more firmly in the history of representation.’

As Bolter and Grusin argue, and as I will demonstrate in my discussion of pre-digital cinema below, the ideals of medium transparency often seen as specific to new media apply equally to media that long predate the digital era. The principle of transparency is in fact central to all forms of art that seek to elicit an immersive response to a diegetic world. By functioning to invisibly communicate a narrative world, transparent media facilitate an immersive engagement with the contents of that world. This medium transparency is central to classical narrative aesthetics, with the 19th Century realist novel a prototypical example. According to Werner Wolf, realist novels have traditionally been accorded a particularly high potential for eliciting immersion as their ‘level of discourse is comparatively inconspicuous, serving mainly to transmit

the storyworlds and to support their consistency and lifelikeness.175 Removing the medium from conscious awareness ‘permits the recipient’s attention to focus, in a relatively ‘easy reception’, on the represented content rather than on the transmission or medium as such.’176 Classical film aesthetics involves the same mode of address, with technique and style relegated to the background of the spectator’s attention, encouraging a focus on the storyworld and content. In *The Classical Hollywood Cinema*, David Bordwell points out that classical cinema’s ‘self-effacing craftsmanship’ foregrounds the storyworld and moves technique and style to the background of the viewer’s conscious awareness.177 Cinema’s formal elements are rendered invisible in their construction of ‘an apparently independent profilmic event.’178 By concealing the production in this way, classical form allows the easy reception of the film’s narrative, facilitating an immersive engagement with a diegetic world.

**Visual Effects Compositing**

The importance of the ideal of medium transparency to popular engagement with cinema means that a significant strand of Hollywood’s technological history can be read as a consistent attempt to remove any evidence of the process of mediation from the final image.179 This is particularly relevant in visual effects compositing. As a procedure involving the combination of carefully selected elements from multiple image sources into a single representation, compositing is particularly susceptible to obtrusive technique that could reveal the image’s constructed status. Visual effects have therefore always existed in tension with the transparent demands of the classical style, and the history of their development, at least from a technical standpoint, can be seen as a consistent struggle to remove the traces of the image manipulation

177 Bordwell, Staiger and Thompson, *Classical*, 4.
179 Bordwell points out that films generally conforming to the classical ‘invisible’ style still have clear reflexive reminders of cinema’s constructed nature, such as opening credits sequences. See Bordwell, Staiger and Thompson, *Classical*, pp. 24–41. And there are of course modes of narration in the Hollywood tradition that deviate more consistently from the invisible style, as shown in comedies and musicals that flamboyantly emphasize artifice and anti-naturalism for comical or romantic effect. The stylized Manhattan skyline that forms the backdrop for the ‘Dancing in the Dark’ sequence from Vincente Minnelli’s *The Bandwagon* (1953), for example, enhances the sense of romantic fantasy.
inherent to the process. The development of visual effects compositing is therefore a valuable site for examining the tension between innovative cinematic technology and the ideal of transparency that is central to the classical style. By analyzing the historical development of the different technologies and methods of compositing we can gain insights into both standards of medium transparency at certain historical and cultural moments, as well as aspects of cinematic style that transcend these immediate contexts.

Compositing is central to the dominant mode of cinematic production today, underlying an enormous range of cinematic effects and processes. In cinema’s early years, compositing usually involved combining different photographic elements using multiple exposures and various matting techniques, but the rapid growth of digital imaging technologies has led to a corresponding increase in the techniques used and the range of different sources from which image elements can be taken. Regardless of the original source, the goal of the compositor is to achieve transparency by blending the elements as seamlessly as possible, producing a composite that hides any trace of the image manipulation that has taken place. Steve Wright, a visual effects artist with credits on over 70 feature films, and author of several practical guides to digital compositing, outlines some of the potential image sources:

- We might be adding an actor or a model from a piece of film or video tape.
- Or perhaps the mission is to add a spaceship or dinosaur that was created entirely in a computer, so it is referred to as a computer generated image (CGI). Maybe the element to be added is a matte painting done in Adobe Photoshop.\textsuperscript{180}

Wright points out that it is ‘the digital compositor that takes these disparate elements, no matter how they were created, and blends them together artistically into a seamless, photorealistic whole.’\textsuperscript{181} The digital compositor’s goal, he writes, is ‘to make them appear as if they were all shot together at the same time under the same lights with the same camera.’\textsuperscript{182} Ron Brinkman, a visual effects supervisor with credits on a number of Hollywood box-office

\textsuperscript{180} Brinkmann, \textit{Art}, 1.
\textsuperscript{181} Brinkmann, \textit{Art}, 1.
\textsuperscript{182} Brinkmann, \textit{Art}, 1.
successes from the 1990s, articulates the same ideals in his book *The Art and Science of Digital Compositing*. He argues that ‘[b]y far the most difficult part of this digital compositing process is producing the *integrated* result - an image that doesn’t betray that its creation was owed to multiple source elements.’\(^{183}\) (italics original). This focus on transparency among practicing compositors is reflected in the more theoretical accounts of the technique. Lev Manovich, for example, describes the aims and functions of the process:

> Once all the elements are ready, they are composited together into a single object; that is, they are fitted together and adjusted in such a way that their separate identities become invisible. The fact that they come from diverse sources and were created by different people at different times is hidden. The result is a single, seamless image, sound, space, or scene.\(^{184}\)

Audiences generally expect visual effects compositing to achieve the level of transparency of other formal elements of cinema, and any break in this principle is likely to be noticed as readily as a break in continuity editing or violation of the 180-degree principle.\(^{185}\) The importance of hiding any evidence of image manipulation can be seen by examining the reaction to a film with particularly obtrusive effects work. Figure 2 shows a composite from David Lynch’s *Dune*, which was met with a predominantly negative critical reception upon its release in 1984, and which Julie Turnock consequently describes as ‘a remarkable historical instance of the impact special effects had on the reception and perceived success of a film.’\(^{186}\) One of the

\(^{183}\) Brinkman, *Art*, 2.

\(^{184}\) Manovich, *Language*, 136. The transparent compositing I’m referring to is specific to classical modes of cinema. Manovich also offers an alternative understanding of compositing, one that is fundamentally different to the way most digital compositors understand the term, at least those working within mainstream cinema production. He writes: ‘Although digital compositing is usually used to create a seamless virtual space, this does not have to be its only goal. Borders between different worlds do not have to be erased; different spaces do not have to be matched in perspective, scale, and lighting; individual layers can retain their separate identities rather than being merged into a single space; different worlds can clash semantically rather than form a single universe.’ Manovich, *Language*, 158.

\(^{185}\) The 180-degree principle is based around the imaginary line in the diegetic space that determines where the camera can be placed when editing together shots taken from multiple camera setups. To ensure spatial continuity in the final edit, the camera must be consistently positioned on one side of the line. See Bordwell and Thompson, *Film Art*, 233-243.

significant problems with the film’s compositing was its failure to seamlessly blend the separate elements within the composition, as shown in Figure 2. The foreground element to the left of the frame has noticeable matte lines around its edge, separating it from the background and clearly revealing the image to be a composite. The result fails entirely in effacing evidence of the manipulation that has taken place, interfering with the illusion of a coherent and fully-formed diegesis that exists outside of its representation as a cinematic image. Roger Ebert described the film as ‘a real mess, an incomprehensible, ugly, unstructured, pointless excursion into the murkier realms of one of the most confusing screenplays of all time.’ But Ebert is particularly critical of the film’s effects work:

The heads of the sand worms begin to look more and more as if they came out of the same factory that produced Kermit the Frog (they have the same mouths). An evil baron floats through the air on trajectories all too obviously controlled by wires. The spaceships in the movie are so shabby, so lacking in detail or dimension, that they look almost like those student films where plastic models are shot against a tablecloth.

Underlying Ebert’s specific criticisms is a more general aversion to such blatant evidence of image manipulation. An expected standard of effects work is implicit in his criticism, particularly in a film with such a considerable production budget. This standard had been cultivated by earlier films such as 2001: A Space Odyssey, Star Wars and Close Encounters of the Third Kind (1977), films that had raised audience expectations regarding medium transparency in visual effects design and composition. The hostile reception to Dune and its disastrous commercial performance is no doubt the result of its failure to achieve such a standard in the visual effects composition of diegetic worlds. As such, Turnock suggests, the film

---


188 Ebert, ‘Dune.’

stands as ‘a prime example of a historical dividing line for marking expectations of photorealistic
special effects.’ Films such as *Dune* can enable us to track the changing standards of
transparency, and reveal the tension between technological developments in compositing and the
transparent ideals of popular cinema.

**Figure 2**

**Synthetic Images and Realism**

Visual effects are increasingly important in constructing cinematic storyworlds, and as
with the more traditional elements of cinematic form such as camera movement, lighting, and
editing, effects in classical Hollywood cinema are generally subordinate to narrative, ideally
functioning to transparently convey a film’s narrative material. As Charlie Keil and Kristen
Whissel write in the introduction to a volume examining historical developments in visual
effects:

> [A]rtists and technicians have historically worked to create effects with a
degree of proficiency, precision, and seamlessness that effectively
conceals both the manufactured nature of an effect (however stylized) and

---

Despite these discrepancies, both sets of figures show that the film was a significant commercial failure, no doubt in
part because of the antiquated effects work. As a comparison, *The Numbers* reports that the top box office success of

the technologies used to create it, thereby allowing audiences to engage in a (more or less) uninterrupted suspension of disbelief as they watch a film.\textsuperscript{191}

The removal of any seams or artifacts that result from the compositing process is central to the transparent mediation of the diegetic world. Furthermore, Keil and Whissel point out that this medium transparency applies regardless of the exact quality of spectator relations encouraged by an effect. ‘Such transparency’ they write, ‘is the goal even in instances where a particularly elaborate effects sequence announces its status as the outcome of groundbreaking techniques and artistry that showcase the power of new technologies.’\textsuperscript{192} Hiding the means by which an effect is created is a primary goal of visual effects artists, regardless of its adherence or otherwise to an illusory realism.

Realism is a concept with significant implications for immersive spectatorship, but its varied understandings among scholars, filmmakers, visual effects artists, and others concerned with cinematic style, means it needs to be carefully theorized and its use contextualized. In cinema and media studies, the concept has a long and complex history, and has been conceptualized in a variety of different ways by numerous theorists, generally focusing on the different aspects of the visual representation of reality, including ontological implications, models of spectatorship, and ethical concerns.\textsuperscript{193} However, these understandings generally differ from the way realism is understood in the visual effects community. Visual effects artists generally understand realism in terms of the perception of an image, with little concern given to its referential status. Many of these artists work entirely with synthetic images that have no extra-filmonic referent, so the idea of preserving the integrity of a representation by maintaining correspondence with the extra-filmic reality it ostensibly represents is irrelevant. A realistic image for these artists is one that structurally corresponds to a hypothetical real-world scene, closely replicating real-world markers of lighting, shading, depth, perspective, and other major visual signifiers. A photorealistic image, furthermore, is one that accurately replicates the way a

\textsuperscript{191} Keil and Whissel, \textit{Editing}, 14.
\textsuperscript{192} Keil and Whissel, \textit{Editing}, 14.
camera captures a scene, including a range of camera artifacts such as barrel distortion, lens flares, and chromatic aberration.\textsuperscript{194} Understanding the realism of an image in this way – in terms of perception rather than representation – reflects an entirely different way of thinking about the ontology of visual imagery, no doubt a consequence of the success and ease with which digital techniques have been used to produce entirely synthetic images.

An example of this understanding of realism in terms of perception can be seen in the work of Andrew Price, a digital artist and online educator with a website and a popular YouTube channel providing tutorials and instructional videos about \textit{Blender}, an open-source 3D computer graphics software program.\textsuperscript{195} In two videos, \textit{Photorealism Explained: the Principles of Photorealistic Rendering} and \textit{The Secret Ingredient to Photorealism}, Price uses the term ‘photorealistic’ to refer to imagery that is created entirely synthetically using computer graphics software, but which still achieves the appearance of authentic photography.\textsuperscript{196} He divides the creation of a virtual scene into four steps – modelling, materials, lighting, and post-processing – each of which can be completed inside of \textit{Blender}, or any other 3D computer graphics software. Price’s overview of these four steps shows that artists working with synthetic imagery draw extensively from real world visual cues when aiming for a photorealistic result.\textsuperscript{197}

\textit{Modelling}

Modelling is the process of constructing three-dimensional geometric representations of objects by arranging collections of points in a virtual 3D space. These points are connected by various shapes such as triangles, lines, and curves, and arranged to form a three-dimensional representation of a particular object, which could be a person, building, or any other animate or

\textsuperscript{194} Ron Brinkman defines each of these camera artifacts. Barrel distortion is ‘[d]istortion of a lens that causes straight lines to bend away from the center of the image.’ A lens flare is ‘[a]n artifact of a bright light shining directly into the lens assembly of a camera.’ Chromatic aberration is an artifact ‘caused by the fact that different wavelengths of light are bent by slightly different amounts as they pass through a lens.’ Brinkmann, \textit{Art}, 634-653.


\textsuperscript{196} Manovich offers a helpful definition of ‘photorealistic’: ‘the industry term for synthetic images that look as though they were created using traditional photography or cinematography.’ Manovich, \textit{Language}, 179.

\textsuperscript{197} Unless indicated otherwise with quotation marks, the explanations that follow are my paraphrases of Price’s ideas.
inanimate figure. To achieve a photorealistic result, according to Price, the resulting model must match ‘the proportions and form of the real world object.’ To achieve this match, Price and other artists recommend using real photographs of the object being modelled as reference images, which can be used to guide the modelling process and help develop the correct form and shape of the model. Real-world reference is an essential part of the process of photorealistically representing synthetic objects, and is frequently used in developing the computer-generated imagery seen in much popular cinema, as will be shown in the discussion of Life of Pi below.

**Materials**

An artist applies materials to an object once it has been modelled, specifying the object’s optical properties including color and specularity, as well as the interaction of light with the object’s surface. Textures are applied along with the materials to capture the unique pattern of the surface being modelled. For example, in creating a wooden table, an artist will first model the table’s basic shape and form, and then apply a wood texture to capture the unique pattern of color variation and the interaction of light and shadow on the table’s surface. As with the modelling, achieving a photorealistic result with the materials means matching the appearance of these elements in the real world. Due to the sensitivity of the human eye to the interaction of light on the surface of objects, Price places particular importance on this step if the goal is a photorealistic result. ‘There’s something about … the way that the light hits an object,’ he suggests, ‘the way that it’s received into the object, that our eyes just pick up on.’ The realism of the resulting image, according to Price, depends on achieving the same appearance with the synthetic materials as those seen by the human eye during real-world perception.

**Lighting**

3D computer graphics programs allow an artist to add a range of different light sources to a scene. These lighting functions simulate the tools available to a real-world cinematographer and give enormous flexibility and control over how the lighting will appear in the final render. Photorealistic lighting, according to Price, is ‘light that matches the color, direction, and intensity seen in real life,’ and as with the previous steps, he suggests using real-world lighting cues to design the lighting of a virtual scene. As an example, he advises using the Kelvin color temperature scale to help match the color of the simulated lights to the color of real world light
sources. For example, if designing a virtual scene with natural sunlight as the light source, a color temperature of around 5500K (the color of sunlight at midday) would help to achieve a result that approximates real-world lighting.\textsuperscript{198}

\textit{Post-processing}

Photorealistic post-processing involves ‘recreating imperfections from real life cameras.’ A real world scene captured with a traditional camera includes a range of imperfections resulting from the apparatus itself, often a result of light moving through its various components. The virtual cameras that are used to capture scenes created in 3D graphics software are free from such imperfections, so to achieve a photorealistic result, a digital artist must artificially add these artifacts during post-processing. Price includes five categories of imperfections – glare (including lens flares and other distortions caused by light shining into the lens assembly), motion blur, depth of field, chromatic aberration, and barrel distortion – each of which can be added artificially during the post-processing of a virtual scene. As with the other steps, the goal is to hide the fact that the image was created artificially, giving instead the impression that it resulted from traditional photographic processes.

In each of these four steps, the digital artist attempts to convince the spectator that the scene depicted in the image actually existed in front of a real camera and was photographed at some point in time. This reflects conceptions of photorealism found more broadly among the visual effects community. ‘The goal of photorealistic rendering,’ according to the authors of \textit{Physically Based Rendering: From Theory to Implementation}, a standard textbook on the topic, ‘is to create an image of a 3D scene that is indistinguishable from a photograph of the same scene.’\textsuperscript{199} Price’s approach to this challenge, as shown by the steps outlined above, is to replicate the real world as closely as possible in the modelling, materials, and lighting, and then in the fourth step, the post-processing, adding a range of camera artifacts to produce an image in a style that has been naturalized as a standard of photographic realism.\textsuperscript{200} This style aims to convince


\textsuperscript{200} For an examination of photorealism as a historically and culturally contingent cinematic style see Julie Turnock, ‘The ILM Version: Recent Digital Effects and the Aesthetics of 1970s Cinematography,’ \textit{Film History} 24, no. 2
the spectator of the reality of the virtual scene depicted in the image, hiding the presence of both the digital artist and the software used in the process of its construction. Price’s videos are aimed at aspiring digital artists, and his ideas offer a valuable insight into the way photorealism is conceptualized by those working most closely with synthetic images. For these artists and technicians, photorealism is both an artistic and technical goal that presents a constant challenge. For Price, photorealism is the ‘most important goal for CG artists.’

This understanding of photorealism reflects large-scale changes in visual culture resulting from the development of digital imaging and its widespread adoption by Hollywood and other industries. This rapidly changing media environment, with changes in both the production and reception of images, has consequences for film and media scholarship. As Stephen Prince points out, the creation of digital composites that have no extra-filmic referent challenges ‘some of the traditional assumptions about realism and the cinema which are embodied in film theory.’ In particular, he argues that theory ‘has construed realism solely as a matter of reference rather than as a matter of perception as well.’ To explore the consequences of this changing media landscape, and its implications for immersive spectatorship, I’ll make a broad distinction between three general forms of cinematic realism - ontological realism, perceptual realism, and photorealism - which can help clarify the alternative forms of realism as understood by both film theorists and by those in the visual effects community.

Ontological realism describes representations that have a clear basis in an extra-filmic reality. This form of realism is most closely associated with André Bazin, who outlined in a series of influential essays the foundations of a theory of realism that would help explain the relationship of spectators to photographic images and the remarkable impression of reality they are able to convey. In a well-known passage, Bazin describes the role of photography in this relationship:

(2012): 158–168. Similarly, Lev Manovich argues that ‘over the course of the last hundred and fifty years we have come to accept the image of photography and film as reality.’ Manovich, Language, 200-201.

201 Prince, Digital, 28.
202 Prince, Digital, 28.
203 For a similar overview of the different conceptions of cinematic realism, see Manovich, Language, 185-198.
The photographic image is the object itself, the object freed from the conditions of time and space that govern it. No matter how fuzzy, distorted, or discolored, no matter how lacking in documentary value the image may be, it shares, by virtue of the very process of its becoming, the being of the model of which it is the reproduction; it *is* the model.\(^{205}\) (original italics)

For Bazin, the object and its representation have a shared ontology, which is preserved through the automaticity of the photographic process. This realism hinges on the spectator’s knowledge that at some point in time the object existed as depicted in the representation. The representation may differ from the original scene from which it derives – it may be ‘fuzzy, distorted, or discolored’ – but the spectator’s knowledge of the photographic process ensures that this ontological connection between the object and its representation remains intact.

The contemporary ubiquity of entirely synthetic representation clearly problematizes this understanding of realism.\(^{206}\) To conceptualize images that appear realistic but that have no extra-filmic basis, Stephen Prince has developed a notion of *perceptual realism*, which shifts the focus away from the referential basis of traditional theories of cinematic realism and onto the spectator’s perception. Shifting the focus in this way can help to describe the spectator’s relationship to images that ‘have no basis in any photographable reality but which nevertheless [seem] realistic.’\(^{207}\) As an example, Prince describes the compositing of computer-generated dinosaurs with live-action footage in Spielberg’s *Jurassic Park* (1993):

> Dinosaurs are not living beings in the age of cinema. They cannot be photographed as sentient creatures. Thus their logical status in *Jurassic Park* is as objects that are referentially false. They correspond to no

---

\(^{205}\) Bazin, ‘Ontology,’ 14.


reality the film’s viewer could inhabit. And yet as depicted in the film they are perceptually realistic.\textsuperscript{208}

Though the dinosaurs cannot have been photographed in any real-world sense, they nonetheless appear in the film as they could reasonably be thought to look in reality. An important aspect of Prince’s notion therefore is that perceptual realism can be achieved with patently \textit{unreal} scenarios: perceptual realism applies to ‘both “realist” and “fantasy” films alike.’\textsuperscript{209} By definition, a composite problematizes the relationship between a representation and its referent that is central to ontological realism, but it can still have a basis in realism if it ‘structurally corresponds to the viewer’s audiovisual experience of three-dimensional space.’\textsuperscript{210} An important part of the perceptual realism of these images therefore is the success of the compositing process and the range of techniques compositors use to seamlessly join the distinct elements into a single, coherent whole. Prince demonstrates this with an example:

When the velociraptors hunt the children inside the park's kitchen in the climax of \textit{Jurassic Park}, the film's viewer sees their movements reflected on the gleaming metal surfaces of tables and cookware. These reflections anchor the creatures inside Cartesian space and perceptual reality and provide a bridge between live-action and computer-generated environments.\textsuperscript{211}

Sophisticated digital compositing techniques ensure the structural integrity of these images, working ‘to match live-action and computer environments and lend credence and a sense of reality to the compositied image.’\textsuperscript{212} Digital technology is fundamental to perceptual realism, as it represents a significant increase in the range of techniques compositors can use to join synthetic imagery with traditional photography. ‘The digital toolbox affords filmmakers ways of crafting more persuasive and convincing effects,’ Prince points out, ‘blending live action and

\begin{itemize}
\item \textsuperscript{208} Prince, \textit{Digital}, 32.
\item \textsuperscript{209} Prince, \textit{Digital}, 32.
\item \textsuperscript{210} Prince, \textit{Digital}, 32.
\item \textsuperscript{211} Prince, ‘True Lies,’ 33.
\item \textsuperscript{212} Prince, ‘True Lies,’ 33.
\end{itemize}
synthetic image elements into scenes that have greater perceptual credibility than what optical printing in the analogue era permitted.  

Julie Turnock offers a third understanding of cinematic realism that aims to challenge Prince’s account of increasing realism resulting from continued technological progress and development. One of her goals is to denaturalize what she calls the ‘ILM aesthetic’, which she argues has become normalized as a standard of cinematic realism. She argues that in contemporary digital effects, ‘filmmakers hearken back to an earlier era more closely associated with the integrity of the photographic image: to the 1970s, and specifically to 1970s materialist docurealism of the sort associated with Hal Ashby, Terrence Malick, and Monte Hellman.’ She argues that this style ‘was developed at the time to accentuate – not hide – the process of filming, and included such techniques as lens flares, handheld cameras and rack focus.’ According to this view, modern digital effects artists artificially introduce camera artifacts in an attempt to replicate an historically specific style of realism. As a result, she argues, it is important to recognize that ‘the powerfully photorealistic style of contemporary effects is in no way objectively more “natural” or “realistic” than any previous style. Instead, it is a historically specific style comprised of component parts.’

In arguing for a historically contingent style of cinematic realism, Turnock dismisses Prince’s notion of perceptual realism as divorced from the historical and cultural influences on cinematic style. ‘Since it relies upon the human eye for its controlling structuring aesthetic,’ she argues, ‘perceptual realism is a wholly inaccurate characterization of the aesthetic strategies involved in cinematic representation.’ Turnock offers valuable insights into the historical contingency of different styles of realism, and her ideas are an important check on the technological determinism that often accompanies descriptions of cinematic realism and predictions about its future. However, her dismissal of Prince’s notion of perceptual realism excludes a theory of contemporary cinema that offers valuable insight into the relationship

213 Prince, Digital, 8.
215 Turnock, ‘ILM,’ 158.
216 Turnock, Plastic, 3.
217 Turnock, ‘ILM,’ 160.
between cinematic technology and visual illusionism. The two perspectives offered by Prince and Turnock are in fact not incompatible, but rather two important and complementary contributions to the analysis of contemporary visual style. In analyzing the relationship between technological development and cinematic realism, it is important to recognize historical contingency, whilst also recognizing the ability of new technologies to provide visual experiences that are more congruent with real-world perception than previous technologies.

As an example of the value of the concept of perceptual realism, I’ll examine a significant visual effects achievement from Life of Pi, which shows how the idea can be used to conceptualize the increasing realism associated with developments in technology. The central narrative line of the film involves Pi Patel, the film’s main character, struggling to survive at sea as he is stranded aboard a lifeboat after his ship is destroyed by a storm. Aboard with him on the lifeboat is a Bengal tiger named Richard Parker. This unique situation - a boy alone with a large Bengal tiger on a small lifeboat - presented obvious practical difficulties for the filmmakers. Bill Westenhofer, the film’s visual effects supervisor, summarizes the challenges of filming with a dangerous animal and the approach the filmmakers took to solving this problem:

Even if you had a perfectly trained tiger and you could get the studio legal department and the insurance company to allow that tiger to interact with a kid, they were going to be trapped together on a small boat. There was no way that could happen. So we knew we were going to have to achieve a completely photorealistic tiger for this show, and that was our goal.218

The decision to create and animate Richard Parker digitally presented an unprecedented visual effects challenge. A photorealistic result would involve a digital tiger that is indistinguishable from a real tiger. Furthermore, shots of the digital tiger would be edited together with footage featuring real tigers, making an indistinguishable result crucial to hiding the fact that the digital tiger was created artificially.219 To achieve such a result, the visual effects artists conducted extensive research on the movements of real tigers as they appeared on camera.

---

Eight weeks of the production involved real tigers, and according to Westenhofer, the footage shot during this time gave the visual effects artists detailed real-world reference to help guide the process of creating and animating the digital tiger:

We got reference photography we would never have gotten otherwise…
No production would pay to have you sit there for eight weeks and take reference footage of paws rolling off benches - but we got all of that. We wound up with about 100 hours of reference.\(^{220}\)

This extensive use of reference footage shows how deeply the synthetic tiger seen in the finished film is grounded in real-world visual cues. The reference footage provided the artists with a clear goal, anything short of which would be identifiable as computer-generated and artificial. Through the remarkable power of digital animation technology and the talents of the visual effects artists, the film realizes the creation of a synthetic tiger that looks and moves more like a real tiger than previous attempts at such animation could achieve. In this respect, rather than being an achievement in an historically variable standard of realism, the success of the result is measured against real-world standards drawn directly from real-world perception. The extensive use of reference footage shows that cinematic realism is not wholly determined by previous styles, but is often firmly grounded in real-world perception. Real-world visual cues play an important role in determining the ultimate appearance of much computer-generated imagery, and an understanding of realism as determined by both naturalistic and historical factors is crucial to understanding the role of developments in technology in determining visual style in the cinema. Despite Turnock’s objections, perceptual realism is therefore a valuable concept in examining changes in film aesthetics as the industry has embraced digital visual effects technology.

Prince does however seem to restrict his notion of perceptual realism to digital composites, and though there is no doubt that digital technologies have greatly increased the abilities of filmmakers and visual effects artists to create believable cinematic illusions, and hide the evidence of their work, the history of effects shows that this ability is not exclusive to those working with digital technology. Though digital technology does represent a significant change

\(^{220}\) Duncan, ‘Life,’ 9.
in the abilities of effects artists to create synthetic imagery, the techniques and methods used today build on those developed through the analogue era. As Lev Manovich puts it, ‘computers do not bring any conceptually new techniques for creating fake realities. They simply expand the possibilities of joining together different images within one shot.’ The history of cinema is full of examples of compositors and special effects artists successfully compelling belief in a fictional world through various matting and multiple exposure techniques, front and rear projection, and other optical and in-camera methods of compositing. As I’ll show below, an immersive response is not entirely determined by technology, though it is an important variable. In many cases, the artists working in the pre-digital era showed remarkable ingenuity in using the most advanced technologies available to create seamless illusions and diegetic worlds in which spectators could become immersed. I will therefore extend Prince’s notion of perceptual realism to the pre-digital compositing techniques that are the focus of the next section.

Compositing in Early Cinema

A brief look at the early history of pre-digital methods of compositing shows that cinema has always made use of virtual spaces to form larger diegetic worlds. Matte shots were one of the earliest techniques used to create virtual spaces, and like later compositors of digital materials, were used by innovative filmmakers to combine elements from multiple sources into a single image. For early filmmakers working with celluloid film, a matte shot involved blacking out or covering selected parts of the frame to prevent them being exposed to light during exposure of the emulsion. This enabled the filmmaker to run the film through the camera during an initial exposure to produce an image on carefully chosen parts of the frame, leaving the masked areas unexposed. The matte could then be reversed, producing a counter-matte, allowing the film to run through the camera a second time to expose the previously unexposed areas of the frame with elements from a different source. The result was a composite of elements from the two separate exposures. As with so many other cinematic innovations, George Méliès was one of the first filmmakers to make use of the technique. His 1899 film The Mysterious Portrait, is an

221 Manovich, Language, 152.
early example of its use in a direct ‘presentational’ address to the audience, with attention directed to the technique’s novel visual effect. The 60-second film starts in a well-decorated interior with a large and elaborate picture frame standing in the center of the setting. Méliès walks around and behind the frame, and then steps through it from behind to show that it is empty. He then takes a canvas with a landscape depicted on it, and places it in the frame, and then positions a stool immediately in front of the canvas. He sits on a stool to the right and just in front of the picture. The canvas within the frame is then blurred momentarily, and an image of another Méliès, dressed identically, miraculously fades into view seated just like the original Méliès. The two figures look and gesture towards each other, their interaction indicating that they miraculously occupy the same diegetic time and space. The canvas then fades away, bringing the film to an end.

The effect was achieved with a multiple exposure and masking technique that is relatively simple but innovative for its time. Méliès placed a mask just in front of the camera lens to black out just those parts of the frame where the canvas would appear, leaving them unexposed during the initial exposure of the film. This meant that Méliès, the picture frame and everything else in the background was captured, but the section of the frame covering the actual canvas of the painting was left unexposed (see Figure 3). The matte was then reversed, with the clear sections masked and the masked sections left clear, and the film rewound through the camera, this time masking everything but the image that appears on the canvas (see Figure 4). This simple matting technique produces a composite of the two images, achieving the illusion that the two figures simultaneously occupy the same diegetic time space (see Figure 5).223

Méliès did not invent the technique but borrowed it from still photography, as did Albert E. Smith and several other early filmmakers that were using the technique at this time. In *Artificially Arranged Scenes*, John Frazer points out that by the ‘1890s the use of a masked lens to create two images of a single figure on a single negative was common practice for Victorian photographers.’ Though the technique pre-existed its first appearance in cinema, *The Mysterious Portrait* represents an early use of it to manipulate cinematic time and space. Brian Jacobson describes this as ‘cineplasticity’ and he points out that Méliès achieved a ‘cineplastic’ image ‘by using film technology to capitalize on the potential embedded in the studio’s plastic pro-filmic space.’ In the sense proposed by Jacobson, Méliès represents a precursor of the modern digital artist that seeks to manipulate the processes of production of the media form.

The malleability of material form often thought to be a new characteristic of digital media can be traced to the earliest examples of cinema.

This notion of a malleable, ‘plastic’ pro-filmic space is fundamental to understanding the ontology of compositied images, and the role of visual effects in cinema more generally. The

---

224 Oscar Gustav Rejlander’s *The Two Ways of Life* from 1857, for example, combined elements from 32 different original sources.

225 Frazer, *Artificially*, 76.


ubiquity of image and video editing software as well as the proliferation of behind-the-scenes documentaries means that effects artists, critics and audiences are increasingly aware that images appearing on screen are malleable and less bound to a stable pro-filmic reality. Turnock refers to this as ‘plastic reality’ and titles her book on the topic accordingly. She traces the contemporary plasticity of digital images to several innovations in optical printing and animation that filmmakers and special effects artists in the 1970s used to gain greater control over the image. There is no doubt that the effects-intensive cinema of the 1970s was significant in establishing many of the conventions of the special effects cinema of today, but as I point out above, this kind of image manipulation can hardly be considered unique to post-1970s cinema. The presence of such techniques in the films of Méliès and other early examples suggests that the plasticity of the image is a fundamental element of cinema that filmmakers have embraced since the beginning of the medium. Furthermore, the cinematic compositing described above demonstrates the interrelationship of technology and aesthetics that has been fundamental to all forms of cinema, and shows one example of how a specific technological development was used to produce an innovative aesthetic mode.

Figure 5
These ‘trick films’ of Méliès demonstrate the role of compositing in constructing a presentational image that works against the immersive ideals of the classical style. According to Gunning, ‘Contemplative absorption is impossible here. The viewer’s curiosity is aroused and fulfilled through a marked encounter, a direct stimulus, a succession of shocks.’

He elaborates on the specific set of spectator relations characterizing this ‘cinema of attractions’:

Rather than being an involvement with narrative action or empathy with character psychology, the cinema of attractions solicits a highly conscious awareness of the film image engaging the viewer’s curiosity. The spectator does not get lost in a fictional world and its drama, but remains aware of the act of looking, the excitement of curiosity and its fulfilment. Through a variety of formal means, the images of the cinema of attractions rush forward to meet their viewers.

The appeal of such films is based on the recognition of an illusion, and thus requires the viewer’s knowledge of the deception that is taking place. This is an entirely different set of spectator relations from the psychological engagement characterizing the narrative style that was soon to develop, and to which newly-developing effects techniques would soon be directed. The separation of the viewer’s space and the space of the diegesis was intrinsic to this ‘presentational’ mode of address, and as Manovich writes, ‘viewers were free to interact, come and go, and maintain a psychological distance from the cinematic narrative.’

Innovative filmmakers quickly recognized however that the same effects used in these early presentational films could be used to construct the illusion of a virtual space in order to deepen narrative immersion. Perhaps the earliest example of this is Edwin S. Porter’s *The Great Train Robbery*, which was shot during November 1903 and first shown in early December. Charles Musser notes that the film achieved unprecedented commercial success, explanations for which range from its status as the first significant Western incorporating story and myth from

---

American history, to its technical achievements. Musser himself attributes the film’s success to its presentation of ‘so many trends, genres, and strategies fundamental to cinematic practice at the time.’ The film appeared after the earliest years of cinema where a fascination with the illusion of movement was the basis of cinematic spectatorship, but it came before cinema had become a predominantly narrative medium. It thus has elements from both the ‘cinema of attractions’ and the emerging narrative foundation of the later classical style. As Gunning suggests, the film points in both directions, ‘toward a direct assault on the spectator (the spectacularly enlarged outlaw unloading his pistol in our faces), and towards a linear narrative continuity.’

Figure 6 shows the interior of a train station as outlaws break in to hold up a train that is about to arrive. On the right of the set there is a window through which the train can be seen arriving at the station and coming to a stop. The scene consists of an interior set inside the station, and the station’s exterior as seen through the window. The scene is represented as a single space, but is a composite of two elements: the interior of the station and the train outside the window were filmed at different times and have been composited together. The scene uses the same technique as the earlier Méliès film to produce an illusion of a coherent and seamless diegetic space, but here the effect is pushed to the background, far removed from the ‘aggressive’ mode of address characteristic of an attraction. The compositing here contributes to the creation of a seamless diegetic world and the illusion of a pro-filmic space that can be photographed by a single camera. There is an attempt to convince the spectator that the pro-filmic space pre-exists its representation by the apparatus of cinema, and the compositing thus represents an important historical move towards the creation of a virtual space.

---

233 Gunning, ‘Attractions,’ 70.
This scene uses the same compositing technique as *A Mysterious Portrait*, demonstrating a similar malleability of the pro-filmic space, but for an entirely different aesthetic end. The compositing here is used to direct the spectator’s attention away from the technology on display and constructs instead a larger cinematic world in which a narrative takes place. The scene demonstrates one of the earliest attempts to use basic matting techniques to achieve one of the fundamental goals of effects technicians throughout the classical period. As Musser describes it, the effect is used here ‘to introduce exteriors into studio scenes.’

This desire to create complete environments comprising interiors and exteriors is a process and motivation that drove the development of several key compositing techniques throughout the classical period, including matte painting, glass shots, traveling mattes, and rear projection, and it continues to motivate the use of techniques such as set extensions and chroma keying that form a large part of the modern digital compositor’s workload. The key motivation for the artist-technician is to

---

achieve a balance between the potentially obtrusive use of the effects technique and the ideal transparency of a classical compositional style.

In shifting from the presentational mode of address shown in *A Mysterious Portrait*, to a representational mode, *The Great Train Robbery* signals the beginning of what Turnock calls the ‘studio functionalist aesthetic’, a key feature of which is a diegetic space ‘where all elements are in the right perspective and are as unobtrusive as possible, forming a seamless backdrop for the all-important actors in the foreground.’ Visual effects at this time followed other potentially self-conscious stylistic devices, such as camera movement, in becoming predominantly subordinate to the narrative imperative of the classical style. As the goal of effects artists was to achieve transparency by blending the disparate elements as seamlessly as possible to give the illusion that the composited elements comprised a singular and whole diegetic space, subsequent technological development in compositing largely consisted of attempts to minimize the artifacts that threatened this illusion, but as the following section will show, this imperative often conflicted with other contingencies of filmmaking, such as production efficiency and cost. Rear projection is one historical example of the tension that arose between the technological apparatus used to mediate the virtual worlds intrinsic to popular cinema, and the seamless aesthetic demanded by this form.

**Rear Projection and Hitchcockian Modernism**

Rear-projection involves projecting background footage against a large translucent screen positioned behind the actors and any foreground elements. This produces a composite of the two images, with the camera capturing both the foreground actors and the projected background simultaneously. It largely replaced earlier compositing methods such as the Williams and Dunning processes, which had been used to combine moving background and foreground elements, and there are a number of reasons for why the process became the compositing standard for studios in the early 1930s and remained so throughout the decades in which it was in use. One of the reasons for its widespread adoption at this time is that compositing a rear-screen

---

projection and a foreground element is a highly technical process, and the technology necessary to produce such composites was not developed until the late 1920s. Another reason was the introduction of sound in the late 1920s. The need for quality sound recording meant increased pressure for filmmakers to control the environment in which they were shooting, which was much easier in a studio than on location. Rear-projection thus emerged as a means of reproducing the varied settings typical of location shooting but in the controlled environment of a studio.

These gains in production efficiency represented industrial reasons for the use of rear projection, and given the industrial/economic imperative driving much of Hollywood cinema this seems a reasonable explanation for the technique’s prominence. But the perceived aesthetic shortcomings associated with rear-projection raise questions about its long-term use throughout the classical era. Judged against the contemporary standards of photorealism set by modern digital compositing, rear-projection is highly problematic as a visual effects technique that ostensibly aims at achieving some level of transparency in its depiction of a diegetic world. Obvious differences between foreground and background elements, including in color saturation and image resolution, and in disjointed lines of perspective between the two image planes, means that contemporary audiences are quick to notice its inherent artifice. Its artifice makes it, in Turnock’s words, ‘a technology contemporary viewers often consider mundane at best and embarrassing at worst.’ However, despite being problematic for modern audiences, the technique was used consistently throughout the classical period. As Manovich writes:

---

238 Farciot Edouart, Paramount’s special effects and process photography specialist throughout the studio era, suggests there were 3 specific technologies that enabled rear-screen projection at this time. These include a variety of ‘electrical hook-ups’ that enabled the synchronization of the background projector and the foreground camera shutters, more sensitive panchromatic emulsions that gave the necessary speed to photograph the rear-projected image, and improved light-sources that increased the brightness of the rear-projected image. Farciot Edouart, ‘The Paramount Transparency Process Projection Equipment,’ *Journal of the Society of Motion Picture Engineers* XL, (June 1943): 368-369.

239 Eliminating the background noise that comes with location shooting was particularly important in the early years of sound given the limitations of the microphones available at the time. Salt writes that as ‘all the microphones in use were omni-directional (that is, they responded to sounds whatever direction they came from), background noise pick-up could be a serious problem with location recording’ Salt, *Film Style*, 234. See also Fielding, *Technique*, 248.

In general, Hollywood cinema has always been careful to hide the artificial nature of its space, but there is one exception: the rear-screen projection shots introduced in the 1930s … The artificiality of rear-screen projection shots stands in striking contrast to the smooth fabric of Hollywood cinematic style in general.241

The technique’s prominence during the classical era is particularly puzzling given its presence in the films of Hitchcock and other classical-era filmmakers that almost invariably upheld the classical style, using other formal elements of cinema, such as continuity editing, to invisibly construct a diegetic world. But it is this discrepancy between the invisibility of the Hollywood style and rear-projection’s garish artificiality that makes the technique an important point for examining the historical contingency of standards of medium transparency. There is evidence to suggest that audiences from the early 30s to the late 60s, when the technique was in widespread use, were satisfied with the composites it produced.242 How could the popular reception of a cinematic technique change so drastically? Analyzing the development and reception of rear-projection, and attempting to explain why it persisted for so long as a standard compositing technique despite its perceived aesthetic shortcomings, shows that audience standards concerning the visibility of rear-projection, and the ideal of transparent technique more generally, have some degree of historical variability.

Hitchcock’s cinema is often the focal point around which much of the scholarly literature on rear-projection focuses, no doubt in part because the director’s career spanned those decades when the technique was in most widespread use. But the director’s use of rear-projection is also an important point in these discussions as he was a Hollywood director who upheld the conventions of the classical style, but one who also demonstrated a degree of formal experimentation throughout his career. His use of the technique thus reveals important insights into the tension between medium transparency and obtrusiveness, and the degree to which formal experimentation would be tolerated within the parameters set by the classical style. Critical

---

disagreement on his use of rear projection tends to be polarized, with some seeing it as excessive and brazen, only serving to rupture the diegetic illusion rather than augmenting and sustaining it, while others viewing it as an essential component of his directorial style. Of the former, Danks argues that this is a common dismissal of rear projection as a dated and aesthetically suspect technique, which ‘relies upon a teleological understanding of film history that favors realism over a more constructivist and synthetic notion of cinematic form.’

This idea certainly problematizes contemporary dismissals of Hitchcock’s use of rear-projection, particularly those dismissals that focus on a perceived lack of realism, but the situation is more complex when considering Hitchcock’s use of the technique later in his career, when ‘it was already considered an outdated special effect.’

Marnie is a film around which there is considerable critical debate in this regard. According to Danks, the film’s release in 1964 came at a time when ‘the technique of rear-projection was becoming somewhat anachronistic and jarring, particularly at the point where many films largely moved out of the studio and onto location.’

In a review for The New York Times, Eugene Archer wrote of ‘glaringly fake cardboard backdrops,’ and suggested that ‘the best technician in the business has faltered.’

Andrew Sarris criticized the film and wrote that the ‘fake sets … have never been more distracting, and the process shots of Tippi on horseback are appallingly dated.’

Others however, justify Hitchcock’s obtrusive use of the technique in Marnie on aesthetic grounds. For example, Yacowar argues that perceived technical errors in Hitchcock’s work are deliberate artistic choices that manifest the director’s modernist aesthetic:

Where Hitchcock’s technical work seems shoddy, what we really have is not a craftsman nodding but an artist extending his resources. Where

---


245 Danks, ‘Being,’ 68.


Hitchcock’s craft seems loose, we usually find his technique subserving his content, his literal realism shading into vibrant metaphor.\textsuperscript{248}

For Yacowar, the artifice of the rear-projection work in \textit{Marnie} perfectly complements the central character’s fractured identity. He argues that ‘the false backdrops in \textit{Marnie} are a concise image of the heroine’s predicament: she lives in dislocation from her surroundings and from her own past. The false backgrounds provide a physical expression of the disjunction in her mind.’\textsuperscript{249} One of the film’s strongest defenders is Robin Wood, who describes it as ‘one of Hitchcock’s richest, most fully achieved and mature masterpieces.’\textsuperscript{250} Wood summarizes the response of critics that dismiss the film as the technically dated product of a disinterested director:

\begin{quote}
The film is full of absurdly clumsy, lazy, crude devices, used with a blatant disregard for realism: hideous painted backdrop for Mrs. Edgar’s street, ugly and obvious back projection for Marnie’s horse-riding… zoom lens for the final attempted theft, red flashes suffusing the screen every time Marnie has a “turn”; thunderstorms arriving coincidentally at climactic moments.\textsuperscript{251}
\end{quote}

Wood argues in response to these critics that the obtrusive rear projection complements the film’s other formal experimentation, and writes that \textit{Marnie}, along with \textit{Vertigo}, is one of ‘the most extreme manifestations’ of Hitchcock’s ‘achievement of “Pure Cinema”.’\textsuperscript{252} Hitchcock’s ‘Pure Cinema’ Wood writes, was ‘the art of putting pieces of film together to create effects.’\textsuperscript{253} It is an aesthetic that prioritizes cinema’s affective qualities over any attempt to accurately recreate the real world. According to Wood, much of the criticism of \textit{Marnie} is an unjustifiable response to its failure to adhere to a realist aesthetic:

\begin{flushright}
\textsuperscript{249} Yacowar, ‘Imagery,’ 30.
\textsuperscript{251} Wood, \textit{Revisited}, 173.
\textsuperscript{252} Wood, \textit{Revisited}, 389.
\textsuperscript{253} Wood, \textit{Revisited}, 389.
\end{flushright}
There is absolutely no valid reason why a film should be “realistic,” why a director shouldn’t use obviously painted sets, back-projection, zoom lens, etc., if the context justifies these things. The question is, then, not “should a serious director stoop to these?” but “Do they work?”

For many audiences and critics at the time of the film’s release, however, these formal experiments did not work. Donald Spoto, for example, in *The Dark Side of Genius*, describes these formal experiments as ‘visually jarring; they mark not a deliberate use of unconventional means, but are simply unpleasant examples of the director’s cavalier disinterest in the final product.’ He describes the defenses of obtrusive technique from Wood and others as ‘tortuous arguments more admirable for their ingenuity than consistent with the facts.’

The diversity of opinion on this aspect of Hitchcock’s cinema is no doubt in part a result of the director’s somewhat unique position as a classical filmmaker working in the tradition of popular studio cinema, but one who also exercised an unusual degree of authorial control. This position is reflected in his films, which tended to follow the conventions of the classical style, but which also at times demonstrated a degree of modernist formal experimentation. Claims of conspicuous technique as a deliberate aesthetic choice are perhaps understandable in relation to a film such as *Marnie*, for which there appears to be thematic justification for the use of a technique that draws attention to the medium. However, Hitchcock’s use of obtrusive technique in other films, particularly those that conformed more closely to classical convention, are more problematic, raising questions about the tension between the invisibility of the classical style and the technical constraints of the medium at particular times in its technological history. The rear projection in *Marnie* works for many critics because it complements the film’s other obtrusive formal elements, including the zoom lens, the use of a red flare during Marnie’s disturbing flashbacks, and the painted backdrop of the Baltimore docklands outside Marnie’s mother’s home, all of which, according to Wood, the film ‘obviously flaunts rather than seeks to disguise.’ They can thus be justified as deliberate breaks in the principle of transparency in the

---

256 Spoto, *Genius*, 476.
service of an alternative aesthetic ideal. However, Hitchcock’s earlier use of rear projection, particularly its use in films that do not demonstrate the formal experimentation of *Marnie*, shows that he frequently used it in an attempt to uphold the classical style’s principle of transparency, rather than in a modernist exposure of technique. The artifice inherent to rear projection is more problematic in such films.

This tension between the transparent ideals of the classical style and the technical limitations of rear projection can be demonstrated by examining its use in *To Catch a Thief*, a film that demonstrates both the creative possibilities and the aesthetic limitations of the technique. Like other directors at the time, Hitchcock likely used rear projection so frequently in his films because of its production benefits and the gains in creative control it represented. But Hitchcock’s aversion to location shooting, and his concern with controlling all aspects of the filmmaking process, made rear projection an ideal match for his directorial style. It enabled him to obtain images of the exotic locations he often sought with the control that came with shooting in a studio. The comments of rear projection technicians at the time suggest a delicate balancing act in uses of the technique, with the resulting artifice an unfortunate but necessary consequence of these significant production benefits.

In *To Catch a Thief*, rear projection is used primarily for its ability to reproduce the exotic images of the French Riviera, augmenting the other spectacular VistaVision location photography. One of the film’s unique qualities is its ability to induce in the spectator a spatial immersion in its exotic setting, which Patrick McGilligan reports in his biography of Hitchcock as among the Hitchcocks’ ‘favorite places in the world.’ Hitchcock in fact sent the film’s

---

258 In *Hitchcock and the Making of Marnie*, Tony Lee Moral shows that during production Hitchcock insisted he was satisfied with the quality of these effects, despite the protestations of many of the crew involved in their creation. Tony Lee Moral, *Hitchcock and the Making of Marnie*, (United Kingdom: Scarecrow Press, 2013), 96-100.

259 Bill Krohn in *Hitchcock at Work* shows that though the popular myth that Hitchcock was a ‘control freak who pre-planned every shot of his films’ is exaggerated, he was nonetheless uncomfortable with the ‘chaotic element’ of filmmaking and did attempt to minimize it through extensive pre-planning. Bill Krohn, *Hitchcock at Work*, (Phaidon Press, 2000), 9.

260 Turnock suggests that though special effects technicians recognized the aesthetic limitations of rear projection, and continued working to improve the technology, directors, producers and cinematographers often had different aesthetic standards. The low status of technicians in the studio hierarchy meant that production efficiency was given priority over their aesthetic concerns. Turnock writes that as a result, rear projection is ‘perfectly consistent with the Hollywood studio production system, but not with its ideal seamless aesthetic.’ Turnock, *Rear Projection*, 162.

writer, John Michael Hayes, to Nice before they began the writing process, in an effort to infuse a degree of authenticity into the script. Also key to the construction of this sense of place are the film’s many scenes set in moving vehicles, which allow the audience to travel through the setting, and allow the film to function as a ‘touristic spectacle.’

Shooting took place on location on the French Riviera principally around Cannes and Nice, with additional studio shooting taking place in Hollywood, after principal photography had concluded. It is both this exotic setting, far away from the Paramount studio in Hollywood, and the film’s numerous scenes set in moving vehicles, that necessitates such extensive use of rear projection. McGilligan reports that soon after the location shooting commenced, dissatisfied executives at Paramount ordered Hitchcock to shoot just the exteriors and the rear-projection plates on location, and save all close-ups for the studio. According to McGilligan, Hitchcock responded by ‘shooting both close-ups and rear projection plates on location, and then back at Paramount shooting alternative close-ups against the plates.’

The resulting film is thus made up of exteriors, rear-projection plates and close-ups shot on location, with additional close-ups shot back in Paramount’s Hollywood studio. This left the responsibility of producing the ‘final blend of reality and artifice’ to the postproduction work of Farciot Edouart, Paramount’s rear projection specialist, and John P. Fulton, who oversaw the film’s special effects.

This combination of location and studio shooting meant that extensive technical effort was needed to blend the disparate elements into a coherent whole. The film thus exemplifies the classical Hollywood style’s creation of an illusion of a coherent and internally consistent diegetic world through the assemblage of distinct elements shot in different times and places. Furthermore, it demonstrates the crucial role rear-projection played in this process. However, as with Marnie, modern audiences, with standards conditioned by a series of technological

---

Times, for example, was typical in its praise of the ‘fantastic, spectacular vistas along the breath-taking Cote d’Azur’ and the ‘shots from great heights down yawning chasms, glimpses of ruins high on hills, views across Mediterranean harbors and, usually, in the background, the blue sea.’ Bosley Crowther, ‘Cat Man Out 'To Catch a Thief'; Grant Is Ex-Burglar in Hitchcock Thriller,’ The New York Times, August 5, 1955, https://www.nytimes.com/1955/08/05/archives/screen-cat-man-out-to-catch-a-thief-grant-is-exburglar-in-hitchcock.html.

262 Spoto, Genius, 349.
263 Danks, ‘Being,’ 81.
264 Spoto, Genius, 349-353.
265 McGilligan, Hitchcock, 495.
266 McGilligan, Hitchcock, 495.
developments beginning in the late 1960s with Kubrick’s *2001*, are likely to be struck by the sharp discrepancy between the projected background and the foreground elements of the film’s process shots. One scene in particular demonstrates both the technical limitations that modern audiences find so jarring, as well as the aesthetic possibilities represented by the technology. Early in the film, John Robie, a famous burglar, and Danielle, the daughter of one of Robie’s friends, escape the police by taking a boat off the French coast and into the Mediterranean Sea. Midway through their escape, the pair engage in a long conversation whilst the boat floats out in the middle of the sea. The scene makes extensive use of rear projection, with the actors performing on a boat filmed in front of rear-projected background plates (see Figure 7). The scene demonstrates the classical use of cinema’s formal elements to facilitate the audience’s immersion in a virtual space. Each of the camera set-ups is positioned on the boat, with the line of sight directed towards the actors, and with the rear-projected plates of the Mediterranean Sea filling in the necessary background detail. The cross cutting between these setups varies the audience’s point-of-view, but limits it to a position on the boat.

For many modern audiences however, the two image planes in the rear projection used in these scenes clearly do not belong together, and the resulting composites fall short of the classical style’s ideal seamless aesthetic. One of the issues seen in figure 7 is the overbright
center of the background image. Barry Salt calls these ‘hot spots’ and points out that it was a common tendency of early rear projection for the center of the projected image to appear brighter than the outer edges. 267 There is also a clear difference in the color saturation of the two image planes, with the rich colors of the foreground contrasting with the slightly washed out background, and there is a noticeable difference in the image resolution of the two image planes. The cumulative effect of these differences is to divide the image into its separate elements, revealing the space to be a construction rather than a coherent whole, depicting a profilmic space that exists independently from its representation as a cinematic image. As Turnock describes it, referring to a different scene in the same film, ‘the foreground makes one flat plane, and the background another flat plane; there is no convincing illusion of a whole.’ 268

For modern audiences, this obtrusive rear-projection photography is a clear violation of classical cinema’s seamless ideal, though unlike Marnie, this appears not to have been a concern for the film’s original audiences. Samuel D. Berns, for example, in a July 1955 issue of Motion Picture Daily, praised the film’s ‘sharp photographic quality,’ and describes, in a direct reference to a scene that makes extensive use of rear projection, ‘an unforgettable automobile chase’ that gives ‘thrilling evidence of the VistaVision range.’ 269 In his review for The New York Times, Bosley Crowther criticized the film’s use of VistaVision, but makes no mention of the use of rear-screen projection. 270 Robert Burks was awarded the Academy Award for Best Cinematography the year after the film’s release, and in a highly critical review that described the film as a ‘significant dud’, Delmore Schwartz failed to mention the film’s process photography, which suggests that even the film’s critics were accepting of its artifice. 271 At the very least, the absence of direct comment on the process photography from these various critics suggests that its conspicuous use in To Catch a Thief did not stand out as a flagrant display of cinematic technique that prevented the audience’s spatial immersion in the diegetic world.

267 Salt, Film Style, 230.
268 Turnock, Rear Projection, 161.
270 Crowther, ‘Cat Man.’
Compositing in the Digital Era

Even though rear projection was largely accepted by 1950s audiences, the technique became dated by the early 1960s, and today has largely been replaced by alternative means of combining image elements in mainstream cinema. Rear projection followed the same trajectory as other traditional methods of combining images in being replaced by digital counterparts that can more easily perform the same functions while resulting in higher quality final composites. Digital compositing consists of a wide range of complex techniques, but in order to understand its role in popular cinema and how it came to be the industry standard, it can be conceptualized as two general processes: **matte extraction** and **integration**. Matte extraction involves preparing the individual elements that are to be combined in the final composite. This could mean isolating a foreground element, for example, such as a person, which would allow them to be placed onto a different background, effectively transposing them into a different setting. Integration refers to the various methods that are used to combine the distinct elements into a coherent and seamless whole. For example, a simple color correction altering the color of a background element so that it matches the foreground can make the two elements appear as though they are lit by the same light source, effectively providing a bridge between them.

Figures 8-9 provide a simple example of how these two general processes are used in popular cinema. The image on the left was produced on set during the filming of *Life of Pi*, and the image on the right is the resulting composite as it appeared in the film, after post-production had been completed. The image shows Pi, the film’s main character, stranded on a lifeboat after his ship has been destroyed at sea. As can be seen in the two images, the final composite consists of two distinct parts - the foreground element of Pi and the boat, and the sea and sky that make up the background. This example demonstrates chroma-keying, one of the most common methods of matte extraction, widely used in mainstream Hollywood studio production, as well as

---

272 Although chroma-keying is perhaps the most common means of achieving digital composites in contemporary practice, rear projection continues to be used. The scenes set in the *Millennium Falcon* in the *Star Wars* anthology film *Solo* (2018), for example, used pre-rendered footage projected on a large screen to depict views out of the cockpit window. In an interview during the film’s marketing and promotion, Rob Bredow, the film’s VFX supervisor, also describes using front projection and LED screens elsewhere in the film to achieve in-camera composites. ‘Rear Projection on Solo: A Star Wars Story,’ Foundry, November 7, 2018, https://www.foundry.com/insights/film-tv/rear-projection-on-solo-a-star-wars-story.
in a number of other areas including news broadcasting and video games. The technique involves filming a foreground element against a background of uniform color, usually blue or green, which can then be ‘keyed’ out during post-production. Keying out the background allows the foreground element to be placed over a different background, resulting in a composite of the two image planes. A number of additional techniques are then applied to the resulting composite to help integrate the distinct elements. In this scene from *Life of Pi*, the foreground element consisting of Pi on the lifeboat was filmed in front of a blue background which was then removed in post-production. The resulting matte could then be used to place the image in front of a different background element, in this case the ocean-sky background. The computer-generated Richard Parker has also been added, producing the final composite as shown in the image on the right.

*Figures 8-9*

As mentioned, digital compositing techniques such as these have replaced traditional methods of combining images due to several advantages they provide filmmakers. One area of advantage is the increased efficiency associated with digital compositing methods. With the rear projection seen in *To Catch a Thief*, for example, both the foreground and background were captured as a single whole at the moment of filming. The distinct components that must match between the two planes - image resolution and color saturation, as well aspects of the mise-en-

---

273 Brinkmann defines keying as ‘the process of algorithmically extracting an object from its background and combining it with a different background.’ Brinkmann, *Art*, 652.

scene such as makeup, costumes, and lighting color and direction - had to be matched prior to shooting. No further manipulation of them as individual elements was possible during post-production. In contrast, digital compositing allows endless manipulation of the distinct elements after shooting has occurred. This helps maintain consistency between the different elements, making the process of achieving a successful composite much easier. A second area of advantage in digital compositing stems from the first, and it is the improved image quality achievable with digital methods. Through the significant increases in control over the compositing process, digital methods are more able than earlier counterparts to achieve the fundamental goal of compositing - to seamlessly blend the disparate elements into a single, coherent whole. The contrast between the final composites seen in Figures 7 and 9 shows the enormous changes in the techniques and methods of compositing separating these two moments in cinema’s technological history. Both films are from major filmmakers working with the production budget of a major Hollywood studio, and both examples demonstrate the capabilities and limits of contemporary visualization technologies to create virtual spaces. The composite from *To Catch a Thief* is immediately obvious as such, while the image from *Life of Pi* successfully hides the fact that it is constructed from multiple parts. The digitally-composited image achieves the structural characteristics of a photographic image captured on a boat at sea with a real camera.

This industry-wide move to digital compositing conforms to the theory of technological development outlined by David Bordwell and Janet Staiger in *The Classical Hollywood Cinema*. Based on their extensive stylistic and industrial analysis of American studio filmmaking to 1960, the authors argue that ‘any technological change can be explained by one or more of three basic processes.’275 The first of these is *production efficiency*, which involves replacing older technologies by newer ones that make the process of filmmaking easier and more efficient. ‘A new technology might cut costs by saving time or physical capital,’ they write, ‘or it might make the results of the work more predictable, or it might solve particular production problems.’276 As outlined above, digital compositing represents significant gains in all three of these areas. Digital technology avoids many of the expenses associated with traditional celluloid, and a digital film can be endlessly copied, altered, edited, and sent electronically, all without resulting in damage

---

275 Bordwell, Staiger, and Thompson, *Classical*, 243.
or degradation to an original and unreplaceable celluloid negative. The example of the endless image alteration available with the chroma-key compositing outlined above shows that digital methods can make the results of filming more predictable, and digital technology can solve particular production problems such as those that arise with filming scenes involving a live tiger on a boat.

Bordwell and Staiger’s second process governing technological change is *product differentiation*. Studios and filmmakers have a clear economic advantage in providing audiences with novel forms of cinema, and as a result, ‘many technological innovations exploited by Hollywood were the result of careful strategies of product differentiation.’

The digital Richard Parker created for *Life of Pi*, for example, was an unprecedented technical and artistic challenge that was prominently featured in the marketing and promotion of the film. An article published in *New York Times* on November 16th, 2012, less than a week before the film was released in American theatres, provides an overview of the process of developing the digital tiger as well as excerpts from an interview with Bill Westenhofer detailing the technical difficulties of the project, which he describes as ‘something that’s never been done before, something as photo-real as anyone has ever done with an animal.’

The digital tiger was also featured prominently in the *Cinefex* article released in conjunction with the film, and the tiger itself was pictured on many of the posters used in the film’s marketing. Hollywood’s major film studios, just as in other industries, have an economic incentive to develop new and innovative experiences that set them apart from competitors.

The third process determining technological change according to Bordwell and Staiger, and the one with most direct relevance to immersive spectatorship, is *adherence to standards of quality*. ‘American cinema’s technological research,’ they write, ‘has been aimed at meeting a commitment to the standards canonized by the classical stylistic paradigm.’

An important aspect of this paradigm is a transparent medium, so a technological development that can more

---

277 Bordwell, Staiger, and Thompson, *Classical*, 244.
279 Duncan, ‘Life.’
280 Bordwell, Staiger, and Thompson, *Classical*, 244.
successfully remove itself to provide unmitigated access to a diegetic world is a more desirable technology, and is more likely to achieve industry standardization. According to this view of technological development, digital compositing methods have become ubiquitous in mainstream cinematic production due to their increased ability to hide the medium and facilitate an immersive engagement with a diegetic world.

The aesthetic paradigm governing Hollywood cinema demands the removal of technique from the spectator’s conscious awareness, allowing unfettered access to the virtual space that inheres in the composit ed image. This facilitates the immersive shift in the spectator’s conscious experience whereby the immediate concerns of the real world are displaced by the exotic virtual worlds demanded by Hollywood storytelling. But immersion in diegetic worlds is not the only aesthetic possibility of combining multiple image elements, and in this sense, earlier artistic and cultural movements can enlighten contemporary aesthetic practices and preferences. Marie-Laure Ryan recounts a prior moment in the history of visual representation when obtrusive technique interfered with the spectator’s unfettered access to the depths of an illusory space. She describes the challenge that impressionism presented to a culture accustomed to the transparent mediation of space that was conventional to visual practice of the time:

The illusion of a penetrable space received a first challenge when impressionism disoriented the eye with visible brushstrokes that directed attention to the surface of the canvas and with shimmering light effects that blurred the contours of objects. Though impressionistic space is still three-dimensional, it opens itself to virtual bodies only after the mind completes a complex process of interpretation and construction of sensory data. For the spectator who has assimilated the lesson of impressionism, visual space can no longer be taken for granted.  

In contemporary Hollywood cinema, in the era of seamless digital compositing, visual space can be taken for granted. But this was not always the case. Directors from earlier eras, working firmly within the classical system and with the budgets of major Hollywood studios, were forced to compromise between the exotic settings they sought for their diegetic worlds, and

---

the invisible technique mandated by the Hollywood style. Impressionistic rear projection during the classical era, where it did exist, was an enforced impressionism, which gifted directors were at times able to steer towards a modernist sensibility. But unlike rear projection, the failure of impressionist painting to establish an illusory world was not a lapse in technique – the disorientation to the eye caused by the visible brushstrokes is key to the impressionist aesthetic.

In a similar vein, rear projection still appears in contemporary cinema, unnecessarily drawing attention to cinema’s status as representation, a gratuitous gesture in an era of seamless digital compositing. The conspicuous rear projection in *Pulp Fiction* (1994), for example, as Butch (Bruce Willis) is driven away from the fight he has just won functions as a self-conscious device, a gesture to an earlier period in Hollywood’s technological history where such visual freedom was not available (see Figure 10). Here the obtrusive technique fits comfortably into the film’s overall aesthetic project of postmodern self-consciousness and reflexivity. As in an impressionist painting, this gesture to the process by which visual illusions are created serves a higher aesthetic end. But it remains an exception in an industry with ready access to the technology to achieve its primary aesthetic goal of immersing audiences in diegetic worlds.

*Figure 10*
Chapter 2 – Camera Movement and the Immersive Long Take

The elaborate long-take that opens Gravity was one the film’s highlights for many of its early reviewers. Critics were impressed by the technological brilliance of the sequence’s visual effects, its ability to simulate the spectator’s presence in space, and the directorial achievements involved in coordinating complex camera movements with the actors and the elaborate LED-light box that was fundamental to the film’s unique visual aesthetic. In a review for Empire magazine, describing it as the ‘film of the year,’ Ian Nathan praised the sequence’s length as ‘enough time to scatter our preconceptions of cinema across the universe before plunging us into the abyss.’

For Gus Lubin, the sequence was ‘a masterpiece’ with a ‘beautiful arc ... beginning with a shot of earth and a mighty astronaut flying above it and ending with a shot of space and a helpless astronaut falling into it.’

Mark Hughes in Forbes suggested that the long take ‘enhances the sense of place and realism,’ giving the audience the experience of ‘the weightlessness and vast distances in space in a way no other film has accomplished before.’

A general consensus saw the sequence as a powerful use of cinematic form, representing an invigorating expansion of the technological possibilities of the medium.

As can be seen from these examples, many of the early responses share an implicit embrace of a Bazinian vision of ‘total cinema,’ a phenomenological encounter with a diegetic world that approaches the real in ‘sound, color, and relief.’ In some cases, Bazin’s vision is explicit, as in this excerpt from a review by James Hoberman in the New York Review of Books:

To watch Gravity on the huge IMAX screen [is] to appreciate the power of illusion - what André Bazin described as “total cinema.” The movie is a virtual reality predicated on the plenitude of absence, the being of nothingness. In an act of technological prestidigitation, Cuarón has

---

284 Hughes, ‘Masterpiece.’
created a sense of unlimited space where the mind knows that none actually exists.²⁸⁶

Bazin’s influence is not limited to the views of the critical community, but is also evident in the filmmaker’s own comments. Speaking in an interview with Wired magazine during the film’s marketing campaign, Cuarón suggests that his use of long takes was partly an attempt to achieve the aesthetic of an ‘Imax-style Discovery Channel documentary.’²⁸⁷ This documentary approach is predicated on the unobtrusive use of the basic recording capabilities of the camera, with minimal use of the inflections of style and directorial intrusion. This includes avoiding the abstractions introduced into cinema through editing. As Cuarón points out, with reference to the film’s setting in space: ‘You don’t have the luxury of cuts when you’re in space. The camera is there; you’re just observing.’²⁸⁸ By minimizing editing in this way, the long take is uniquely able to preserve what Cuarón calls ‘a sense of real time.’²⁸⁹ In another interview with The New York Times, he specifies the spectator response he hopes to achieve with such an approach: ‘We wanted to slowly immerse audiences into first the environment, to later … immerse them into the action, and the ultimate goal of this whole experiment was for audiences to feel as if they are a third character that is floating with our other two characters in space.’²⁹⁰ It is clear from these comments that Cuarón seeks to remove the barriers separating the audience from the diegetic world. It is an attempt to maximize the illusion of an autonomous reality unfolding on screen, of which the audience is a part. It is an aesthetic approach based on the kind of immersive relationship to a diegetic world that I have so far been describing.

The relevance of Bazinian theory for a film like Gravity - a big-budget sci-fi thriller constructed almost wholly out of digital visual effects - can of course be challenged on a number of fronts. The film’s reliance on compositing virtual spaces immediately problematizes Bazin’s ideal of an objective camera faithfully capturing reality as it unfolds in pro-filmic space.

²⁸⁸ Cuarón cited in Roper, ‘Why.’
²⁸⁹ Cuarón cited in Roper, ‘Why.’
Similarly, the film’s overt genre appeals immediately foreground a range of spectatorial practices and expectations, a significant departure from the Bazinian ideal of an encounter with unmediated reality. Furthermore, the presence of Sandra Bullock and George Clooney - two of Hollywood’s highest-paid and most instantly recognizable actors - conflicts with the use of unfamiliar non-professionals that Bazin praised so highly in the cinema of the Italian neo-realists, for example.\footnote{Bazin, ‘Reality.’} But as suggested by the reviews excerpted above, and by the scholarly attention already given to the film, there is an aesthetic principle governing \textit{Gravity} that encourages a Bazinian interpretation. In particular, the film’s opening long take, and Cuarón’s general preference for sequences of extended duration, shows the respect for the continuity of space and time that is so consistently praised in Bazin’s writing.\footnote{See for example Bazin, ‘Evolution.’}

In this sense, the film represents a challenge to what is perhaps a dominant visual style in contemporary cinema, a style characterized by rapid editing and visual fragmentation that Steven Shaviro labels ‘post-continuity.’\footnote{Shaviro, ‘Post-Continuity.’} This cinema, Shaviro points out, sacrifices spatio-temporal continuity in delivering a series of shocks to the audience. It is characterized by fragmentation and disjunction, and ‘a preoccupation with immediate effects [that] trumps any concern for broader continuity.’\footnote{Shaviro, ‘Post-Continuity,’ 51.} Shaky hand-held camerawork, composited digital images, extreme camera angles, and deliberately mismatched shots are rapidly cut together such that ‘there no longer seems to be any concern for delineating the geography of action, by clearly anchoring it in time and space.’\footnote{Shaviro, ‘Post-Continuity,’ 51.} Shaviro concludes that ‘the classical values of continuity simply don’t matter to certain contemporary filmmakers anymore.’\footnote{Shaviro, ‘Post-Continuity,’ 56.} Post-continuity is an accurate descriptor of much of contemporary cinema, but the enormous commercial and critical success of \textit{Gravity}, as well as the significant scholarly attention it has already attracted, shows not only that an alternative to the fragmentation and disjunction of the post-continuity style exists, but that this alternative is appreciated by broad sections of the public, as well as critical and scholarly audiences.\footnote{As I outlined in the previous chapter, Bordwell, in contrast to Shaviro, sees contemporary cinema as largely continuous with the principles of classical storytelling. He does allow for some stylistic changes but argues that}
This chapter examines the technology and aesthetics that underlie this stylistic alternative to post-continuity, conceptualizing it as a cinematic form conducive to what Marie-Laure Ryan calls ‘spatio-temporal immersion.’ Spatio-temporal immersion involves the spectator’s relocation to the space and time of the narrative event, and I argue that the long take, the shot of extended duration, is a particularly powerful means of eliciting such a response. The long take has an extensive history in cinema, and has been conceptualized in a number of different ways by various theorists. Bruce Isaacs points out that the long take fulfills a number of philosophical and aesthetic functions, and has been subject to a variety of interpretations: ‘The desire for realism, the mark of the pro-filmic event, experiential immersion in the diegetic world, and spectatorial ambiguity have all filtered through competing discourses surrounding precisely what constitutes the long take.’ It has been seen as a mark of stylistic distinction for auteurs, functioning according to what Lisa Purse describes as ‘an extra-textual, celebratory discourse of the technological mastery of the auteur filmmaker,’ and it has a strong association with high aesthetic ideals. Mark Le Fanu for example, relates the long take to ‘a certain kind of art cinema’ epitomized by directors such as Dreyer, Bergman, and Antonioni, who offered audiences the chance ‘to linger, to explore, to risk boredom in the search for epiphany,’ a mode of spectatorship ‘that not so long ago was part and parcel of serious cinemagoing experience.’

This association of the long take with a mode of contemplative spectatorship is often traced back to Orson Welles, whose cinema, according to Bazin, ‘restored to cinematographic illusion a fundamental quality of reality - its continuity.’ I begin my discussion then with Citizen Kane (1941), a film that was arguably the first to make extended duration fundamental to its aesthetic. I then examine the cinema of Stanley Kubrick, whose directorial style included a preference for long takes shot with a moving camera. I offer an extended analysis of The Shining

298 Ryan, Virtual Reality.
302 Bazin, ‘Reality,’ 28.
(1980), a film that uses the distinct qualities of the newly-invented Steadicam to immerse audiences in the haunted spaces of its hotel setting. The Steadicam provides a distinct sensation of movement that is liberated from the material constraints that govern conventional camera movements. In combination with shots of extended duration, Kubrick uses this novel form of movement to inscribe a sense of presence - that essential trope of the haunted house film - at the level of the film’s form. And finally, I examine the digital long take in the cinema of Cuarón, a director whose work represents the culmination of the technological and aesthetic trajectory explored throughout the chapter. Cuarón shows a preference for long takes throughout his cinema, but particularly in Gravity, where he extends the available technologies of cinematic illusion in the service of a powerful spatio-temporal immersion. At the time of the film’s release, it represented the apotheosis of immersive cinematic ideals, exhibiting a mode of spectatorship made possible by the visualization capabilities of a modern digital cinema. In this sense, the film represents, as Bazin said of Italian neo-realism, a ‘progress in expression, a triumphant evolution of the language of cinema, an extension of its stylistics.’

Spatio-Temporal Immersion

Spatio-temporal immersion involves the transportation of the spectator to both the space and time of the diegetic event. According to Ryan, spatio-temporal immersion places the spectator ‘on the scene and at the time of the narrative window - to the heart of the storyworld.’ (italics original) Cinema’s status as both a spatial and temporal medium - its ability to depict both a space and a series of events occurring in that space - make it an ideal

303 My focus in this chapter is on the use of digital visual effects to extend duration, but the film’s stereoscopic imagery – its other main point of technological innovation – is also clearly central to the powerful sense of immersion. In an article on the epistemological and affective implications of stereoscopic imagery, Kristen Whissel positions the film within a longer historical tradition of stereoscopy, in which positive parallax, or perception into the depths of screen space, aligns with ‘the desire to see and know’, and negative parallax, or the perception of images into theatre space, aligns with ‘heightened emotion and sensation.’ (236) She concludes that the film ‘provides the spectator with the illusory experience of sensory immersion within the radically inaccessible, sublime location of outer space – a setting that constitutes the limits of the visible and knowable and makes locatability a matter of life and death.’ Kristen Whissel, ‘Parallax Effects: Epistemology, Affect and Digital 3D Cinema,’ Journal of Visual Culture 15, no. 2 (2016): 233. For more on the immersive implications of 3D technology, see Carter Moulton, ‘The Future is a Fairground: Attraction and Absorption in 3D Cinema,’ CineAction 89 (2012): 4-14.

305 Ryan, Virtual Reality, 93.
medium for producing a powerful spatio-temporal immersion in the spectator. The manner in which those events are narrated is central to the spectator’s relationship with them. According to Ryan, a fundamental aspect of narration is its ability to vary the spectator’s proximity to the events being narrated. ‘One of the most variable parameters of narrative art,’ she writes, ‘is the imaginative distance between the position of narrator and addressee and the time and place of the narrated events.’ This has significant implications for an immersive aesthetics. As Ryan suggests, ‘Spatio-temporal immersion takes place when this distance is reduced to near zero.’

An examination of the temporal dynamics of cinematic narration can help to understand how spatio-temporal immersion can be induced. According to a traditional view of cinematic narration, a view Andre Gaudreault describes as a ‘blind spot’ of film theory, communication in the cinema exists solely in the present tense, with the cinematic narrator limited to depicting narrative events at the time and place in which they are occurring. Gaudreault has shown however that cinema offers greater flexibility concerning narrative temporality than understood in this view. Gaudreault divides storytelling into ‘narration’ and ‘monstration’, the former applying most readily to written or scriptural narrative, which has a clearly identifiable narrator responsible for producing the words that communicate the narrative, and the latter applying to theatrical narrative, which has no clearly identifiable narrator and where the ‘flesh-and-blood characters’ move about on the stage ‘quite autonomously.’ Written narrative demonstrates a wide range of temporal registers, allowing flexibility in how the narrator positions the spectator in relation to narrative events, while monstration is limited to the present tense, communicating events as they are occurring in real time. Gaudreault points out that though cinema tends to be monstrositive, a degree of temporal flexibility is possible and can be achieved by placing a series of individual shots into a longer sequence. Each of the individual shots exists in the present tense, but once placed into a sequence they can acquire a more variable temporality. ‘One must recognize that if the shot is in the present tense,’ he writes, ‘then certain editing operations (though clearly not all) allow for the mastery of time which is one of the possibilities of all

---

306 Ryan, Virtual Reality, 93.
307 Ryan, Virtual Reality, 93.
309 Gaudreault, ‘Narration,’ 29.
narrational activity." By joining together individual shots cinema can approach the temporal range of written narrative.

The relationship between narration and monstration, and the implications of their interplay for an immersive aesthetics, can be demonstrated through the famous sequence from *Citizen Kane* (1941) depicting Susan’s failed attempt to establish a career as an opera singer. The first half of the sequence is a montage of fades between various shots. Welles cross-fades shots of Susan performing on stage, shots of Susan rehearsing, and headlines in Kane’s newspaper *The Inquirer* falsely announcing Susan’s success as an opera singer, condensing 3-years of storytime into a 40-second sequence. By placing these individual shots into a sequence, Welles is able to convey key story information through an abstraction of space and time that Bazin describes as ‘an abstract metaphorical or symbolic montage.’ The cross-fading of the shots, the abstraction of space and time, and the headlines that overtly signify key story elements all indicate the presence of a narrator, characterizing the sequence with the past tense of overt narrational communication.

In contrast to this, the second half of the sequence shifts abruptly to the time and place of a specific narrative event: Susan’s suicide attempt. This sequence shot transports us to the event’s time and place, in a demonstration of the Wellesian long-take aesthetics that Bazin places at the center of his realist phenomenology. We are positioned at the side of Susan’s bed, with a medicine bottle and drinking glass in the foreground, and with Susan lying on the bed and breathing heavily in the midground. Far in the background of the shot is the closed bedroom door, to which our attention is soon directed by a knocking sound (see Figure 11). After the knocking intensifies, Kane breaks through the door, and rushing to Susan’s side discovers the truth of the situation: after three years of continuous pressure and repeated failure, Susan has finally broken down, acknowledging the futility of her attempt to achieve success as an opera singer. The power of the shot lies in the abrupt shift in temporality represented by its juxtaposition with the montage in the first half of the scene. As Bazin describes it: ‘after the

---

310 Gaudreault, ‘Narration,’ 32.
series of superimpositions encapsulating three years of torture for Susan and ending on a light going out, the screen thrusts us brutally into the drama of Susan’s suicide attempt.\footnote{Bazin, ‘Orson Welles,’ 81.} In contrast to the past tense of the clearly foregrounded narrational activity of the montage, the sequence shot is monstrative: the action unfolds in the present tense, which, as Ryan indicates, is ‘inherently more immersive than the past.’\footnote{Ryan, Virtual Reality, 97.} We are positioned as witnesses to the drama, learning of Susan’s despair with Kane as he discovers it in real-time. The sequence shot reveals the truth of Kane and Susan’s fractured relationship in a drastic shift from the symbolic and abstract to the concrete and real, taking us to the time and place of the diegetic event.\footnote{Robert Carringer, in The Making of Citizen Kane, notes that this shot, celebrated by Bazin for its realism, is in fact an in-camera matte shot, filmed using the process I described in chapter 1. Carringer describes the process used to achieve the shot: ‘First, the foreground was lighted and focused, and shot with the background dark. Then, the foreground was darkened, the background lighted, the lens refocused, the film re-wound, and the scene reshot.’ Carringer concludes that Welles is not a ‘photographic realist’ but a ‘master illusionist.’ Robert L. Carringer, The Making of Citizen Kane (Berkeley: University of California Press, 1985), 82.}

According to Bazin, the sequence shot originates with Orson Welles’ first two films, Citizen Kane (1941) and The Magnificent Ambersons (1942), and he argues that it represents ‘a
new unit in film semantics and syntax. He attributes Welles’ preference for shots of extended duration to his background in the theatre:

One may imagine that the intuition of the sequence shot ... grew out of the vision of a director accustomed to placing the actor within the decor, who experienced traditional editing no longer as a fluency or language but as a loss of efficacy, a mutilation of the spectacular possibilities of the image. For Welles, each scene to be played forms a complete unit in time and space. The acting loses its meaning, is deprived of its dramatic blood like a severed limb, if it ceases to maintain a living and responsive connection with the other characters and the decor.

The sequence shot, as Bazin points out, provides an aesthetic experience more congruent with real-world perception than a conventionally edited sequence of individual shots. When describing the sequence shot, he writes that ‘independently of the contents of the image, its structure is more realistic.’ By eliminating the need for editing, deep focus cinematography brings ‘the spectator into a relation with the image closer to that which he enjoys with reality.’ Deep focus is based on ‘a respect for the continuity of dramatic space and, of course, of its duration,’ and is therefore ‘far superior to anything that could be achieved by the classical “cut”.’ The ability to capture both the foreground and background image planes in crisp, sharp focus, is so significant for Bazin, that he famously describes its addition to the range of available cinematic techniques as a ‘dialectical step forward in the history of film language.’

Deep focus cinematography forms just one part of Bazin’s realist phenomenology. Another key part is camera movement, which also enhances realism by ensuring temporal continuity. In his discussion of Renoir’s Grand Illusion (1937), Bazin praises the mobile camera’s ability to eliminate the need for editing, which he believes separates reality ‘into

---

316 Bazin, Orson Welles, 68.
317 Bazin, Orson Welles, 68.
318 Bazin, ‘Evolution,’ 35.
319 Bazin, ‘Evolution,’ 35.
320 Bazin, ‘Evolution,’ 34.
321 Bazin, ‘Evolution,’ 35.
successive shots which [are] just a series of either logical or subjective points of view of an event. As an example, he describes the sequence from the film in which the camera looks down from a high angle on Cartier in the courtyard as he shouts off-screen to an unseen prisoner (Figures 12-15). As their brief dialogue comes to an end, the camera pulls back to reveal first a window frame through which the camera has been filming, and then the interior of a room where two prisoners are seated, one of whom was shouting to Cartier moments earlier. The movement shifts the focus of the scene from Cartier in the courtyard to the conversation between the two prisoners in the room above, effectively linking two dramatically and spatially distinct scenes. Bazin sees this kind of movement as fundamental to preserving dramatic and phenomenological unity. ‘By moving the camera to “reframe” the scene instead of cutting,’ he writes, ‘Renoir is able to treat the sequence not as a series of fragments but as a dramatic whole.’ As with deep-focus, realism is achieved by eliminating the need to introduce an ‘obviously abstract element into reality’ through editing. ‘It is through such techniques,’ Bazin argues, ‘that Renoir attempts to portray realistically the relations between men and women and the world in which they find themselves.’

Figures 12-15

Proprioceptive Aesthetics

The immersive implications of camera movement go far beyond linking spatially distinct scenes and ensuring spatial and temporal continuity. The spectator’s sensation of movement

---

322 Bazin, ‘Reality,’ 28.
324 Bazin, ‘Reality,’ 28.
325 Bazin, Renoir, 64.
through a virtual space also has a powerful visceral quality that fully activates the body’s sense of movement. Due to the varied effects camera movement can produce, it has become central to a wide range of cinematic forms, ranging from the classical style that aims to transparently depict events occurring in a specific time and place, to more abstract styles that avoid any kind of figurative representation. In an article on the historical development of camera movement, John Calhoun points out that the ability to ‘reframe a continuous image’ is in fact unique to cinema, and filmmakers were experimenting with movement as early as the late 1890s, which suggests the technique has been central to the medium since its inception.\textsuperscript{326} A short sequence depicting movement in \textit{Panorama du Grand Canal pris d’un Bateau} from 1896, ‘likely … the first travelling shot in the history of cinema,’\textsuperscript{327} was achieved by attaching the camera to a gondola and capturing a series of buildings as the camera floated along the canals of Venice. The shot is typical of early examples of movement, which used available means of transport as the method of moving the camera, due to an absence of panning heads and other specialised equipment. One of the most popular methods of achieving movement was attaching the camera to a train, and Nielsen points out that early in the 20th century there developed two-subcategories of train mobility: ‘panoramas’ – shots filmed from the side of a moving train, which could articulate a spatial layout – and ‘phantom rides’ – shots filmed from the front of a moving train, which offered the thrill of ‘spectacular viewing positions.’\textsuperscript{328} These examples of movement provided audiences with a quick and temporary thrill, in accordance with the ‘cinema of attractions,’ but the storytelling possibilities represented by a mobile camera’s ability to articulate a spatial layout and describe a setting were soon recognized, and the technique was quickly integrated into the narrative paradigm that was to become predominant around 1907. Barry Salt points out that most movements at this time were unexceptional and were most likely used to ‘keep the action in frame,’\textsuperscript{329} but more ambitious camera movements became common after around 1915-16, influenced by the Italian film \textit{Cabiria} (1914), whose extensive camera movements, according to Roger Ebert, ‘helped free movies from a static gaze.’\textsuperscript{330} The introduction of sound in the late

\textsuperscript{326} John Calhoun, ‘Putting the “Move” in Movie,’ \textit{American Cinematographer} 84, no. 10 (2003): 73.
\textsuperscript{327} Jakob Nielsen, ‘Camera Movement in Narrative Cinema: Towards a Taxonomy of Functions’ (PhD diss., University of Aarhus, 2007), 89.
\textsuperscript{328} Nielsen, ‘Camera Movement,’ 89.
\textsuperscript{329} Salt, \textit{Film Style}, 88.
1920s posed a challenge to the mobile camera, as adding sound recording equipment to the basic production setup added significantly to the size and weight of the apparatus, making movements more difficult and more expensive to achieve.\textsuperscript{331} Despite this significant production obstacle, Hollywood filmmakers continued to move the camera, and Salt points out that there was in the end remarkably little discontinuity in the use of camera movement across the transition to sound, which, he argues, attests to ‘the vigour with which a burgeoning fashion could be pursued in the face of technical obstacles.’\textsuperscript{332}

Bordwell suggests that the quick adoption of the technique and its continued use was due to several distinct effects that made it central to classical forms of cinematic storytelling. He argues that camera movement ‘was needed for a specific representational system, that of an “illusory realism” related to narrative time and space.’\textsuperscript{333} Reframing an image can remove spatial incoherencies from a scene, clarifying the relative layout of people and objects in the diegetic space. It also tends to specify a ‘unified perceptual viewing position’, providing ‘a powerful surrogate for the active locomotion which we surrender upon settling into our cinema seat.’\textsuperscript{334} It can also produce an unbroken view of distinct narrative events, achieving a Bazinian continuity in the depiction of the diegesis. These features - coherent space, unified viewing position, and narrative continuity - were, as Bordwell writes, ‘canonized by the classical narrative style of filmmaking.’\textsuperscript{335} Camera movement was therefore quickly integrated into classical modes of cinematic storytelling and has remained central to mainstream filmmaking practice.

Camera movement as described here functions according to the classical style’s preference for transparent continuity in depicting space and time. But Bordwell also suggests that a mobile camera can provide a powerful simulation of the spectator’s own movement, providing a substitute for the ‘active locomotion’ we surrender as a necessary precondition of cinematic

\textsuperscript{332} Salt, ‘Film Style,’ 204.
\textsuperscript{333} Bordwell, ‘Camera Movement,’ 30.
\textsuperscript{334} Bordwell, ‘Camera Movement,’ 30.
\textsuperscript{335} Bordwell, ‘Camera Movement,’ 30.
spectatorship. Michele Guerra sees this aspect of a moving camera as essential to the spectator’s immersion in the diegesis:

The reason filmmakers started moving the camera was to emphasize the presence of the viewer within the shot: while the still camera provided just a strong impression of reality, by adding movement to photography camera movements highly changed the mode of presence of film, because of their implied anthropomorphism.336

This view of the immersive potential of camera movement is tied to classical forms of storytelling, but here Guerra is expanding the understanding of camera movement beyond its role in revealing story elements and spatial layout, to include the embodied responses it activates. According to Guerra, this appeal to the spectator’s sense of embodiment is what facilitates the powerful sense of immersion, the ‘presence of the viewer in the shot,’ that a mobile camera can provide.337 In its more potent manifestations, this kind of camera movement activates the spectator’s sense of proprioception, the perceptual system we use to orient ourselves in space. According to Scott Richmond, proprioception is central to the way a spectator constructs and understands the world of a film. ‘Proprioceptive aesthetics,’ he writes, ‘lies at the heart of the cinema as an aesthetic medium and as a technical system - in both its historical continuity and its contemporary uses.’338 Films such as 2001 and Gravity, which are profoundly interested in the relationship between the spectator’s body and the diegetic space, make proprioception their governing aesthetic principle. These films maximize the sensation of immersion in the diegesis of which ‘reframing a continuous image’ is capable.

337 Guerra, ‘Modes,’ 6.
2001 was revolutionary in its exploration of the aesthetic and philosophical implications of the movement of the spectator’s body, as noted by Annette Michelson in her 1969 essay *Bodies in Space: Film as Carnal Knowledge*, perhaps the most celebrated piece of critical writing on the film. Writing three decades after this early enthusiastic review, she describes the film as a ‘magnum opus’ that marked ‘an important moment in the history of cinema.’ At the centre of Michelson’s analysis is the film’s ‘corporeality’ and ‘the manner in which it implicates the spectator’s body.’ In this sense, the film promotes an aesthetic response centered around the experience of space, in contrast to a more classically-inflected cinema that relies on the temporal dynamics of a narrative. This use of movement to produce a powerful embodied response can be seen when Bowman makes the first attempt to check the antenna control device that HAL reports is malfunctioning (see Figure 16). During Bowman’s movement through space towards the antenna, the distant stars move slowly from the bottom of the frame to the top, a significant effects challenge involving coordinating multiple image elements into a final composite, but an effect that is essential in creating a gentle sensation of movement in the spectator. In combination with a range of other formal devices, including a breathing sound on the soundtrack and careful framing of the shots to confuse the spectator’s spatial orientation, this

---

340 Michelson, ‘Bodies,’ 194.
movement powerfully simulates the spectator’s own movement through space. Scott Bukatman describes it as a ‘haptic experience’:

The expansion of the visible field to cineramic proportions, the removal of perceptual clues to verticality and other conditions of physical orientation, the sustained evocation of bodily weightlessness, the imposition of the rhythms of respiration and circulation on the soundtrack all contributed to the profound redefinition of haptic experience undergone by the voyagers in the audience.341

For its early audiences, the film represented a novel form of cinematic experience made possible by the effects developed specifically for the film. For the first time, movement was combined with the virtual spaces produced through effects images in a powerful extension of the immersive capabilities of the medium. According to the reviewers in The Harvard Crimson, like Michelson’s essay a notable early exception to the almost unanimous critical disparagement of the film, ‘No film in history achieves the degree of three-dimensional depth maintained consistently in 2001’342 The effects are essential in allowing the construction of a virtual space that extends into the illusory depths beyond the surface of the image, and into which the body of the spectator can be projected. The film’s historical importance, and its significance in the technological development of the medium, is in adding movement to the virtual spaces enabled by the compositing of multiple image elements, and it thus represents a significant extension of the possibilities regarding the simulation of movement. To an extent not previously achieved, 2001 enables movement through a virtual space, a significant step towards the navigable spaces of contemporary digital cinema.

Aesthetic Immersion and Distanciation

2001’s role in developing the technology that would allow the compositing of multiple moving image elements makes the film a milestone in the history of cinematic effects and the development of technologies of visual simulation more generally. The film’s interplay of movement and space makes it a powerful exploration of the aesthetic implications of the spectator’s relationship with moving images, a theme that interested Kubrick throughout his career. His cinema as a whole demonstrates a predilection for camera movement, and a consideration of his directorial style beyond 2001 can demonstrate the wide range of aesthetic responses that can be produced through a moving camera. The innovative use of the zoom in Barry Lyndon (1975), for example, produces an alternative aesthetic response that works against the illusory realism and sense of embodiment described above. At designated moments throughout the film, the zoom functions to undermine the transparent mediation of the diegesis, which in other ways is carefully constructed and communicated to the viewer in a way designed to maximize authenticity in its historical setting. The film’s frequent use of long drawn-out zoom movements provide an external perspective on the drama, encouraging, as Chris Pliatska puts it, ‘a reflective engagement with the film instead of a relatively unthinking immersion in the film.’ In this section I’ll examine Barry Lyndon as an alternative to the immersive ideals that camera movements conventionally uphold. I’ll make a contrast between three distinct responses – immersion, aesthetic immersion, and distanciation – which will add to the phenomenology of immersion that I’m developing.

In an article entitled Bionic Eye, John Belton clarifies the distinction between tracking and zooming. He writes that with a tracking shot, ‘the camera moves boldly through space, producing a two-dimensional image through a three-dimensional filming process which endows that image with an illusion of depth.’ Belton here explicitly links the movement from a tracking shot with the illusion of a virtual space: it is the spectator’s sensation of movement

343 For details on the meticulous effort to achieve realism in the cinematography, see Ed DiGuilio, ‘Two Special Lenses for Barry Lyndon,’ American Cinematographer 57, no. 3 (1976): 276-7.
through space that helps give depth and dimensionality to the diegetic world. In contrast, the zoom produces the illusion of movement through changes in the focal length of the lens, rather than through actual movement of the camera, ‘creating an image which progressively alters the original space being photographed and which subverts the illusion of depth.’ The zoom flattens the image, exposing its inherent artifice and preventing the illusion of a virtual space. The unique effects produced by these different movements have profound implications for the spectator’s immersion in the diegetic world. As Belton summarizes, the viewer ‘feels distanced from or outside of the action shown through a zoom and feels involved with or inside of an action shot with a moving camera.’

The important aesthetic implications of these two movements can be seen in the way Barry Lyndon makes the zoom movement and its distinct effect central to its philosophical concerns. Critical discussion of the film often labels it as ‘distancing’ or ‘alienating,’ adjectives that highlight its divergence from an immersive aesthetic and suggesting a disinterested spectator. This generally unenthusiastic response goes back to the film’s early critical reception. In a 1976 review in The New York Review of Books, for example, Michael Wood describes the film as ‘austere, stately, beautiful, faintly inhuman’ and argues that ‘it just hangs there on the screen for three hours, a monument to Kubrick’s patience and pedantry and rather laborious good taste, but signifying very little else.’ This conventional view of the film as cold and distancing is challenged by Michael Dempsey and many others who emphasize the film’s unique power and emotional depth. In a review in a 1976 edition of Film Quarterly, Dempsey suggests that the early negative critical response to the film was because of its ‘stately pace, the painterly framing, the detached tone, the unpsychoanalyzed characters, and the lack of dramatization.’ But Dempsey highlights the way the film operates in alternative aesthetic register, making use of a variety of formal devices to deliberately inhibit the spectator’s immersed identification with the characters and their struggles.

---

347 Belton, ‘Bionic,’ 22.
One of the film’s strongest visual strategies is the repeated use of slow, reverse zoom movements that begin with a close-up on a particular object or person, before slowly drawing back into an extreme long shot. A typical example occurs in the first half of the film as Barry and Captain Quin are preparing to fight their duel. The scene starts with a close up of the guns being prepared, and over 45 seconds the zoom slowly withdraws until Barry and Captain Quin are small figures enveloped by the surrounding environment (see Figures 17-19). A similar movement could be achieved with a tracking shot, with the camera pulling back to achieve a similar distancing effect from the subject of the image’s focus. However, this would involve an entirely different register of movement. The zoom movement flattens the image, pushing both the foreground and background elements into a single image plane, removing the spectator from the diegesis and exposing the virtual space as an illusion. In contrast, a tracking movement would move the spectator through the space, retaining the distances between the elements in the image, and maintaining the spatial coherence of the diegetic world. Both movements maintain temporal continuity, but the zoom movement heightens the sense of distanciation from the diegesis by problematizing the scene’s spatial integrity.

Alan Spiegel, another of the film’s early reviewers, describes the distancing effect produced by the zoom:

[In the grandest, saddest, most elegiac of motions, the camera slowly pulls back (from near to far), and takes leave of the human struggle; that is to say, the motion of the camera begins in drama and ends in spectacle,
starts off with an action and finishes with a design, converts human value to aesthetic value and a utilitarian image into a self-reflexive image.\(^{350}\)

The zoom’s distinct movement turns a diegetic space into a ‘self-reflexive image,’ problematizing emotional engagement with the characters and their individual struggles. It is a unique formal strategy that has been interpreted in various ways. One interpretation comes from Dempsey, who praises the film as ‘a meditation on the transience of life.’\(^{351}\) The film is not a ‘drama or a character study,’ but a contemplation of life’s ephemerality and a somber deliberation on our place in history.\(^{352}\) Dempsey highlights the reverse zoom as one of the film’s most affective visual strategies, suggesting that it places ‘humanity and its activities in a timeless perspective, highlighting their evanescence.’\(^{353}\) By distancing us, the film discourages an immediate involvement with the characters in favor of reflection and contemplation of them and their place in history. With the long, drawn-out movement from drama to spectacle, we see the characters not as individuals with unique personalities and individual struggles, but as lonely, isolated figures, dwarfed by the sweeping historical forces to which they are subject.

A second interpretation comes from Chris Pliatska, who focuses on the spectator’s distanciation from the diegetic world in terms of the philosophical idea of the absurd. The philosophical form of the absurd differs from the conventional understanding of the term, in that the philosophical sense involves a degree of self-awareness that comes from adopting an external perspective on our lives. Pliatska points out that there is a natural viewpoint that we take as participants in life. ‘From this perspective,’ he writes, ‘human life is permeated with meaning and value. We show intense concern about the kinds of people we are or would like to be, and we pursue myriad projects with varying degrees of interest.’\(^{354}\) However, as humans we have the capacity to step back from our lives and adopt an external perspective on them. When we take on an external perspective in this way, ‘the lives that are ordinarily so permeated with meaning take on an arbitrary quality.’\(^{355}\) Several formal characteristics in *Barry Lyndon*, including the reverse

\(^{350}\) Alan Spiegel, ‘Kubrick’s *Barry Lyndon*,’ *Salmagundi*, no. 38/39 (Summer-Fall 1977): 204.
\(^{351}\) Dempsey, ‘Barry Lyndon,’ 49.
\(^{352}\) Dempsey, ‘Barry Lyndon,’ 49.
\(^{353}\) Dempsey, ‘Barry Lyndon,’ 50.
\(^{354}\) Pliatska, ‘Shape,’ 187.
\(^{355}\) Pliatska, ‘Shape,’ 187.
zoom, encourage such an external perspective. Throughout the film we are generally aligned with Barry’s perspective and thus begin, Pliatska writes, to ‘identify with his struggles and ambitions.’ But when the camera begins to pull away, ‘we have access to a perspective that is unavailable to Barry and thus can comprehend the larger world of which he is mostly unaware.’ Through this we recognize the absurdity of the struggle in which he is engaged. Furthermore, the film’s formal strategy of distanciation also encourages us to adopt a similar perspective on our own lives, ‘and from this perspective, the rituals that we consider significant will also seem trivial and arbitrary.’ The unsettling tension created through this recognition reveals ‘the absurdity of the human condition.’

These two interpretations of the film – as meditation on the transience of life, and as confrontation with the absurd – emerge from the unique dispositions of the individual spectators, but both responses clearly develop from positions of profound emotional engagement with the film. Both responses show levels of deep reflection on the film’s philosophical and aesthetic interests, and can hardly be described as the detached or critical responses of disinterested spectators. They are affective responses that emerge from an emotional engagement not with particular diegetic characters and their drama, but with the more general concerns of the human condition. I would argue therefore that Barry Lyndon is not the cold unfeeling film that a distanced aesthetic suggests, but its various distancing devices – the ironic voice-over narrator, the self-conscious allusions to the period’s artwork, the zoom’s breach of spatial integrity –

356 Pliatska, ‘Shape,’ 190.
357 Pliatska, ‘Shape,’ 191.
358 Pliatska, ‘Shape,’ 192.
359 Pliatska, ‘Shape,’ 192. A third interpretation of the film is suggested by Maria Pramaggiore, who points out that the film provides a critique of military culture. Its focus on ‘the accoutrements surrounding the enterprise of war’ such as ‘looting, desertion, and the awarding of medals for bravery,’ aligns it with ‘the military culture critique of Dr. Strangelove.’ Furthermore, its coming of age trajectory from innocence to experience ‘places it firmly within the sensibility of Paths of Glory and Full Metal Jacket.’ She also suggests that the film emphasises ‘the dialectics surrounding masculine bonding and conflict, including dueling, brawling, intergenerational disputes, and formal battles.’ Pramaggiore does not explicitly refer to the zoom in this interpretation of the film, but its conspicuous detachment from the concerns of the characters encourages the critique that she raises. See Maria Pramaggiore, Making Time in Stanley Kubrick’s Barry Lyndon: Art, History, and Empire (New York and London: Bloomsbury, 2015), 35.
360 A cognitive theorist here might explain these higher-level interpretations in terms of the brain’s neural processes. Grodal’s PECMA flow model of film experience, for example, describes the process by which audio-visual data are received by ‘innate, sealed modules that are totally impervious to cultural influence and learning,’ and are then processed by higher-level associational networks, where ‘culture and personal experiences play a much greater part.’ Grodal, Embodied, 147.
produces a spectator response that nonetheless deviates in important ways from the ideal of immersion that is my primary focus. Specifically, it disrupts the transparent mediation of the diegetic world that is necessary for the spectator’s relocation to the time and place of the narrative event. The film produces a powerful affective response, but generally avoids recentering the spectator in the diegetic world as in characteristic immersive responses.

To clarify the similarities and differences between what I will call the ‘aesthetic immersion’ in *Barry Lyndon* and a more standard or conventional immersion, I’ll examine the use of camera movement in Jean-Luc Godard’s *Weekend* (1967), a film that employs camera movement as a distancing device but for an end entirely different from *Barry Lyndon*’s philosophical and aesthetic project. Godard’s aggressive rejection of emotional engagement is located throughout the film in his deliberate and consistent subversion of classical convention, but the film’s celebrated tracking shot serves as an appropriate counter to the immersive ideals I’m associating with more conventional long take aesthetics. The shot tracks from left to right to follow a bourgeois couple as they struggle with a traffic jam while driving on a road through the French countryside. Godard’s departure from classical convention is clear and sustained. The length of the shot far exceeds the classical norm, and the camera repeatedly loses and finds the characters with whom it is ostensibly concerned, demonstrating an indifference to the characterological function of more conventional Hollywood storytelling. The camera’s extended lateral movement runs parallel to the trajectory of the characters, an effect that flattens the image and avoids acknowledging the depth of space. As Brian Henderson points out in an article titled *Toward a Non-Bourgeois Camera Style*, Godard’s ultimate aim is political. For Godard’s critique to work, the tracking shots must be understood dialectically, as a response to the moral depth and ambiguity inherent in the Wellesian sequence shots I examined above. ‘Godard's tracking shot format,’ Henderson writes, ‘insists on a single perspective and on the sufficiency of a single comprehensive survey for understanding of the transparent, easy-to-understand bourgeois world.’ 362 The sustained lateral movement collapses the world into a ‘cinema of one plane’, which is a ‘demystification’ and ‘an assault on the bourgeois world-view and self-image.’

---

And this is where I see the distinction between these two examples of distanced spectatorship. The zoom in *Barry Lyndon* collapses the frame in a powerfully affecting visual strategy, revealing the tragic reality of life’s ephemerality and the angst and sorrow that come with contemplating the absurdity of existence. In contrast, the lateral tracking shot in *Weekend* activates the critical awareness of the spectator by *disengaging* the emotions. The two films display a similar mode of address, but each makes a fundamentally different aesthetic claim on the spectator. *Barry Lyndon* is therefore not the cold and distant film of critical convention, but a powerfully affecting film that makes use of cinema’s formal qualities to evoke profound aesthetic and philosophical reflection. However, the spectator’s external perspective on its drama clearly distinguishes it from the immersive ideals of classical cinema, so it is not contained within the phenomenology of immersion I’ve so far been describing. Its powerful affect coupled with its distanced address means it lies somewhere between the poles of immersion and distanciation. I’ll therefore describe it as an example of ‘aesthetic immersion,’ a label that indicates an aesthetic response, but indicates a degree of departure from the full relocation to the virtual world found in more characteristic immersive responses. Retaining the word ‘immersion’ indicates the heightened involvement with the aesthetic object and differentiates this kind of response from the staunch rejection of emotional engagement in Godard’s more radical and sustained distanciation.

**The Shining and the Aesthetics of the Steadicam**

Aesthetic immersion is a response that I would argue characterizes Kubrick’s cinema more generally. It can be seen in *The Shining*, a film representing a fundamentally new form of cinematic movement made possible by developments in the early 1970s in the stabilization of hand-held cinematography. These developments were perfected in 1975 in the Steadicam, a device that combines the mobility of a hand-held camera with the stability of a dolly, producing a graceful and fluid image in motion, while simultaneously allowing the free movement of the camera through space. Serena Ferrara points out in *Steadicam: Techniques and Aesthetics* that the extra stability of the device is achieved through three basic principles: shifting the camera’s center of gravity, spreading the camera mass, and isolating the camera from the movements of
The three principles eliminate the unstable movements that usually characterize handheld camerawork, but do not limit movement in the way the dolly or other more stable conventional tracking methods can. This leads to the smooth, free-flowing movements of a device that, as Geuens describes it, ‘responds to the demands of one’s hands with the grace and fluidity of a jazz dancer.’

The device was developed in the early 1970s and was used on a number of films from 1976 onwards. None of these films however made extensive use of its unique visceral qualities, and it was not until the release of The Shining in 1980 that the Steadicam was used as a sustained formal strategy central to a film’s aesthetic. The Steadicam’s fluid movements through the threatening spaces of the Overlook Hotel have been described as ‘one of the most spectacular applications of technological invention to the illusionist seductions of the medium.’ The film’s many lengthy tracking shots through the hotel’s hallways provide a visceral sense of immersion in the threatening space. Kubrick’s specific achievement in the film is the recognition of the new technology’s ability to intensify the spectator’s sense of immersion in the cinematic image. However, as with many of Kubrick’s films, The Shining was met with a combination of confusion and disappointment upon release. Gregg Smith sums up the response: ‘some critics complained the film was too complicated and didn’t make sense, others that it was too slow, still others that it was not scary enough.’ Variety wrote that Kubrick destroys ‘all that was so terrifying about Stephen King’s bestseller,’ and Dave Kehr found that the ‘imagery

---

366 The first Steadicam shot in a feature film is a 2-minute tracking shot in Bound for Glory (1976). Beginning from a raised platform, the camera descends to the ground and then follows David Carradine as he moves through a crowd. The device was also used in Marathon Man (1976) to add dynamism to the shots of Dustin Hoffman running through Central Park and the streets of New York. Rocky (1976) features the celebrated sequence in which the Steadicam follows Rocky as he climbs the stairs to the Philadelphia Museum of Art.
367 One possible exception to this may be John Carpenter’s use of the Panaglide in the famous opening shot of Halloween (1978).
– with its compulsive symmetry and brightness – is too banal to sustain interest.® 371 Given the
film is a long, ponderous story about unremarkable and unlikeable characters, it is hardly
surprising that some of the response was negative. Its release also coincided with a surge in
popularity in ‘slasher’ films, and as a ‘thinking person’s horror film,’® 372 The Shining was
unlikely to appeal to audiences accustomed to films like The Amityville Horror (1979),
Halloween (1978) and Friday the 13th (1980), all of which were released in the three years prior
to The Shining to significant commercial success.

Smith adds that it was not just popular critics that disliked the film, but that academic
critics were disinterested because they saw it as ‘a horror film and as such not worth paying
attention to.’® 373 Smith’s point here is that the contemporary critical dismissal of The Shining as a
genre film was a simplistic reduction of its thematic and stylistic complexity. As with previous
‘genre’ films in his career, the horror genre to Kubrick was hardly a strict system of rules and
conventions to which a film must conform, but rather a framework on which to construct a
unique conceptual vision. Richard Jameson argues that categorizing The Shining as a horror film
is as helpful as describing Dr Strangelove as an ‘anti-war film’, or 2001 as an ‘outer-space pic’,
or Barry Lyndon as a ‘costume picture.’® 374 He argues that ‘The Shining is a horror movie only in
the sense that all Kubrick’s mature work has been horror movies – films that constitute a
Swiftian vision of inscrutable cosmic order.’® 375 Contemporary audiences expecting a
conventional horror film similar to the commercially successful slashers of the late 1970s were
thus unlikely to be impressed by a film that not only dispensed with the conventions of the horror
genre, but one that strategically deviated from the classical Hollywood style to conform more
closely to what Bordwell would describe as an ‘art’ film.® 376

371 Dave Kehr, ‘The Shining,’ The Chicago Reader, accessed November 14, 2017,
https://www.chicagoreader.com/chicago/the-shining/Film?oid=2686751.
372 Amy Nolan, ‘Seeing is Digesting: Labyrinths of Historical Ruin in Stanley Kubrick's The Shining,’ Cultural
373 Smith, Horrorshow, 300.
374 Richard Jameson, ‘Kubrick’s Shining,’ Film Comment 16, no. 4 (July/August 1980): 29.
376 David Bordwell, ‘The Art Cinema as a Mode of Film Practice,’ in Film Theory and Criticism, 6th edn., ed. Leo
Smith and Jameson both suggest that audiences judging the film according to aesthetic criteria based on classical convention were unlikely to appreciate The Shining’s narrative irresolution, ambivalent point-of-view, unlikeable characters and general ambiguity. Furthermore, the film’s overt self-consciousness was likely to further alienate audiences accustomed to a style that effaces all traces of the author – a ‘rule’ ignored by Kubrick throughout his career. Jameson makes the point that it is more helpful to categorize The Shining as ‘A Stanley Kubrick Film’ than as an example of the horror genre.\(^{377}\) Mamber puts this another way: ‘Behind all the hotel doors setting ghastly images in motion, dripping blood out of elevators, providing the unexplained means of escape to frequently trapped characters, lies the director himself, a parody puppeteer in the shadows.’\(^{378}\) Kubrick’s frequent inclusion in lists alongside the great auteurs of cinema – Hitchcock, Welles, or Antonioni, for example – is in part due to this authorial presence, or directorial ‘signature’. By rejecting the effacement of the author typical of the classical style, Kubrick’s cinema demonstrates a modernist sensibility that frequently pushes beyond the conventions of classical style.\(^{379}\)

One example of where this modernist sensibility is shown is in the abundant visuality provided by a fully mobile camera. Like the zoom in Barry Lyndon, the Steadicam in The Shining acquires a self-consciousness that exceeds purely descriptive camera movement. An early example of this can be seen in the first post-credit sequence of the film, which depicts Jack entering the doors of the Overlook Hotel’s spacious and well-lit lobby (see Figures 20-25). He enters the lobby left of frame and the camera tracks laterally to follow him as he approaches a reception desk that soon enters the frame from the right. He has a short conversation with a receptionist who gestures to Ullman’s office off-screen and behind the camera. The camera then follows Jack as he moves towards the office, first panning anti-clockwise to give a panoramic view of the expansive set. The camera then follows Jack through one doorway, pauses as he hesitates at a second, and follows him into the office as he is invited in by Ullman. The camera

\(^{377}\) Jameson, ‘Shining,’ 29.


\(^{379}\) I use the term in accordance with Bordwell’s understanding of a modernist cinema: ‘that set of formal properties and viewing protocols that presents, above all, the radical split of narrative structure from cinematic style, so that the film constantly strains between the coherence of the fiction and the perceptual disjunctions of cinematic representation.’ Bordwell, ‘Art Cinema,’ 780-781.
finally becomes stationary inside the office to observe a brief conversation between the two characters.

A relatively simple classical arrangement would suffice in conveying the sequence’s straightforward narrative content. A long shot would establish Jack’s entrance to the lobby; a cut to a mid-shot would show his brief conversation with the receptionist; a cut to the inside of Ullman’s office would show Jack entering the room moments later. This causal montage would effectively establish the spatial contours of the environment and direct the gaze of the spectator toward its key significatory elements. Instead, Kubrick films the sequence as a continuous shot, showing the entirety of Jack’s movement from the front doors of the hotel lobby to his arrival in Ullman’s office nearly a minute later. Kubrick’s decision to film the sequence without a cut recalls Bazin’s analysis of camera movement in *Grand Illusion* described earlier. It could be said that Kubrick’s decision to film the sequence as a continuous take functions to portray ‘realistically the relations between men and women and the world in which they find themselves,’ as Bazin described Renoir’s use of camera movement. But this sequence has none of the subtlety of Renoir’s movement. Kubrick instead revels in the exuberant visuality provided by a fully mobile apparatus. In blocking Jack’s movement as he moves through the lobby
towards Ullman’s office, Kubrick ensures he walks around the camera, allowing it to perform a wide circle to follow him, orientating the spectator to the vastness of the set and establishing the camera’s ability to go anywhere and see everything. Furthermore, the Steadicam’s eerie floating sensation betrays an autonomy unachievable with conventional hand-held cinematography. Not only does visuality here thoroughly exceed the demands of the scene, it ‘superimposes over the conventional action a panoptic demonstration of pure visual presence.’380 The camera here is not just describing a setting; it is announcing a vast and precarious cinematic space of which it is unmistakably the center.

The Steadicam and Diegetic Presence

The incongruity of a ‘difficult’ modernist sensibility with a conventionally marketed genre film goes some way toward explaining The Shining’s early mixed reception, but as with other Kubrick films it has risen in both popular and critical esteem in the decades subsequent to its release. It has become the subject of numerous critical studies and works of scholarship, with critics focusing on the film's narrative elements – plot, story, and characterization – in discussing its thematic concerns. These analyses have interpreted the film as a complex exploration of American history, patriarchal repression, and the American nuclear family,381 which connect to broader ideas that recur throughout Kubrick’s cinema, such as humanity’s preoccupation with war, the maintenance of male-dominated systems of power, and the incompatibility of social institutions with human nature. High scores on both Rotten Tomatoes and IMDB, cinema’s more informal indices for gauging popular success, show continued appreciation of the film beyond just a scholarly context.

The belated embrace of the film and its continued resonance shows that it has an effect that transcends its immediate historical context. In Stanley Kubrick Directs, Alexander Walker writes that a reason Kubrick’s cinema has endured is that he is one of few filmmakers with a

380 Geuens, ‘Visuality,’ 15.
unique conceptual talent, which enables him to ‘crystallize every film [he] make[s] into a cinematic concept.’ This talent involves the use of cinematic form to ‘exhibit the maker’s vision in an unexpected way, often a way that seems to have been the only possible one when the film is finally finished.’ In The Shining, the form that ‘exhibits the maker’s vision’ is the Steadicam’s dematerialization, its liberation from the material density of conventional camera movements that so effectively simulates the spectator’s presence in the hotel. In this respect, the film makes use of two strategies of immersion, both of which result from the unique use of the Steadicam. The first strategy is the consistent use of movement following characters throughout the hotel, which aids in the spectator’s construction of a 3D space. As the camera moves it activates several depth cues that are essential to our real-world perception of the depth and dimensionality of space. The second strategy used to simulate our presence in the hotel is the unique quality of movement of the Steadicam, which neurophysiological studies suggest may be the form of image capture most able to simulate real-world human vision. These two strategies work together to produce the uncanny immersion that pervades the film.

The first strategy relies on the movement enabled by the Steadicam to establish the depth that is so important to our perception of a virtual 3D space. Garret Brown, the Steadicam’s inventor, describes the new qualities the technology brought to the cinema:

> When the camera begins to move, we are suddenly given the missing information as to shape and layout and size. The two-dimensional image acquires the illusion of three-dimensionality and we are carried across the divide of the screen, deeper and deeper into a world that is not contiguous to our own.

The Steadicam is of course not the first device capable of producing a moving point-of-view, but it represents a radical increase in the freedom of movement. Kubrick exploits this freedom throughout the film, making use of movement on both the x-axis (movement horizontal to the spectator) and on the z-axis (movement into and out of the space) to enhance the illusion

---

of three-dimensionality. An example of x-axis movement occurs early in the film as Jack and Wendy are being introduced to the hotel by Ullman. Here the camera tracks laterally to follow the movement of the characters as they move from right to left in the midground of the shot. As they move through the space, a pillar momentarily blocks our view of them (see Figures 26-28).

Stephen Prince points out that in both real-world and cinematic perception ‘we perceive spatial layout according to numerous depth cues that act in concert to provide redundant information about the positioning of objects in a volume of space.’ \[385\] This camera movement from *The Shining* activates at least two depth cues – occlusion and motion perspective – both of which help construct the illusion of a virtual space. According to James Cutting, occlusion occurs when an object in a scene is positioned in front of another object, obscuring part of the object behind it. Occlusion gives some indication of the relative position of objects in depth, though the amount of information is limited. ‘Camera position and the layout of clutter in a scene,’ Cutting writes, ‘will dictate to the observer (and camera) which objects occlude or partly occlude others.’ \[386\] In this scene, the pillar momentarily occludes the characters in the midground of the shot, who in turn occlude objects in the more distant background, providing a series of depth cues, the cumulative effect of which is a clear sense of the depth of the space. Cutting indicates that occlusion alone cannot give detailed information as to the exact position of objects in a scene, but it can indicate the relative positions of objects in depth. Depth cues usually work in conjunction, redundantly providing information regarding depth. Motion perspective, a second depth cue relevant here, involves differences in the apparent speed of movement of different

---

385 Prince, *Digital*, 199.
objects depending on their distance from the observer. Cutting writes that most effective in producing motion perspective in cinema are dolly or tracking shots ‘where near objects and textures move faster than far ones, and their velocity is inversely proportional to their distance from the camera.’ The motion perspective in this sequence causes the pillar in the foreground to move more quickly across the frame than the characters in the midground, who in turn move more quickly than the windows in the distant background. This gives further information regarding the depth of the space, adding to the information provided by the occlusion caused by the pillar, and further enhancing the illusion of a virtual space.

In addition to occlusion and motion perspective, Cutting describes a third depth cue that contributes to our sense of space. Height in the visual field, an example of which can be seen in figure 19, involves the relative positions of different objects in a scene with regard to their height from the spectator’s point-of-view. Cutting points out that in cinema, height in the visual field is determined by both the position of the camera and its angle in relation to the ground plane. He uses varying camera height and angle throughout The Sound of Music to demonstrate the role that height in the visual field can have in determining the viewer’s sense of space. In the first half of the film, when ‘viewers are supposed to identify with the children’, the camera is generally placed at a level slightly less than the height of the adults. The second half of the film, which is more focused on the developing romance between Maria and Captain Von Trapp, ‘is shot mostly from the eye height of an adult.’ The changing height of objects in the frame that results from these variations in camera placement is crucial in developing the viewer’s subjective sense of the cinematic space. Cutting writes that ‘the relations among objects, particularly as revealed by height in the picture plane, tell us where our eye is – and thus help tell us whether we, the film audience, are “children” or “adults”.’ Kubrick uses height in the visual field in The Shining to align the spectator with Danny’s subjective view of the threatening spaces of the Overlook Hotel. In Figure 29, the low camera positions us close to the ground with an angle of vision turned upwards towards the vast spaces of the hotel that tower above. This follows a convention of cinematic language whereby the position of the camera and its angle determines

387 Cutting, ‘Perceiving,’ 15.
388 Cutting, ‘Perceiving,’ 11.
389 Cutting, ‘Perceiving,’ 11.
390 Cutting, ‘Perceiving,’ 11.
the power relations between the subjects in the scene. As Cutting writes, ‘The height of the camera and its angle, in turn, place the perceiver in a subjective position - high often indicating dominance ... and low a more submissive role.’

These depth cues offer some insight into the way depth and space are constructed in the cinema, and the way that camera movement can help produce the sense of depth that is essential to immersion in a virtual space. However, *The Shining* has a distinct form of movement that affects the reader beyond just mapping out the space, an effect that would be achievable with more conventional methods of moving the camera. The film also makes use of the Steadicam’s distinct quality of movement to add to the simulation of presence. Empirical research into cinematic spectatorship has begun to illuminate the exact processes through which different aspects of film style, including camera movements, affect an audience. Some studies show correlations between neurophysiological activity in the brain of a spectator and subjective experiences. For example, a study led by Katrin Heiman at the University of Parma examined the neurophysiology of spectators as they viewed a scene filmed with different camera movements. The authors were able to draw some general conclusions about the effects of camera movement based on comparisons between the results of the study and the subjective accounts of the

---

391 Cutting, ‘Perceiving,’ 11.
spectators’ experiences of the scene. In the experiment, a scene showing an actor performing a ‘goal-directed hand action’ was filmed in four different ways: from a fixed position, zooming in on the scene, moving towards it with a tracking shot, and moving towards it with a Steadicam. The participants viewed each of the sequences while their motor cortex activity was recorded through electroencephalography, a standard method for measuring neurophysiological responses to stimuli. The participants were then asked a series of questions about their experience of each of the different movements, including ‘How much did you feel involved in the scene?’ and ‘How much did you feel as if you yourself would approach the scene?’ After examining the motor-cortical activity that was recorded during each of the movements, and comparing the results with a baseline level of activity, the authors concluded that the Steadicam, when compared with the other movements, produces ‘a visual experience closest to natural human vision during movement.’

The study was centered around the idea that the mirror mechanism, the structure in the brain that includes the motor cortex, is activated when a subject both performs and perceives someone else performing an action. Vittorio Gallese and Michele Guerra point out that the cortical motor system was traditionally understood as the controller of basic physical aspects of movement, but this understanding was revolutionized by the discovery of mirror neurons that activate not only when performing movements and actions, ‘but also during their perception when executed by others.’ This principle has been used to explain a wide range of cinematic phenomena, from the empathic identification we experience with fictional characters, to cinema’s powerful sense of embodiment. The authors of this study suggest that it is also crucial for our experience of camera movement. The resonance between the neural activity recorded during cinematic camera movements and during actual movement suggests that on some level spectators experience cinematic phenomena vicariously. Gallese and Guerra write that this shared motor activation is largely the cause of ‘the impression of “being there,”’ and exploring

---

394 Gallese and Guerra, ‘Feeling,’ 108.
the film space and measuring its time. It provides, they argue, ‘empirical ground to the notion of the capacity of the camera to simulate the virtual presence of the viewer inside the movie.’

The application of neurophysiological studies to cinematic spectatorship has a short history, and limitations are immediately apparent. For example, recognizing the neural correlates of certain aspects of film style confirms that camera movement has an effect on the spectator that can be measured at the neurological level, but does little to explain the subjective experience of the movement. There is also a large degree of variation in camera movement in the cinema, far beyond the limited scenario tested during the study explained above. Studies such as these may be of limited use in explaining the function of camera movement more generally in cinema. However, these studies should also not be entirely dismissed. The authors’ conclusion that the Steadicam produces a sensation of movement that resembles natural vision corresponds with more anecdotal accounts of the Steadicam’s unique ability to simulate human vision. It adds further evidence to the claim that the device is able to simulate real-world vision more closely than other forms of moving image capture. Further research of this type could produce a greater understanding of different camera movements, and of film style more generally, and help arrive at a more comprehensive explanation for the subjective experience of cinema.

In The Shining, the Steadicam’s unique ability to simulate natural human vision works in combination with its freedom of movement to produce a powerful simulation of the spectator’s presence in the haunted spaces of the Overlook hotel. I’ve already suggested that Kubrick’s use of the Steadicam’s exuberant mobility represents a radical departure from conventional spatial aesthetics, but the significance of this stylistic characteristic can be more clearly understood through its contrast with a more conventional style of representing space. The ‘tableau’ style of

---

395 Gallese and Guerra, ‘Feeling,’ 106.
396 Gallese and Guerra, ‘Feeling,’ 103.
397 Murray Smith points out that just confirming that neural activity occurs during an aesthetic experience does little to illuminate our understanding of it. He argues that studies that observe and record neural activity must be accompanied by phenomenological and psychological accounts from the subject in whom the experience takes place, a ‘triangulation’ that can do much to shed light on aesthetic phenomena. In this way, we can start to provide answers to questions about consciousness and subjective experience, domains that often elude the methods of the hard sciences. Murray Smith, ‘Triangulating Aesthetic Experience,’ in Aesthetic Science: Connecting Minds, Brains, and Experience, ed. Arthur P. Shimamura and Stephen E. Palmer (New York: Oxford University Press, 2011), 80-102.
filmmaking, which Bordwell identifies as a predominant style of the silent era, maintains a strict demarcation between the spectator and the diegetic space. According to Bordwell, ‘the camera is set fairly far back from the action,’ and there is very little cutting, allowing the actors to ‘play out the drama in prolonged shots.’ Storytelling in this mode consists of carefully directing the spectator’s attention towards significant parts of the frame at relevant moments during the unfolding scene. The spectator is rarely brought over the threshold into the diegetic space through a cut or a camera movement. Camera movement generally, but the Steadicam’s abundant movement in particular, problematizes this strict demarcation between diegetic and extra-diegetic space. Z-axis movement is the deliberate attempt to reconfigure this conventional relationship between the diegesis and its observer, which places the diegesis on one side of the frame, and the spectator firmly on the other. By moving forward or backward in space the spectator is invited to probe the cinematic world presented in the diegesis.

Kubrick’s consistent attempt to breach the threshold separating the diegesis from the spectator is shown through a trope that recurs throughout his cinema: a character walks towards the spectator through a symmetrical tunnel as the camera tracks backwards. Kubrick often combines this movement with a wide-angle lens, which distorts the frame by emphasising objects in the foreground and pushing the background into the distance. The combination of an exaggerated sense of depth produced by the wide-angle lens and the tracking movement of the camera gives the illusion that the sides of the tunnel are accelerating from the edges of the frame towards a central distant vanishing point. An early example of this in Kubrick’s cinema is the sequence showing General Mireau marching through the trenches inspecting his troops in Paths of Glory (1957). The camera tracks backwards to anticipate the General’s movement, pausing occasionally as he talks to the soldiers, whom he has decided will soon be going over the top of the trench to meet German machine-gun fire. As the General marches past, the wounded and demoralized soldiers lining the sides of the trench accelerate from the edges of the frame out of

---

399 The exceptions to this rule, Bordwell points out, included close-ups of things that were too small to be seen in the overall shot. Bordwell, ‘Viewer’s,’ 31.
400 For more on the spatial distortion caused by different lenses, see Brinkman, Art, 42-45.
shot towards the distant vanishing point. The camera movement thus captures the General’s indifference to the suffering of the men under his control.

Kubrick’s innovation in *The Shining* is to reverse the direction of the tracking. The film contains frequent shots of Danny riding through the hotel, positioning the spectator behind him, inches above the ground. As Danny rides forward we follow him toward the distant vanishing point in the center of the frame. The vanishing point is what Jameson labels a ‘tear in the membrane of reality’ in his description of one of the film’s visual motifs:

> Virtually every shot … is built around a central hole, a vacancy, a tear in the membrane of reality: a door that would lead us down another hallway, a panel of bright color that somehow seems more permeable than the surrounding dark tones, an infinite white glow behind a central closeup face, a mirror, a TV screen … a photograph. 401

The use of the wide-angle lens in these lengthy forward-tracking shots makes the walls of the tunnel accelerate towards us, producing the sensation that we are being pulled into the frame towards the distant vanishing point. These tears in the ‘membrane of reality’ thus contribute to the hotel’s threatening sense. Paradoxically, the extra stability provided by the Steadicam accentuates this sensation, as it is free from the materiality and groundedness that permeate a dolly shot. Like the zoom movement in Barry Lyndon, the distinct movement of the Steadicam in *The Shining* is central to the film’s aesthetic. The combination of the low camera angle, the relentless forward tracking movement through the distorted tunnel, and the Steadicam’s dematerialization - its liberation from the material density of the setting - intensifies and makes tangible the hotel’s threat to subsume the spectator. 402

401 Jameson, ‘Shining,’ 30.

402 Bazin comments on a similar use of wide-angle lenses and low camera angles in *Citizen Kane*: ‘The stretching of the image in depth, combined with the nearly constant use of low angles, produces throughout the film an impression of tension and conflict, as if the image might be torn apart.’ Bazin, *Welles*, 74.
Moving Through Virtual Spaces

So far I’ve dealt exclusively with camera movement in the pre-digital era, but how do the aesthetics of camera movement as described above fit into the history of virtual spaces outlined in chapter 1? The aesthetic modes as well as the production practices that govern the production and reception of the virtual spaces of contemporary digital cinema have roots in much earlier forms and practices. The challenges of smoothly integrating multiple image elements into a single composite are magnified with the addition of movement to the individual shots. To add movement to an effects shot, each of the individual elements to be combined needs to be filmed with exactly the same camera movement prior to their integration. Moving a spectator through a virtual space has therefore been a consistent technical challenge for effects artists and filmmakers. The technology that allows for exactly repeatable camera movements is usually seen as starting with motion control photography, first appearing as the ‘Dykstraflex’ system designed by effects artist John Dykstra for Lucas’ Star Wars (1977).403 The Dykstraflex system uses computerized mechanisms to produce exactly repeatable camera movements, allowing each of the individual elements to be filmed individually, but with the same camera movement applied to each of them. Nora Lee, in an article on motion control in American cinematographer, defines the process as ‘the reliance on some mechanical, electrical or computer-assisted contraption to repeat reliably a camera movement or series of camera movements for the purpose of achieving various special effects.’404 Repeatable movements are necessary in order to achieve matching camera moves on each of the individual elements, which can then be combined into a single, unified composite. The difficulties associated with synchronizing the different moving elements meant that prior to the consolidation of the technology in the 1970s, effects sequences were almost invariably shot with a locked-off camera.405

403 However, electronic control of movement was available as early as the 1940s with the Dupy Duplicator, a device that allowed repeatable camera movements by adapting technology used to synchronize sound recorders with motion picture cameras. It was used in Samson and Delilah (1949) and An American in Paris (1951). See Jeffrey A. Okun and Susan Zwerman, eds., The VES Handbook of Visual Effects: Industry Standard VFX Practices and Procedures (Focal Press, 2010), 8.
404 Nora Lee, ‘Motion Control,’ American Cinematographer 64, no. 5, (May 1983): 60.
405 Brinkman defines a locked-off camera as a ‘camera whose position and lens settings do not change over the duration of the shot’. Brinkman, Art, 653.
2001 is again a landmark in the development of this technology, but as Lee points out, the technology ‘did not explode on the scene in 1968 as most would believe,’ nor did it suddenly emerge during the pre-production of Star Wars, another misconception regarding the development of the technology. Rather, it is part of a long history that stretches back as far as the elaborate clocks of the 15th century, which ‘mechanically repeated a complex series of motions with amazing accuracy - over and over - for years.’\textsuperscript{406} Although there was some development in repeatable movements during the early silent era, with James Brautigan, one of Thomas Edison’s cinematographers, working on aligning multiple double-exposures involving movement, the major technical developments would be made during an intense period of experimentation in the late 1940s.\textsuperscript{407} Various pieces of machinery were added to the basic setup during this time, allowing for tighter control over the movements of the camera, a series of developments that would eventually lead to the application of the technology on a much wider scale in 2001. In contrast to previous uses of these systems, such as in Samson and Delilah (1950), which saved the technique for spectacular and climactic moments, in 2001 effects images involving the combination of multiple moving elements were central to the film’s overall aesthetic.\textsuperscript{408} According to Turnock, the effects team of 2001 ‘rethought the overall purpose and use of special effects in films, by integrating them into an overall programmatic aesthetic approach, interspersing and blending the effects throughout with the live action.’\textsuperscript{409} The film had over 200 effects shots out of a total of 602, a remarkable percentage for the time, showing the unprecedented degree to which effects were interspersed with live-action photography.\textsuperscript{410} The movement through space that permeates the film and forms a central part of its aesthetic meant that the technology that would allow control of multiple moving image elements was an essential part of the film’s production.

\textsuperscript{406} Lee, ‘Motion,’ 60.
\textsuperscript{407} Lee, ‘Motion,’ 60.
\textsuperscript{408} Lee, ‘Motion,’ 61.
\textsuperscript{409} Turnock, Plastic, 70.
Herb Lightman, in the June 1968 edition of *American Cinematographer*, describes the elaborate setup that was constructed during *2001*’s pre-production to achieve the replicable camera movements that were essential to the compositing process:

For this purpose a camera animating device was constructed with a heavy worm-gear twenty feet in length. The large size of this worm-gear enabled the camera mount of the device to be moved with precise accuracy. A motorized head permitted tilting and panning in all directions. All of these functions were tied together with selsyn motors so that *moves could be repeated as often as necessary in perfect registration.*

This kind of precision was necessary to avoid the artifacts that would interfere with the transparent mediation of the virtual space and prevent the audience from accepting unproblematically the film’s vision of the future of space travel. The perfect registration of the individual elements helped ensure the integrity of the images of the diegetic world, into which the audience could become immersed. But rather than completely rejecting the kind of spectacle that so often accompanies technological developments, the film indulges in a series of images that encourage a focus on the novelty of the technology used in their production. The film thus has a complicated mix of spectatorial addresses, ranging from the ‘presentational’ mode of address of the ‘cinema of attractions’, to an immersed involvement in the film’s diegesis, as during the main narrative episode centering on the conflict between Hal and the astronauts.

For example, the film’s first images of space - the five and a half-minute ‘Blue Danube waltz’ sequence extending from the famous jump cut up until the first interior of the Space Station V - eschews the illusory aesthetic of a classical narrative mode and makes a direct claim on the audience, directing attention towards the surface of the images, rather than into their illusory depths. There is no attempt to engage the audience in a developing narrative, or to promote empathy with a character. Rather, the sequence appeals directly to the audience’s ability to appreciate the most visually rich images of space cinema had yet been capable of producing. The slow ponderous movements and a high average shot length for the sequence encourage the

---

412 Gunning, ‘Attractions.’
audience to consider and scrutinize the images carefully, turning them into objects of contemplation, rather than utilitarian images serving to convey narrative content. This is a departure from the style of previous effects-heavy films, where the cutting often speeds up during effects sequences, often an attempt to hide the artifacts that result from the technological limitations in developing the heavily composited effects images. But in this sequence from *2001*, the images are held on screen, with attention directed to their novelty and technological brilliance, a combination of music, image and spectacle in a celebration of filmmaking and the technology of cinema.

While the effects processes developed for *2001* enabled its majestic images of humanity’s future in space, the effects technology of the time was insufficient for the intricate coordination of spaceships, lasers, and exotic settings that Lucas envisioned for *Star Wars*. To achieve the kineticism and visual dynamism he sought for his film, Lucas needed a more precise means of achieving the exactly repeatable movements that was essential in combining multiple moving elements. According to Turnock, effects artists in the years immediately prior to 1977, including Douglas Trumbull, who had designed the ‘stargate’ sequence in *2001*, and those associated with *Industrial Light and Magic*, including Dykstra, Richard Edlund, and Dennis Muren, revived the optical printing methods that had been developed in the 1920s and 1930s, but which had fallen out of favor soon after. The revival of these techniques led to a shift away from principal photography as the part of the filmmaking process most important to a film’s visual aesthetic, to post-production as the most important determinant of the final look of a film, a displacement that has only increased in the digital era. The revival of optical printing techniques, in addition to the desire for a more kinetic visual style, made the control of the multiple moving image elements that made up the heavily layered effects images an essential part of principal photography. In response to the demands of Lucas and other filmmakers at the time, effects artists developed a series of technologies that increased the control and precision in capturing multiple elements independently but with the same motion applied to each of them. When printed together, these

---

413 The entire sequence consists of twenty shots, giving an average shot length of around 17 seconds, well above average for a film of this era. As a rough comparison, Barry Salt found an ASL of 9.3 seconds in a statistical analysis of a random sample of films released in 1959. See Salt, *Film Style*, 279-284.

moving elements produced the illusion of a unified, pro-filmic space that Lucas and others sought.\textsuperscript{415}

As with most of the technologies examined, the motion control technology developed for \textit{Star Wars} represents a development of previous effects technologies and practices, rather than a break from them. The practices that led to the increased control over the photography of moving elements on \textit{Star Wars} can therefore best be understood as augmentations of the practices used by the effects team of \textit{2001}. While the ‘worm-gear’ of the earlier film allowed the painstaking photography of multiple moving image elements to produce its slow-moving and majestic images, the motion control technology developed for \textit{Star Wars} represents an intensification of this process, allowing, as Turnock suggests, ‘more moving elements, along more directional axes, making more intensive layering of elements and motion effects possible.’\textsuperscript{416} While representing developments on, rather than breaks from, the practices of previous eras, the developments introduced for \textit{Star Wars} were sufficiently different from previous iterations of the technology for Turnock to claim that it is ‘difficult to underestimate the importance of the introduction of motion control camera technology in the 1970s within the history of special effects in American cinema.’\textsuperscript{417} As I’ve already suggested, the various developments that occurred during the 1960s and 1970s were important steps in the development of the virtual spaces of contemporary cinema.

\textbf{Digital Movement}

The increasing control over the creation of effects images, including those that simulate camera movement, has made CGI an increasingly affordable and attractive filmmaking production method, rendering the traditional notion of a pro-filmic space an outdated conception of cinematic representation. Digital techniques for simulating camera movement through and

\textsuperscript{415} Turnock notes that motion control is usually associated with the “Dykstraflex” system developed for Star Wars, but that the independent effects company Robert Abel and Associates developed a similar system that was used on their “Bubbles” 7-Up television ad of 1975, and furthermore that Douglas Trumbull was ‘almost simultaneously’ developing the technology for use on Spielberg’s \textit{Close Encounters}, released the same year as \textit{Star Wars}. See Turnock, ‘Auteur,’ 120.

\textsuperscript{416} Turnock, ‘Auteur,’ 120.

\textsuperscript{417} Turnock, ‘Auteur,’ 119.
around photorealistic virtual spaces are now a common occurrence in contemporary modes of visualization, both in the cinema and outside of it. This recent development has its roots in the drastic changes in production methods and technology that occurred throughout the 1960s and 1970s, with the production and popular reception of films such as *2001* and *Star Wars*. These significant developments in motion control technology freed the camera from the restrictions previously needed for effects shots, which has led to the dynamic visual style of contemporary Hollywood cinema. As Turnock writes, ‘Motion control, by combining traditional photographic techniques and aesthetics and the animated frame-by-frame manipulation of the picture plane, is one of the primary sites for understanding how the aesthetics of moviemaking changed in the 1970s.’\(^{418}\) She argues furthermore that ‘by detaching the camera motion from live action, it furthered the notion of what we call now “virtual” camera movement and virtual mise-en-scène.’\(^{419}\) The camera’s increased freedom from the material constraints that governed pre-digital cinema has led to movement - now simulated by a virtual camera - becoming more common and more pronounced, at times challenging, but also frequently extending the classically-mandated continuity central to classical forms of cinematic storytelling.

The increasing freedom from material restraint that characterizes simulated movements challenges the already tenuous relationship between a cinematic image and the pro-filmic reality it ostensibly represents, and is the source of much of the critical skepticism regarding the use of effects in contemporary cinema. The long-take that opened *Gravity*, for example, though largely celebrated by audiences and critics as shown in the opening paragraphs of this chapter, was not without its detractors. Brad Stevens, writing for the British Film Institute’s *Sight & Sound* website, claims that digitally augmented long-takes such as those found in *Gravity* lack materiality and are therefore fundamentally different from their non-digital equivalents, with important consequences for the way they are received. ‘I found myself distinctly unimpressed by the ‘bravura’ opening shots of Alfonso Cuarón’s *Gravity* (2013),’ he writes, ‘since I knew perfectly well that [it was] only playing without cuts thanks to CGI.’ He contrasts the opening of *Gravity* with those of Orson Welles’ *Touch of Evil* (1958) and Alfred E. Green’s *Union Depot* (1932), which, he argues, ‘have a monumental quality because we are intensely aware of the

---

\(^{418}\) Turnock, *Plastic*, 135.

\(^{419}\) Turnock, *Plastic*, 132.
effort that went into filming them, aware that what we are seeing involved real actors moving through real spaces in real time.\textsuperscript{420} The simulated long-takes of contemporary cinema, according to this view, represent a significant departure from a Bazinian realism based on an objective camera faithfully recording events as they unfold in real-time. However, though virtual movement through virtual spaces does offer a radical challenge to traditional notions of cinema’s ontology, the phenomenological implications of these developments are less clear. In the remainder of this chapter, I’ll suggest that digital technology offers not a threat to the continuity of classical aesthetics, but rather an extension of it. I’ll show not only that a classically-inflected conception of space, one grounded in spatial and temporal continuity, remains common in contemporary effects-heavy cinema, but also that digital technology represents an intensification of classical continuity. In many cases, digital technology and effects are used to enable and extend continuity, rather than displace it.

The cinema of Cuarón is a test case for an examination of these ideas, as he is a filmmaker who embraces the radical imaging possibilities represented by digital technology, as well as demonstrating the realist aspirations more readily associated with a pre-digital cinema. \textit{Children of Men}, for example, involves extensive use of the imaging technologies involved in the simulation of camera movement, using them in the service of a naturalistic aesthetic, in contrast to the more overt uses of effects associated with blockbuster filmmaking. In an interview with \textit{Cinefex} magazine, Frazer Churchill, the film’s visual effects supervisor, points out that Cuarón and director of photography Emmanuel Lubezki were aiming for the style of \textit{The Battle of Algiers} (1966), ‘which they loved for its naturalism and documentary feel.’\textsuperscript{421} However, this sort of naturalism is difficult to achieve in shots permeated with effects, as would be needed to construct the futuristic dystopian England that is the film’s setting. The wide-ranging camera movements Cuarón sought added a further challenge to the effects artists. The digital effects team used a range of techniques to allow the compositing of moving footage, shot during principal photography, with the computer-generated imagery necessary for the construction of the diegetic world.


In the *Art and Science of Digital Compositing* Ron Brinkmann points out that the most fundamental operation in the simulation of movement is ‘tracking,’ a process that involves ‘selecting a particular region of an image and determining that region’s movement over time.’

Motion control solved the problem of applying the same movement to multiple elements, all of which could then be combined into a seamless whole. However, as Brinkmann suggests, motion control ‘doesn’t do you any good if you want to synchronize the movement of something you’re adding to a scene with the movement of some object in the scene that already has its own movement.’

In order to add an element, such as a computer-generated image, to a scene with movement, the movement must be derived from the original footage, and then applied to the element that is to be added. This operation, which Brinkmann calls ‘camera tracking,’ effectively produces ‘an exact match, in a virtual 3D space, of the original move that the camera used to shoot the scene went through.’ The opening scene of *Children of Men* demonstrates how this form of tracking can be used to construct a larger diegetic world. This crucial scene, which serves as the audience’s introduction to the world of the film, was shot with a handheld camera that follows Clive Owen’s character as he exits a cafe and enters a futuristic rendering of London’s Fleet Street. To create the futuristic world of the film, animated commercials were created and then added to the sides of buildings, an effect that was achieved by tracking the camera’s movement during the shot, and then applying the tracking data to the animations, which locked them in place on the sides of the buildings (see Figure 30).

---

This kind of camera tracking was possible within a traditional optical compositing system through a painstaking process of manually hand-tracking an object frame-by-frame. Digital compositing methods make this form of image manipulation much easier to accomplish. However, in addition to making the process of tracking easier, and providing increased control over the final image, digital effects offer more radical developments in the simulation of movement, extending the imaging possibilities of cinema and allowing a more immersive aesthetic. The opening scene of *Children of Men* was shot over two days, with the interior of the cafe shot on day 1, and the exterior involving the explosion shot on day 2. Frazer Churchill points out that there were around 300 extras, 80 vehicles, and 20 special effects cues in the shot, and the difficulty in choreographing the sequence led to the division into 2 days. The 2-day shoot also allowed the special effects crew to reset the cafe interior with explosives and rubble, and to prepare the extras injured in the explosion with appropriate make-up and effects. However, this division into two days presented a problem for Cuarón, who wanted the scene to be experienced by the audience as a continuous take. To achieve the phenomenological continuity Cuarón sought, the visual effects team included a seamless digital transition to ensure continuity between the two takes.

Churchill describes the process in detail:

Alfonso didn’t want the “waiter with the tray of drinks” approach, where a random passerby wiped frame at the cut… Instead we decided to make the blend where the camera momentarily panned off Clive, becoming his point of view as he stepped outside the cafe. In the moment we panned off him, we did a digital transition around the cafe doorframe… As Theo stepped outside, we were seeing the pavement and people from Day 2

---

426 Fordham, ‘Children,’ 34.
427 For a more detailed description of the methods used to ensure seamlessness, see Fordham, ‘Children,’ 34.
composited with the bus from Day 1. As the bus passed by, the rear of the bus wiped on the Day 2 buildings and traffic.\textsuperscript{428}

Through a series of complex compositing processes the effects artists built a virtual 3D space through which the camera can move, achieving with digital methods the continuity and duration praised by Bazin. The film contains several such seamless transitions, which were essential to achieving what are perhaps the most celebrated aspects of the film: the several long takes that occur throughout, usually at moments of heightened intensity. Each of these long takes is an example of digital cinema’s tendency to incorporate multiple elements, often filmed at different locations and times, into a coherent whole, aiming for the illusion of a single unified pro-filmic space that was captured in one take by one camera. But these long takes demonstrate a deeper significance in the history and development of compositing technology. In Cuarón’s long takes, the seamless combination of multiple separate takes, each filmed separately at different times, extends the duration of the shot beyond what was captured during principal photography. These moments are distinct in form from the compositing I examined in chapter 1, though both share the same basic principle. The rear projection and chroma-keying techniques used to create the virtual spaces of \textit{To Catch a Thief} and \textit{Life of Pi} involve the combination of two or more spaces into a coherent whole. The long takes in Cuarón’s cinema are examples of a ‘temporal compositing’ that involves the combination of distinct temporal moments into a seamless continuity. And as with spatial compositing, the digital long takes of extended duration are phenomenologically distinct from their pre-digital equivalents. Digital technology here extends the aesthetic range of cinema, achieving the seamless duration that presented a technological obstacle for previous generations of filmmakers.

Again a contrast with Hitchcock’s cinema can be instructive here. In \textit{Rope} (1948), Hitchcock sought the combination of movement and extended duration that digital camera tracking and other methods of simulating movement would later make possible. The 80-minute film takes place in real time and was shot as 10 distinct long takes, each of which was no more than 10-minutes in length, 10 minutes being the maximum time possible with one reel of film. To achieve the illusion of a continuous take, Hitchcock attempted to hide the transitions between

\textsuperscript{428} Churchill cited in Fordham, ‘Children,’ 34.
the shots using the ‘waiter with the tray of drinks approach’ rejected by Cuarón. Figures 31-34 show an example of one of the film’s transitions. The first movement occurs as the camera moves down from the position shown in the first image to that shown in the second image. There is a small pause on the second position, before the camera moves up and to the right, centering on the character’s back, as shown in the third image. While the frame is predominantly filled with the character’s back, a cut is made, allowing a transition to a new roll of film and a new take, and the camera moves away from the character and into the position shown in the final image.

![Figure 31-34](image)

The sequence is readily identifiable as a transition upon close inspection, failing to achieve the seamlessness that would have made the film an unmitigated technical achievement. Comments from contemporary reviewers reveal that though considered an interesting technical exercise, it was generally seen as an aesthetic failure. Donald Spoto summarizes the response to the film at the time: ‘When it was released in August 1948 - to mixed reviews and a lukewarm public response - there was some talk about the single set, but no one paid much attention to the ten-minute takes.’ In The New York Times Bosley Crowther described it as a ‘stunt’ that is ‘neither effective nor does it appear that it could be.’ In his biography of Hitchcock, Patrick McGilligan reports that the director himself was ambivalent about the film, but felt a sense of pride in its technical accomplishments. The critical reception has improved in the decades since the film’s release, and in his final analysis, McGilligan describes it as a ‘near masterwork, not without flaws, not for all tastes, but the singular experiment of a ceaselessly questing

429 Spoto, Genius, 308.
431 McGilligan, Hitchcock, 421.
Despite this varied critical response, *Rope* represents a significant moment in cinema’s technological and stylistic development, a gesture towards a cinema free from the editing that Bazin saw as an unwelcome intrusion into an audience’s relationship with the reality depicted through the screen. Hitchcock’s attempt at spatial and temporal continuity demonstrates the same aesthetic mode as the long takes of the contemporary digital era. The gradual accumulation of digital tools that increase an artist’s control over the image has made this kind of extended duration much easier to facilitate, a trajectory eventually culminating in Cuarón’s digitally-augmented, seamless long takes, and the 13-minute opening of *Gravity*, with which I began this chapter.

**Gravity: Virtual Reality Cinema**

At the time of its release in October of 2013, *Gravity* represented the most compelling depiction of space the cinema had yet produced. Part of the film’s extraordinary power lay in its provision of an entirely new cinematic experience, representing cinema’s latest gesture toward the ‘total cinema’ that Bazin had identified as the medium’s enduring myth. The film’s unique aesthetic involved a degree of stylistic and technological innovation that ensured much critical attention on its release, but it also demonstrated a degree of formal experimentation unusual in blockbuster cinema, leading some critics to identify a modernist sensibility governing its design. Kristin Thompson, for example, in a two-part exposition that appeared soon after the film’s release, describes the film as ‘an experimental blockbuster’ and outlines several modernist tendencies in Cuarón’s treatment of space. Thompson points out that the constantly moving camera offers little guidance with regards to spatial orientation, rendering any given screen direction ephemeral. Furthermore, the diffuse movement of the characters through the gravity-free space made conventional blocking impossible, so any line of action linking the two characters is almost immediately compromised. There are exceptions to this general pattern of spatial disorientation – there is a conventional pattern of shot/reverse-shot editing when Stone,

---

433 Bazin, ‘Total Cinema.’
caught in a parachute cord, barely manages to hold on to Kowalski (George Clooney) moments before he detaches himself and drifts off into space. But Thompson concludes that for the most part the ‘disorienting simulation of weightlessness for characters and camera dominates the scenes outside the vehicles and creates a style that can truly be called experimental.’

This removal of directional indicators encourages the sensation of actually floating in space, a quality universally recognized by the film’s critics. The spatial disorientation along with the gentle camera movements activates the body’s proprioception system in a way similar to that experienced during the space flight scenes described in \textit{2001} above. But these experimental camera movements are clearly motivated by the narrative context, which as Thompson has suggested, may have helped with their reception by a wide audience.\footnote{Kristin Thompson, ‘GRAVITY, Part 1: Two Characters Adrift in an Experimental Film,’ \textit{Observations on Film Art}, November 7, 2013, http://www.davidbordwell.net/blog/2013/11/07/gravity-part-1-two-characters-adrift-in-an-experimental-film/} However, the film does have other moments that show a more radical deviation from a classical aesthetic. Throughout the film there are abrupt shifts from a representational mode of address to the presentational, revealing a willingness to sacrifice narrative immersion in favor of a modernist exposure of technique. One example occurs 41 minutes into the film, after Stone has just entered the International Space Station after losing Kowalski. As Stone floats through the gravity-free environment, she moves a bag of water, sending water drops floating through the space towards the camera. Stone remains in the background of the shot, while an overt shift in the camera’s focus draws attention away from the character and towards the foreground where the water drops float miraculously in the air. This momentarily displaces character identification as the scene’s governing aesthetic principle in an overt gesture towards the visual marvel occurring on screen (see Figures 35-37).
This moment of direct audience address, a remarkable occurrence in a mainstream Hollywood blockbuster, can be understood by considering cinema’s historical tendency to celebrate novel technologies of visual illusion, a tendency that extends back to the origins of the medium. In *An Aesthetic of Astonishment*, an essay on spectatorship and illusion in early cinema, Tom Gunning recounts the legend surrounding the terrified spectators of the Lumières’ first screening of *Arrival of a Train at a Station* (1895). Gunning questions the accuracy of conventional accounts of this event, but does acknowledge that these early screenings caused ‘shock and astonishment, [and] an excitement pushed to the point of terror.’437 The enthusiasm for the new medium was heightened by the common tendency to begin these early screenings with the film’s first frame projected on the screen as a still image, only cranking the projector to bring the image to life after a brief introduction from the exhibitor. Gunning writes that this ‘strongly heightened the impact of the moment of movement,’ enhancing the audience’s

---

astonishment at ‘the new illusion of projected motion.’ Gunning concludes that the early enthusiasm for the cinema was not the result of a naïve belief in the reality of the image, but emerged from a fascination with the medium’s illusionary power.

Part of Gravity’s appeal for its contemporary audience lies in this enduring fascination with technologies of visual illusion. In the shot described above, our immersion in the ongoing narrative is temporarily suspended as our focus shifts to the astonishing illusion of drops of water, rendered in the latest visual imaging technologies and seemingly liberated from the frame, float in the air towards us, and miraculously, splash onto the now-highly visible camera lens. This shift from the representational to the presentational is a momentary diversion from a classical aesthetic, reinforcing the film’s status as a technological novelty and drawing attention to its powerful extension of the aesthetics of the medium. The sense of awe and astonishment at this moment is not a result of an investment in the film’s narrative but wonder at the unbelievable visual spectacle occurring before our eyes. The fascination here is with the illusion of cinema itself.

But the moment also serves the film’s more general aesthetic project of radically exploring cinema’s ability to render an illusion of spatial depth. By shifting the focus between the background and the foreground, Cuarón emphasizes the depth between these planes and enhances the illusion of penetrable space. This illusion of a space that extends into the far depths of the image, as well as beyond the limits of the frame, forms one half of the spatio-temporal immersion that is the film’s governing aesthetic principle. One technique the film uses to enhance this illusion of spatial depth is depicting objects moving between the foreground and background, linking the two planes and effectively demonstrating the distance separating them. One dramatic example of this occurs as the space shuttle is first hit by the orbiting debris. Figure 38 shows the shuttle just as it comes into contact with the debris, with Stone in the foreground trying to detach herself from the arm connecting her to the shuttle. As the shuttle begins to fall apart, spinning chaotically, Stone is pulled in a dramatic 360-degree arc through space as the shuttle disintegrates around her. In 5-seconds of screen time, Stone moves from the foreground

---

438 Gunning, ‘Astonishment,’ 34.
439 I refer here to a ‘camera lens’ out of convenience, but it is of course a ‘virtual camera.’ See Jones, ‘Vanishing Point.’
of the shot as shown in Figures 38 and 39, to the background as shown in Figures 40 and 41, and then back to the foreground as shown in Figures 42 and 43. Stone traverses an enormous amount of space in this movement from foreground to background to foreground, an effect conveyed through her reduction to a small and distant figure as she is propelled far away from the camera. It is a powerful moment that indicates the rapid escalation in the developing crisis, but it also effectively accentuates the depth of the space in which the drama is occurring.

Figures 38-43

The second half of the film’s aesthetic principle of spatio-temporal immersion centers around the audience’s temporal alignment with the unfolding diegetic events. Thompson has calculated that the events in the film take place in a little over 3 hours, while the film’s runtime is around 83 minutes, a difference that is bridged through several ellipses that occur throughout the
film. With these ellipses aside, the audience’s temporal experience of the diegetic events aligns with their unfolding in real time, an effect enhanced by the extensive use of extended takes. This is an aspect of the film’s aesthetic that its cinematographer, Emmanuel Lubezki, has commented on explicitly. In an interview with American Cinematographer, he outlines the motivation for the use of long takes:

> The main thing about the [long take] is that it is immersive. To me, it feels more real, more intimate, more immediate. The fewer the cuts, the more you are with [the characters]; it’s as if you’re feeling what they’re going through in real time.\(^{441}\)

Lubezki is hinting at an aesthetic principle that extends back to Bazin’s celebration of the Wellesian sequence shot. But like in Children of Men, Gravity’s long takes are synthetic, entirely reliant on cutting-edge technology and the resources of a major Hollywood studio. The digital visual effects in Children of Men are crucial in extending the duration of its long takes beyond what is possible with conventional photography. It represents a significant advance in the cinema’s historical goal of achieving phenomenological continuity in the spectator’s encounter with a diegetic space. Gravity presents a further technical challenge in this regard, as the only photographic elements for large parts of the film were the performers, with the sets, the backgrounds and even the costumes generated digitally.\(^{442}\) But despite this technological virtuosity, the film is still bounded by the principles of classical continuity, adhering to the aesthetics of the Wellesian sequence shot, and grounded in the language of the classical style.

And furthermore, despite its occasional presentational flourishes, the film operates primarily in a mode of address in accordance with Hollywood’s classical aesthetic. Rather than a modernist severing of the contents of an image and its means of representation, much of the film’s formal experimentation, including its extreme shot lengths and exuberant camera movement, represents an extension of the continuity principles that have governed Hollywood

\(^{440}\) Thompson, ‘Part 2.’
\(^{441}\) Cited in Benjamin, ‘Facing,’ 40-41.
cinema since the consolidation of the classical style in the 1910s. In this way, the film represents the kind of ‘intensified continuity’ that Bordwell has argued characterizes the cinema of the late 20th and early 21st centuries more generally.\textsuperscript{443} \textit{Gravity} remains for the most part anchored to a classical illusory realism, generally encouraging an immersive spectatorial response, both perceptually – providing the phenomenological equivalent of presence in space – and narratively – encouraging an immersed identification with Stone and her dilemma.

And because of this general adherence to the principles of classical continuity, the film represents an important development in the immersive potential of cinema. This chapter has focused on the development of technologies of camera movement that have extended the cinema’s tendency towards spatial and temporal continuity, expanding the medium’s aesthetic limits. \textit{Gravity} represents the cinema’s latest iteration in cinema’s ongoing development of increasingly immersive spaces. Earlier in the chapter I examined Kubrick’s forward tracking shots in \textit{The Shining} as an extension of cinema’s ability to depict the illusion of penetrable space. Kubrick’s aesthetic achievement in the film was in exploiting the Steadicam’s exuberant visuality to open up the spaces of the diegetic world, providing a powerful sensation of spatio-temporal immersion in the corridors of the Overlook hotel. But Kubrick’s visuality was limited on the vertical axis. Cuarón’s virtual camera, unhindered by the materiality that pervades conventional camera movements, represents another evolutionary step forward in Hollywood cinema’s spatial aesthetics.

\textsuperscript{443} Bordwell, ‘Intensified.’
Chapter 3 – Emotional Immersion: The Response to Character

In August of 2004, a month before the release of Sky Captain and the World of Tomorrow (2004), Stuart Klawans of The New York Times wrote a skeptical review of some of the innovative CGI that was expected to appear in the film. ‘When Olivier is no longer captured on film but manufactured on the computer,’ he wrote, ‘we lose the very thing that art was supposedly preserving: our point of contact with the irreplaceable, finite person.’ Klawans was referring to a scene in the film featuring the digital recreation of Laurence Olivier as the film’s villain, the sinister mad scientist Dr. Totenkopf. Olivier died in 1989 but was recreated for the role through a careful melding of existing footage of the actor, an effect obviously disguised in the film by the character’s appearance as an electric pulse, and clearly intended as a minor in-joke for those able to recognize the famous actor. However, this digital recreation of a recognizable but long-deceased person is, as Klawans implies, an important moment in our developing relationship to the technologies of visual simulation. In the Language of New Media, Manovich writes that ‘throughout the history of computer animation, the simulation of the human figure has served as a yardstick for measuring the progress of the whole field.’ The human form, to which we are evolutionarily predisposed to be acutely sensitive, is perhaps the most difficult challenge for the visual effects artists attempting to create visual simulations of reality. However, the increasingly realistic simulations of the human form that are appearing more and more frequently in popular cinema and elsewhere, shows that the technology is developing rapidly, and of all the domains in which digital simulations have encroached upon their naturalistic counterparts, it is the representation of the human form that raises the greatest concern. The use of digital simulations in the cinema upsets long-held assumptions about our emotional engagement with cinematic characters, challenging the idea, according to Klawans, of


445 Manovich, Language, 196.
the greatness of a performer as an expression of individuality,’ and compromising ‘the fate of the actor’s soul.’

Since Olivier’s appearance in 2004, there have been several similar digital reanimations involving high-profile figures. Audrey Hepburn, who died in 1993, was featured in a 2013 advertisement for the British chocolate brand Galaxy. Digital recreations of Paul Walker were added to existing footage for Furious 7 (2015) after he died in 2013 midway through the film’s principal photography. And most recently, in 2016’s Rogue One, a Star Wars anthology film, Peter Cushing, who died in 1994, appeared as a recreation of the character he portrayed in the franchise’s original film from 1977. All of these cases attracted considerable public attention and debate, ranging from celebration of the artists and designers involved in achieving such technological milestones, to moralistic aversion to what was held to be a profane act of technological reincarnation.

The artificial recreation of the human form has a long history in the cinema, but as David Rodowick points out in his 2007 book The Virtual Life of Film, digital technologies have led to a significant increase in the frequency with which these recreations have been attempted:

[T]echnological transformations of the film actor’s body in contemporary cinema are indicative of a sea change that is now nearly complete. One could say that the body of the film actor has always been reworked technologically through the use of special makeup, lighting, filters, editing, and so on. Contemporary cinema, however, is taking this process to new levels.

---

446 Klawans, ‘Dead Stars.’
447 Consider for example these comments from a writer on ReelRundown on the digital recreation of Audrey Hepburn: ‘There is, for me, something unnerving in watching this image that is almost a perfect replica of Hepburn, but which very evidently has no soul… On a deep emotional level, I don’t want to go there. It’s as if I’m being tempted by something extremely seductive, but just the slightest bit unclean.’ Ronald E. Franklin, ‘Audrey Hepburn Resurrected for a New TV Commercial – Is this a Good Thing?’ ReelRundown, last updated January 29, 2019, https://reelrundown.com/celebrities/Audrey-Hepburn-Resurrected-For-A-New-TV-Commercial-Is-This-A-Good-Thing.
The concern with digitally recreating the human form in a cinematic context no doubt stems from the belief that it violates a fundamental element of the aesthetics of cinema. André Bazin famously argued that cinema’s ontology derives from its photographic basis, a technological apparatus that preserves an essential connection to the human figures appearing beyond the screen. For Bazin, a cinema without that human connection is a fundamentally different medium. Likewise, for many contemporary viewers, something fundamental to the aesthetics of the medium is lost when a living human ceases to be the object of its focus. I am not unsympathetic to this view. However, it is a view that places the spectator’s recognition of flesh-and-blood actors at the center of the emotional engagement we have with the fictional characters of cinema. In this chapter I will argue that the sources of our emotional engagement in the cinema encompass a broader range than is sometimes accounted for in this view. The recognition that a fictional character is portrayed by a flesh-and-blood actor is just one means of producing the kinds of emotional responses that can lead to a powerful immersion in the diegetic world.

The characters that populate diegetic worlds have been and continue to be important sources of popular engagement with the cinema, something numerous Hollywood script consultants have noticed and placed at the forefront of their advice for aspiring screenwriters. According to Syd Field in *Screenplay*, his popular screenwriting manual first published in 1979: ‘Character is the essential internal foundation of your screenplay. The cornerstone. It is the heart and soul and nervous system of your screenplay. Before you can put one word down on paper, you must know your character.’ Such screenwriting guides have been criticized for hyperbolic literary styles and an overreliance on formulaic conventions, but in focusing on character as essential to the creation of compelling drama, they are correct in identifying one of the most important aspects of our relationship to fictional narratives. Moreover, this focus on character

---

449 Outside of a cinematic context, perhaps the most concerning aspect of this technology is the prospect of its use for political ends. In a 2017 SIGGRAPH paper, researchers from the University of Washington report on their attempt to use a neural network, ‘trained’ on footage of Barak Obama’s weekly presidential addresses, to synthesize a video of the former President speaking words of their choosing. Although the authors’ claim that the procedure produces ‘photorealistic results’ could be debated, the results do raise urgent ethical and legal questions about the development and use of such technology. See Suwajanakorn Supasorn, Steven M. Seitz, and Ira Kemelmacher-Shlizerman. ‘Synthesizing Obama: Learning Lip Sync from Audio.’ *ACM Transactions on Graphics* 36, no. 4 (July2017): 1-13. DOI: http://dx.doi.org/10.1145/3072959.3073640.

450 See for example Bazin, ‘Ontology.’


452 For a critique of popular guides to screenplay writing see David Bordwell, *The Way Hollywood Tells It: Story and Style in Modern Movies* (Berkeley and Los Angeles: University of California Press, 2006), 27-30. See also
can be found among the more measured and scholarly approaches to the question of what makes fictional narratives engaging. For example, in *The Art of Dramatic Writing* (1946), a book frequently cited by these guides as a source of inspiration, Lajos Egri writes:

> There must be something to *generate* tension, something to *create* complication, without any conscious attempt of the playwright’s part to do so… We think we know what that force is: human character, in all its infinite ramifications and dialectical contradictions.⁴⁵³ (italics original)

Likewise, David Bordwell places character at the center of narrative structure in *The Classical Hollywood Cinema*, writing that ‘Character-centered – i.e., personal or psychological – causality is the armature of the classical story.’⁴⁵⁴ Bordwell further defines Hollywood story construction as ‘causality, consequence, psychological motivations, the drive toward overcoming obstacles and achieving goals.’⁴⁵⁵ Similarly, Janet Murray in *Hamlet on the Holodeck*, her analysis of the future of narrative in an era of increasing technological mediation, discusses what she calls ‘movie-rides,’ the immersive experiences found in theme parks and based on movie series such as *Back to the Future* and *Batman*. She points out that these rides maximize perceptual immersion, but usually at the expense of narrative immersion. ‘[M]ovie-rides are providing evidence that audiences are not satisfied by intense sensation alone,’ she writes. ‘Once people do go “into” the movie, they want more than a roller-coaster ride; they want a story.’⁴⁵⁶

This chapter argues that the spectator’s relationship to the characters that inhabit diegetic worlds is a powerful source of immersion. In chapter 1 I examined spatial immersion and the role of compositing in constructing the virtual spaces that form diegetic worlds, and in chapter 2 I examined spatio-temporal immersion and the use of the long take to enhance phenomenological

---


⁴⁵⁴ Bordwell, Staiger, and Thompson, *Classical*, 13. However, Bordwell notes that 'narrative causality could also be impersonal as well.' Among impersonal bases for story action he includes natural causes such as floods, a natural rhythm or cycle of life such as found in the work of Yasujiro Ozu, and finally, a social causality of 'institutions and group processes.'


⁴⁵⁶ Murray, *Hamlet*, 52.
engagement with those spaces. In this chapter, I examine the spectator’s relationship with the characters that populate the virtual worlds of cinema. Fictional characters are potent catalysts for the range of human emotions that allow a spectator to become invested in the events of a diegetic world. Narrative immersion relies on the projection of the spectator’s body into the virtual world of the film, and a fictional character often functions as the point around which such a projection materializes. An effective way to send a spectator to a diegetic world is therefore to populate it with characters that successfully engage the emotions.

But not all emotional responses are equally conducive to diegetic immersion. The first aim of this chapter therefore is to sharpen the distinction between two types of affective response in cinema that are often conflated in critical discourse under a broad notion of ‘immersion.’ The first type consists of the strong affective responses to cinema of the kind I argued characterize *Barry Lyndon* in the previous chapter. These responses inhere in powerful emotions such as sympathy for the characters and their sufferings, and a sorrowful acknowledgement of life’s ephemerality, but despite these powerful affective responses, *Barry Lyndon* generally avoids the kind of character-identification that encourages the projection of the spectator’s perspective to a point within the diegetic world.\(^{457}\) The film is characterized by powerful emotional responses and strong engagement, but it does not lead to the kind of immersion in which a spectator’s conscious experience of the real world is displaced by concern for the characters and drama of the virtual world. The spectator’s powerful emotional response to *Barry Lyndon* comes from a position external to the diegetic world and its characters. I thus described it in the previous chapter as an example of ‘aesthetic immersion,’ indicating the response’s aesthetic quality and signifying its departure from imaginative relocation to the world of the film.

\(^{457}\) Due to its general acceptance, I retain the term ‘identification’ to describe the process by which a spectator aligns with the subjectivity of a diegetic character, but I do acknowledge the term’s conceptual issues. As Noel Carroll has pointed out, identification ‘suggests that we fuse with characters or become one with them which would suggest, at least, that we duplicate their emotional states.’ (96) Carroll continues: ‘the audience’s overall emotional response … diverges from the object of the character’s emotional response in ways that qualitatively differentiate the responses. That the two responses overlap in terms of certain elements … does not indicate that the overall emotional states are the same, or that the audience member takes herself to be the protagonist.’\(^{457}\) Noel Carroll, *The Philosophy of Horror: Or, Paradoxes of the Heart* (New York and London: Routledge, 1990), 93. For an argument in defense of the term, see Berys Gaut, ‘Identification and Emotion in Narrative Film,’ in *Passionate Views: Film, Cognition, and Emotion*, ed. Carl Plantinga and Greg M. Smith (Baltimore: The Johns Hopkins University Press, 1999), 200-216.
Aesthetic immersion is characterized by an aesthetic engagement with the film, rather than an immersion in the world of the film. It is therefore not in itself conducive to the spectator’s relocation into a virtual world that underpins diegetic immersion. For an emotional response to lead to full relocation to the virtual world, the spectator must identify with the subjectivity of a diegetic character. Using the example of Clarice Starling from *The Silence of the Lambs* (1990), I’ll examine two approaches to understanding how such relocation can emerge as a response to the emotional experiences of a fictional character. The first approach involves the cognitive processes involved in the conceptual recognition of others as conscious beings, an ability that psychologists and neuroscientists generally refer to as Theory of Mind (ToM). The second approach focuses on the embodied simulation mechanisms involved in the immediate and pre-conceptual understanding of the emotional experiences of others. These processes allow the spectator to experience the diegetic world through an immersed identification with the subjectivities of fictional characters.

The second aim of this chapter is to explore the prospect of emotional immersion as a response to virtual human-like characters. Visual effects have traditionally been associated with the simulation of non-organic forms, but the increasing application of the technology to the augmentation of actors’ performances has raised the possibility of entirely synthetic creations displacing human actors as the focus of an audience’s emotional investment in the cinema. Attempts to digitally-augment performances have generally been unsuccessful, with several high-profile directors making films featuring digital characters widely regarded as aesthetic failures. The negative reaction to these films has often focused on the characters’ inability to bridge the uncanny valley – the uncomfortable sensation that accompanies the reception of human-like characters that fall just short of achieving perceptual realism. In falling into the uncanny valley, these characters fail to achieve the level of transparency necessary for the technique to recede into the background of the spectator’s conscious awareness. It was not until the extensive effects work used to age Brad Pitt in *The Curious Case of Benjamin Button* (2008) that the technology was widely seen as having achieved a level of proficiency sufficient to enable emotional immersion in a developing narrative. I’ll examine the aesthetic achievements of *Benjamin Button* to argue that the digital simulation of fictional characters presents not a barrier to emotional immersion in the cinema, but an expansion of the medium’s creative and aesthetic horizons.
The Varieties of Aesthetic Immersion

An emotional response experienced independently of a character is a comparatively rare occurrence in classically constructed cinema. Characters are generally the centerpiece around which emotional investment in this form of cinema revolves. Nonetheless, an aesthetic analysis of examples of emotional responses that are experienced independently of diegetic concerns can help clarify those emotional responses that do emerge from an engagement with the diegetic world. Often these ‘extra-diegetic’ emotional experiences involve an authorial voice intruding into the otherwise-transparent narration, a deliberate shift in focus away from diegetic concerns and onto the medium itself. In some cases this occurs at the beginning of the film, before the spectator’s focus becomes locked onto the subjectivities of its characters. The famous opening shot of George Lucas’ Star Wars Episode IV (1977), for example, serves to introduce the diegetic world, but it is also gratuitous, functioning autonomously as an excessive display of technological virtuosity. The sequence opens with a shot of a starry background, before the camera tilts downward to reveal a planet in deep space, and then another planet towards the left in the midground, and then finally the surface of a third planet, which fills the bottom quarter of the frame. John Williams’ rousing score is accompanied by off-camera sounds of spaceships in battle. A small spaceship then enters the frame, moving from the top right corner towards the vanishing point in the center, and then the enormous Star Destroyer enters from the same direction, slowly and majestically filling the frame in a powerful demonstration of the scale and grandeur of the diegetic world we are about to enter (see Figures 44-47).
It is a remarkable display of the technological sublime, an image Scott Bukatman describes as ‘too large for the screen – or our consciousness – to hold.’ And there is a level of communication that exceeds the bounds of narrative, affecting the spectator at a level distinct from the narrative concerns of plot and character. As Bruce Isaacs points out in his analysis of ‘the technological image’ of cinema:

The astonishment of the reveal derives not only from the impressive size and clarity of the image, but from the perfectly regulated, machinic movement that precedes the reveal. The cinematic image of outer space is a complex composition of models, mattes and compositing that draws attention to its technological construction.

The moment encourages a powerful affective response, operating independently of the concerns of character and plot through which our subsequent engagement with the film will be filtered. It displays the presentational mode of address I examined in Gravity in the previous chapter, in a similar gesture to the power of cinematic illusionism, but from an earlier era in the medium’s technological history.

Another example of an affective response that can result from extra-diegetic communication is the use of symbolism. In my previous discussion of Gravity, I argued that though the film generally adheres to the invisibility of technique mandated by Hollywood’s classical style, there are moments of direct audience address that problematize the film’s general tendency to communicate meaning at the diegetic level. One example of this includes the use of symbolism to convey much of Stone’s emotional journey. Figure 48 is taken from just after

---

458 Bukatman, Matters, 93. For more on the technological sublime, see David E. Nye, American Technological Sublime (MIT Press, 1996).


460 While the film received unanimous critical acclaim, some critics took issue with the banality of Stone’s characterization. Hoberman, for example, in his review in The New York Review of Books, refers to a ‘distracting surplus of back story.’ Hoberman, ‘Drowning.’ Similarly, Scott Foundas pointed at that ‘no one is telling their friends to see Cuaron’s film because of its great story.’ Foundas, ‘Avant-Garde.’ However, Thompson points out that the screenplay’s classical features, including a main character with a simple dramatic need with which the audience can easily identify, undertaking a series of clearly defined challenges and identifiable goals, was just enough to ‘sustain the experimental style and motivate that style enough to hold the attention of a popular audience.’ Thompson suggests that this is a general tendency in Hollywood cinema. She writes that ‘strongly innovative films
Stone has managed to enter the International Space Station, moments after Kowalski has been lost to space. Earlier, she had shared with Kowalski the painful account of her daughter’s death, revealing the grief she is suffering. This introduces her ‘dramatic need’ and clearly outlines for the audience the emotional journey she will undertake throughout the film.\textsuperscript{461} Her grief, in addition to her relative inexperience in space, has meant that up until this point she has been reliant on Kowalski, both for his leadership in their work prior to the start of the crisis and for emotional support in dealing with the crisis once it has started. Her success in making it into the ISS without Kowalski thus represents an important step in her journey to achieving emotional independence. Cuarón communicates the significance of this event with an allusion to a fetus in utero, a symbol of Stone’s ‘rebirth.’\textsuperscript{462}

\textit{Figure 48}

\begin{flushright}
\textsuperscript{461} Screenplay manuals distinguish between ‘dramatic wants’ – the external goals a character actively tries to achieve – and ‘dramatic needs’ – the internal psychological struggles a character must overcome, but about which they are often themselves unaware. See for example Field, \textit{Screenplay}, 40-41.

\textsuperscript{462} This is an allusion Cuarón has acknowledged in interviews. See, for example, Meredith Woerner, ‘Gravity’s Ending Holds a Deeper Meaning, says Alfonso Cuaron,’ \textit{Gizmodo}, October 8, 2013, https://io9.gizmodo.com/gravitys-ending-holds-a-deeper-meaning-says-alfonso-c-1442690788?IR=T.
\end{flushright}
This use of symbolism can problematize the immersive moment by too directly revealing the presence of the filmmaker. The spectator’s appreciation of this symbolism requires their movement from what Werner Wolf calls the ‘content level’ of an illusionist representation, to its ‘transmission level,’ effecting their removal from the diegetic world. It is an affecting moment in the film, but it operates outside of an immersive relationship to the film’s diegetic world. The technological image of Star Wars and the symbolic image of Gravity are therefore problematic in the context of an immersive aesthetics. Marveling at the celebration of technology in Star Wars, or understanding Stone’s rebirth through symbolism, involves encountering the cinematic image as an aesthetic object, displacing the utilitarian function it conventionally adopts in the classical style. It turns the immersive image of narrative cinema into an object of aesthetic contemplation.

Other-Directed Emotions

These examples of extra-diegetic communication are rare in classical cinema, and they represent deviations from the heightened engagement with the development of character and plot that characterizes diegetic immersion. In addition to these examples of aesthetic immersion, I would include a range of responses that emerge from the diegetic world, but involve distanced responses to characters, and are distinct from an immersed identification with a character’s subjectivity. These responses are felt about a character, including responses to their values and personalities, and judgements of their behavior. At their broadest, these kinds of emotions can be categorised into either like or dislike for a character, and approval or disapproval of their behavior, but more refined examples include admiration, contempt, sympathy, gratitude, pity, amusement, anger, disgust and Schadenfreude. As an example, I’ll example sympathy, a problematic emotional response in this context because of its overlap with empathy. Alex Neill,

463 Wolf, Immersion, 37.
464 Carl Plantinga describes the emotional responses that emerge from an external perspective on a film’s diegetic world as ‘artifact emotions.’ Artifact emotions are the ‘emotional responses that can be solicited directly by the artifactual status of the film as opposed to the content of the fiction.’ He suggests that examples of artifact emotions might include ‘exhilaration at a particularly brilliant camera movement, disdain for a hackneyed screenplay, anger at the seeming contempt the filmmakers have for the audience, or admiration for the excellence of a film.’ Contrary to my argument here about the challenge these kinds of responses present to the spectator’s immersion in a diegetic world, Plantinga suggests that ‘audiences can have artifact and fiction emotions simultaneously.’ Carl Plantinga, Moving Viewers: American Film and the Spectator’s Experience (University of California Press, 2009), 74.
who distinguishes between sympathy and empathy in a discussion of emotional responses to cinema, writes that among ‘other-focused’ emotional responses, ‘we may distinguish between sympathetic responses (such as those in which I fear for you), and empathetic responses (for I may also fear with you).’\(^{465}\) (italics original) A sympathetic response comes from an external perspective on a character’s situation, whereas an empathetic response involves understanding a character from within. According to Neill, ‘in responding empathetically to another I come to share his feelings, to feel with him; if he is in an emotional state, to empathize with him is to experience the emotion(s) that he experiences.’\(^{466}\) (italics original). A sympathetic response involves a difference between what the character is experiencing and what the spectator is experiencing, while an empathetic response involves a direct correspondence of emotional experience.

The spectator’s moral and ethical predispositions play a role in determining their response to a character, as a spectator with ethical sympathies towards a particular type of person or group of people, for example, may find it easier to form a relationship with such characters, thus deepening their affective response to the diegetic world. But though the range of prejudices and partialities a spectator brings to the cinema will affect their relationships to the characters, an affective engagement with a diegetic world is dependent on more than a direct correspondence between the ethical sympathies of a spectator and a fictional character. The long history of anti-heroes in popular culture – from Travis Bickle in *Taxi Driver* (1976) to Tony Montana in *Scarface* (1983) to Walter White in *Breaking Bad* (2008-2013) – suggests that a spectator’s interest in a fictional character is due to much more than a simple alignment of ethical sympathies. In some cases, a sympathetic relationship is possible despite a character’s immoral behavior, due to the context surrounding the immoral behavior. In other cases, a spectator’s engagement with a fictional character is likely because of their immoral behavior, a result of the fascination and interest such behavior often elicits. In an article on the relationship between perversion, morality and the emotions in cinema, Murray Smith offers a number of explanations for how we may ‘experience sympathetic emotions toward morally perverse or undesirable


\(^{466}\) Neill, ‘Empathy,’ 175.
characters. He describes what he calls the “attractive-bad” character structure, in which a genuinely vicious, corrupt, or immoral character is nevertheless made attractive in some way.

The cannibal and sadist Dr. Hannibal Lecter from The Silence of the Lambs (1990), for example, engages in a range of behaviors that are likely to elicit disgust and aversion from the audience. His cannibalism is clearly outlined towards the beginning of the film, including a vivid description of an incident in which he ate the tongue of a nurse. His cannibalism is again foregrounded when he fondly recalls an incident in which he ate a victim’s liver ‘with some fava beans and a nice chianti.’ He is also shown berating an anxious mother, Senator Ruth Martin, whose daughter has been kidnapped by the serial killer Buffalo Bill. ‘Lecter’s monstrosity is,’ Smith writes, ‘firmly established,’ and yet the audience maintains a sympathetic relationship with the character throughout the film. In a 2001 review, Roger Ebert describes Lecter as ‘a malevolent but somehow likable presence,’ and one of ‘the most memorable characters in movie history.’ He remains a cultural icon whose popularity extends well beyond the film.

Smith offers several explanations for how representations of the perverse, or acts that deliberately violate accepted moral precepts, can engage the interest of the audience.

Representations of the perverse allow audiences to imagine experiences ‘that we lack the opportunity or courage to experience in reality.’ A fictional representation of a forbidden act allows us to undergo a form of ‘imaginative slumming,’ in which we vicariously partake in immoral acts, safe in the knowledge that we are constrained by the fictionality of the representation. Such representations provide, Smith writes, ‘the emotional thrill of imagined-but-controlled danger.’ Furthermore, representations of the perverse can satisfy an audience’s

---

468 Smith, ‘Gangsters,’ 225.
469 Smith, ‘Gangsters,’ 225.
471 Smith, ‘Gangsters,’ 236.
472 Smith, ‘Gangsters,’ 224.
473 Smith. ‘Gangsters,’ 228.
‘curiosity about the extremes of possible or conceivable experience.’ And finally, an audience can be compelled by the ‘delight in provocation’ that representations of the perverse can evoke. Smith traces this aesthetic back to the Aestheticist and Decadent movements of the late nineteenth century and the work of writers such as Poe, Baudelaire, and Wilde, who rejected conventional moral assessments of art and embraced the ‘(perverse) discovery of beauty in acts of violence and depravity.’ In all these ways, representations of morally perverse characters and their actions can be powerful sources of interest and attention.

The Neurophenomenology of Emotional Immersion

These responses to characters, however effective in securing interest and attention, do not facilitate the spectator’s relocation to a point within the diegetic world. They are responses to characters from an external perspective, encouraging engagement with the film, and a heightened interest in the cinematic image as an aesthetic object, but they are not conducive to the diegetic immersion that is my primary focus. For diegetic immersion to occur, a spectator must adopt a viewpoint internal to the diegetic world, a shift in perspective facilitated by the presence of characters into whose subjectivity the spectator can be projected. Filmmakers working within the Hollywood tradition have developed a complex language of formal techniques – POV shots, close-ups, shot/reverse-shot editing – that encourage such a projection to take place. It is a process that turns characters into avatars through whom spectators can experience the vicarious pleasure we associate with a gripping narrative.

The immersive implications of emotional responses to characters are captured by a comment from an anonymous reviewer on an IMDB forum expressing their disappointment after seeing Alfonso Cuarón’s Roma (2018), a film that avoids the engaging characters of conventional Hollywood storytelling in favor of a slow-paced meditation on the director’s childhood:

Underlying cinema is a story. The story has characters, that represent the complexity of the human condition – they have hopes, disappointments,

---

474 Smith, ‘Gangsters,’ 236.
475 Smith, ‘Gangsters,’ 228.
successes and failures. They have different traits. Roma is a work of visual art. It has no story that captivates you. It has no characters that have any real sense of depth or complexity. If it’s [sic] leading actress wins the Oscar it will be a shame… In a nutshell – it is boring.476

This reviewer suggests that complex characters with ‘hopes, disappointments, successes and failures’ are central to successful cinematic storytelling, and films that avoid this form of characterization fail at an essential part of what makes cinema enjoyable. I would speculate that the reviewer expects diegetic characters to be clearly articulated, with their actions and behavior motivated and consistent with expectations about the kind of person they represent. I would further speculate that the reviewer expects to identify with those characters, experiencing a vicarious enjoyment in their experiences throughout the narrative. In describing their disappointment with Roma, this reviewer clarifies the expectations that surround classical spectatorship and lays out the parameters within which this form functions. In doing so, the reviewer implicitly highlights the importance of emotional immersion to most forms of popular engagement with cinema and demonstrates the disappointment that can result when those expectations are not fulfilled. The following section is an attempt to theorize the kind of emotional response this anonymous reviewer was so disappointed in not achieving.

In examining emotional responses to cinema in general, but specifically those that encourage an alignment with the subjectivity of a diegetic character, I’ll necessarily be addressing one of the more problematic aspects of our relationships with fictional characters. This is an aspect of fictional engagement Noël Carroll refers to as the ‘paradox of fiction’, and it concerns the question of how it is possible ‘for us to be moved emotionally by fictions since we know that what is portrayed in a fiction is not actual.’477 Ryan describes the paradox in a similar way: ‘The fate of fictional characters can generate emotional reactions with physical symptoms, such as crying, even though readers know full well that these characters never existed.’478 As I have argued throughout this thesis, our immersion in diegetic worlds depends upon the same perceptual and emotional apparatus that anchor us in the real world. This is a phenomenological

477 Carroll, Philosophy, 87.
478 Ryan, Virtual Reality, 10.
approach to fiction that regards characters, as Ryan puts it, ‘as pseudo human beings rather than as collections of semantic features.’ For a spectator to respond emotionally to a cinematic character, the character somehow emerges from the pixels of the cinematic image as an autonomous individual, capable of generating genuine emotional responses in an audience.

Two theories have been suggested to account for the emotional involvement that underlies our relationship to fictional characters. The first, a classical cognitive perspective on intersubjective understanding, involves the cognitively-demanding process known as Theory of Mind (ToM). ToM has been subject to different interpretations and understandings depending on the theorist, but it generally refers to our ability to recognize and understand the beliefs, intentions, and mental capacities of others. For example, the neurobiologist Robert Sapolsky writes that ToM involves understanding that ‘other people have different thoughts, beliefs, and knowledge’ than our own. The cognitive psychologist Steven Pinker describes ToM as ‘one of the brain’s most striking abilities,’ linking it to important functions involved in intersubjective understanding such as ‘allowing us to predict people’s behavior from their beliefs and desires’ and our ‘ability to empathize.’ The philosopher and neuroscientist Sam Harris offers a more nuanced understanding, distinguishing between two forms of ToM. The first he refers to as ‘standard’ theory of mind, which aligns with the general use of the term. He describes it as attributing mental states to other people, and the phrase ‘theory of mind’ captures this, it seems that we make a much more basic attribution first, and perhaps independently: We recognize that other people are (or can be) aware of us. (italics original)

---

479 Ryan, Virtual Reality, 112.
481 Pinker, Blank Slate, 224.
483 Harris, Waking Up, 111.
The recognition that others can be conscious of us has important implications for cinematic spectatorship. As Harris makes clear, this fundamental theory of mind is absent in our relationships with the fictional characters of cinema and television:

We can view the actions of others, along with the minutiae of their facial expressions – even to the point of making eye-contact with them – without the slightest risk of being observed ourselves. Movies and television magically transform the primordial context of face-to-face encounters, in which human beings have always been subjected to harrowing social lessons, allowing us, for the first time, to devote ourselves wholly to the act of observing other people. This is voyeurism of a transcendental kind.\footnote{Harris, Waking Up, 112-113.}

The important point here is that the absence of a fundamental ToM in cinematic spectatorship does not diminish the heavy emotional investment we so often make in such moments. In fact, as Harris makes clear, it is likely \textit{because} of the absence of this fundamental ToM during cinematic spectatorship that we are able to fully appreciate the emotional experiences of those we are watching. An emotional immersion in the experiences of a fictional character emerges from the ability to fully understand them as autonomous individuals with thoughts, emotions, and feelings just like ourselves, but without the self-consciousness that comes with recognizing that they too might be conscious of us. And as I’ll explore further below, it allows us to simulate the emotions of diegetic characters and experience those emotions as our own.

Filmmakers working within the Hollywood tradition have developed a complex language for encouraging audiences to activate ToM processes in order to foster an emotional connection to diegetic characters. This can be seen in \textit{The Silence of the Lambs}, winner of the 1991 Academy Award for Best Adapted Screenplay, and a film frequently praised as a model of classical storytelling.\footnote{See for example Thompson, Storytelling.} Earlier I examined the spectator’s relationship to Hannibal Lector, describing it as a distanced relationship based on a fascination with Lector’s deviance. But the
film fosters a different relationship with Clarice Starling, a student training to become an FBI agent who is assigned to track down the notorious serial killer Buffalo Bill. Various ToM processes are involved in the development of Starling as a distinct character. The audience is encouraged to align with Starling from the opening credits sequence where she is seen running through the training grounds of the FBI Academy. While she is running, she is summoned to the office of her mentor, Jack Crawford, and is given a briefing on Hannibal Lecter, whom she will interview as part of an ongoing project to develop a ‘psycho-behavioral profile’ of serial killers. As Kristin Thompson points out in *Storytelling in the New Hollywood*, this scene effectively serves as the setup for the narrative that is about to unfold, introducing important narrative information and establishing the film’s dark tone. For example, moments prior to this conversation, as Starling waits in Crawford’s office for him to arrive, she scans the gruesome photographs of the victims of Buffalo Bill that line the walls of the office. Brief glimpses of headlines such as ‘Bill Skins Fifth’ introduce the central line of narrative action involving the hunt for the serial killer who will become Starling’s main antagonist. The gruesome images also establish the violence and danger associated with the situation with which Starling is about to become involved.

Thompson points out some of the ways this opening sequence is important for Starling’s characterization. Her enthusiasm when questioned about her future career plans establishes her ambition and drive and her overarching goal of becoming an FBI agent. This reinforces an impression of Starling formed moments prior to this dialogue, when she is shown walking through the corridors of the FBI building towards Crawford’s office, looking from side-to-side and eagerly observing her surroundings as she passes. This shows, Thompson writes, that ‘she is active, inquisitive, and attuned to her surroundings.’ Furthermore, her dialogue with Crawford makes clear that she is a talented and capable student – in the ‘top quarter of [her] class’ – and that she has earned the respect of her instructors. These examples all help characterize Starling as ambitious, conscientious, and popular among those with whom she works, contributing to the audience’s emerging positive relationship with the character.

---

486 Thompson suggests that Starling’s confrontations with both Buffalo Bill and Hannibal Lecter make *The Silence of the Lambs* a rare example of a film with ‘parallel antagonists.’ Thompson, *Storytelling*, 104-105.
487 Thompson, *Storytelling*, 110.
Contemplating the beliefs, intentions, and desires of others and forming opinions based on observable behavior is important in developing an understanding of a character. However, it cannot fully account for the powerful emotional experience that comes with identifying with the subjectivity of a character. These are what Torben Grodal calls ‘observer theories’ of character relationships. Grodal writes that ‘observer theories often give impoverished descriptions of the film viewer’s experience, focusing mainly on emotions such as pity, admiration, or fear for the protagonist – emotions shaped by the viewer’s distanced relationships to other people.’

The problem with these theories, he argues, is that they ‘implicitly rule out feelings arising from the viewer’s own immersed experience of first-person emotions such as love and fear.’ Theorists such as Grodal and others have advocated approaches to studying spectator-character intersubjectivity that are less reliant on cognitive activity on the spectator’s behalf.

In a paper advocating the introduction of neuroscientific research into literary theory, the neuroscientist Vittorio Gallese and literary theorist Hannah Wojciehowski note that relying solely on ToM to understand emotional engagement with fictional characters runs ‘the risk of reintroducing a backdoor Cartesianism to our understanding of literary texts, a separation of mind and body into distinct divisions and an implicit privileging of one over the other.’ These authors argue that in the classic cognitive understanding of intersubjectivity ‘the understanding of other minds is conceived solely as a predicative, inferential, theory-like process,’ and that it is therefore ‘highly questionable that this model fully captures the expression and functional architecture of the human mind.’ In response, they advocate an approach to humanities scholarship that incorporates research from neuroscience, an approach they call ‘neurohumanism.’ The authors argue that this approach is an important challenge to the view that reason and cognition are disembodied, which they and others view as an inadequate account of the human mind that

488 Grodal, Embodied, 182.
489 Grodal, Embodied, 182.
490 Amy Coplan, for example, argues that emotional contagion, ‘an automatic and involuntary affective process that can occur when we observe others experiencing emotions,’ plays a fundamental role in emotional responses to cinematic characters. Amy Coplan, ‘Catching Characters’ Emotions: Emotional Contagion Responses to Narrative Fiction Film,’ Film Studies 8, (Summer 2006): 26.
492 Gallese and Wojciehowski, ‘Stories.’
493 Gallese and Wojciehowski, ‘Stories.’
discounts key aspects of human cognition and intersubjective understanding. They argue that the human mind is equipped with faculties that can provide a more immediate access to understanding the emotions of others than is often assumed in these disembodied accounts. They write that ‘a more direct access to the meaning of others’ behavior is available – more direct, that is, than the explicit attribution of the propositional attitudes of [ToM]. In particular, they advocate incorporating into narrative theory the neurocognitive process known as ‘embodied simulation’, which they define as ‘a mandatory, prerational, non-introspective process – that is, a physical, and not simply ‘mental’ experience of the mind, emotions, lived experiences and motor intentions of other people. It is a means of intersubjective understanding that bypasses the disembodied accounts of cognition that have been the conventional approaches in literary and film theory, but which have been found to be inadequate in accounting for the experience of emotion and the affective aspects of literary and film aesthetics.

Embodied simulation theory derives from the discovery of mirror neurons, which are the premotor neurons that discharge when a subject performs a motor action as well as when they observe someone else is performing it. As Gallese and Wojciehowski write, ‘[o]bserving an action causes in the observer the automatic activation of the same neural mechanism that is triggered by executing that action oneself.” It has been proposed that this neural mapping mechanism enables the spectator to understand, on an intuitive and immediate level, the actions of others. But the authors go further and suggest that through enabling understanding of the motor actions of others, mirror mechanisms may also be involved in understanding the emotional and sensory experiences of others and are therefore linked to ‘our capacity to empathize with others.” This is a finding with profound implications for the understanding of emotional engagement with fictional characters, and for the immersive power of cinema more generally.

The similar neural expressions involved in both real and fictional imagery, which underpins embodied simulation theory, also accords with my argument for a congruence between real-

---

494 For one of the earliest and strongest critiques of disembodied cognition, see Antonio Damasio, *Descartes’ Error* (London: Vintage, 2006).
495 Gallese and Wojciehowski, ‘Stories.’
496 Gallese and Wojciehowski, ‘Stories.’
497 Gallese and Wojciehowski, ‘Stories.’
498 Gallese and Wojciehowski, ‘Stories.’ The exact role played by mirror neurons in intersubjective understanding is strongly debated. For a sceptical view of their importance, see Sapolsky, *Behave*, 535-542.
world and cinematic perception. Gallese and Wojciehowski suggest that this has implications for the barrier separating the spectator from the diegetic world:

[D]ata show that typical human cognitive activities such as visual and motor mental imagery, far from being exclusively symbolic and propositional, rely and depend upon the activation of sensory-motor brain regions. Visual imagery is somehow equivalent to simulating an actual visual experience, and motor imagery is also somehow equivalent to simulating an actual motor experience. 499

The deep alignment between the spectator and Starling in *The Silence of the Lambs* emerges from these embodied responses to the character’s emotional experiences. One example of how this occurs is the innovative use of an editing pattern of POV shots, a stylistic choice that recurs throughout the film, reinforcing the spectator’s empathy with the protagonist. This editing construction can be seen in Figures 49-50, both of which are taken from the dialogue in which Crawford introduces the case Starling is about to undertake.

![Figures 49-50](image)

As Crawford describes Lecter and provides background on Starling’s assignment, a shot-reverse shot editing pattern alternates between close-ups of the two characters. A subtle distinction between the characters’ eyelines helps align the audience emotionally with Starling. During her close-up Starling looks slightly off camera at Crawford, who is positioned to the left

---

499 Gallese and Wojciehowski, ‘Stories.’ The authors also argue that these findings have implications beyond just the processing of visual images. ‘[T]he premotor cortex exhibiting the Mirror Mechanism,’ they write, ‘is also involved in processing action-related words and sentences, suggesting that mirror neurons, together with other parts of the sensory-motor system, could play a relevant role in the semantics of language.’
of frame. This is typical of classical shot/reverse-shot editing patterns that alternate between the two characters in the dialogue. However, this scene introduces what Bordwell calls ‘point-of-view cutting’ into what is otherwise a conventional editing pattern for a dialogue scene. Bordwell writes that with point-of-view cutting, ‘[t]he first shot shows the character looking at something offscreen; the second shot shows what the character is seeing, but more or less from the character’s optical vantage point.’\footnote{Bordwell, Staiger, and Thompson, \textit{Classical}, 57.} This means that the cut away from Starling’s close-up to Crawford’s close-up is a POV shot, with the spectator occupying Starling’s position within the scene. This reveals Crawford looking straight at the camera, a rare occurrence in a classically-constructed film.

This unique editing construction has several effects. Crawford’s eye-contact reflects his dominant position in this mentor-student relationship, so the editing reinforces the power dynamics operating within the scene. Crawford’s eye-contact with the audience also challenges the classically-mandated removal of the spectator from the diegesis, resulting in the unsettling experience of having another person stare intently at oneself, a sensation usually absent in classical cinematic spectatorship. This contributes to the uneasiness that Starling is presumably experiencing and associating with the investigation she is about to undertake. The same effect occurs in the next scene, as Starling meets with Dr. Chilton, the sneering and arrogant director of the hospital in which Lecter is being held. During their dialogue, Chilton looks directly at the camera, while Starling looks off camera at Chilton. Figures 51 and 52 show Chilton’s close-up as he appraises Starling’s physical appearance, the first in a series of comments that demean Starling and challenge her professional competence. Starling’s close-up reaction shot shows her uncomfortable response. Here again the audience experiences the unease produced by direct eye-
contact with a diegetic character, an unconventional technique that instills in the audience the same unease that Starling is experiencing in the scene.

Close-ups are a particularly powerful means of inducing an embodied response in the spectator, and they are used effectively throughout the film to encourage the spectator to simulate Starling’s emotions. In Figures 53-58, taken from a scene just after Lecter has made a deal with Dr. Chilton to reveal the identity of Buffalo Bill, a close-up of Starling conveys the painful childhood trauma that has been motivating her pursuit of Buffalo Bill throughout the film. In the scene, Lecter has been moved to a Tennessee Courthouse where he is held in a cage in the center of a large hall. Unlike the other police officers involved in the investigation, Starling has realized that Lecter failed to reveal the identity of Buffalo Bill, despite agreeing to do so as part of the deal with Dr. Chilton. Starling brings Lecter’s drawings to him, hoping he will reveal the killer to her. The scene represents the final encounter between Starling and Lecter, and it is an important moment in the development of Starling’s dramatic need – acknowledging the trauma she experienced as a child when she was unable to save the dying lambs. During their dialogue, Lecter makes explicit the role this trauma has had in her decision to become an FBI agent and in her effort to find Buffalo Bill. It is during this scene that she painfully acknowledges for the first time the significance of this childhood memory, an important moment in her character arc that will only fully be resolved when she finds Buffalo Bill and rescues Catherine Martin in the film’s climax.

501 In a discussion of emotional immersion, Grodal points out the importance of close-ups in inducing an empathetic response in a spectator: ‘The depth of our simulation of the characters in a film is deeply influenced by the degree of salience in their representation. A close-up revealing strong interest or emotion in the face of a character is much more salient than a long shot, since the resolution with which, for instance, facial expressions and eye movements are shown is so much higher.’ Grodal, *Embodied*, 201.
The scene follows a conventional pattern of shot/reverse-shot editing, clearly delineating the positions of the two characters and directing the spectator’s attention to each of them at appropriate moments in their dialogue. As the scene progresses, a combination of gentle camera movements and closer framings lead to a steady movement in towards the characters, with their faces occupying increasingly large proportions of the frame. Figures 53 and 54 show Starling at the start of the scene, as she hesitantly approaches Lecter hoping he will cooperate with her, and Lecter’s position behind the bars, carefully watching her as she attempts to extract information from him. As Lecter begins to berate Starling, his face occupies more and more of the frame, culminating in his close-up in Figure 56, where he stares at both Starling and the spectator in a domineering manner. Figure 57 shows Starling’s timid effort to stand up to him. Her unease and pain throughout the scene culminate in Figure 58, where she finally gives in to Lecter’s pressure, revealing the personal information that he wants.

The full emotional resonance of the moment is conveyed using the close-up in Figure 58. Starling’s wounded expression as she reveals her painful childhood trauma activates the embodied simulation processes through which we intuitively understand the emotional experiences of others. This is a level of emotional communication that is pre-cognitive, a means
of understanding the character that bypasses the more cognitively demanding ToM mechanisms that are sometimes seen as central to intersubjective understanding. As Gallese and Wojciehowski suggest, ‘[w]hen we see someone acting or expressing a given emotional or somatosensory state, we can directly grasp its content without the need to reason explicitly about it.’ But this emotional response is not entirely decoupled from ToM mechanisms – the cognitive activity involved in recognizing and understanding Starling’s verbal account of her trauma is essential for contextualizing the painful emotions revealed by her close-up. Viewing this close-up independently of Starling’s dialogue would minimize its emotional resonance. But it is through the close-up – cinema’s most salient means of conveying a character’s emotional experience – that the spectator is able to simulate the emotions of the character, fostering the kind of identification that reinforces an emotional immersion in the diegesis. Embodied simulation thus helps explain how the emotional experience of the spectator can be aligned with that of a character, providing the point around which an immersed identification can take place. In this way a spectator can adopt a viewpoint internal to the diegetic world, showing how, as Gallese and Wojciehowski suggest, ‘the border between real and fictional worlds is much more blurred than one would expect.’

Simulation and its Discontents

An emotional immersion in the experiences of diegetic characters is central to classical storytelling, and it is an aspect of popular engagement with the cinema that has bridged the digital and pre-digital eras. Throughout his book on digital visual effects, Stephen Prince makes the point that in recognizing the changes that have occurred as a result of the transition to digital technology, we should not overlook those aspects that have remained unchanged. Prince elaborates on one important aspect of characterization that has remained consistent:

Actors provide the human element in cinema, a medium that otherwise is heavily dependent upon machinery for creating light and color and capturing images and sounds. And yet the actor’s presence is

---

502 Gallese and Wojciehowski, ‘Stories.’
503 Gallese and Wojciehowski, ‘Stories.’
paradoxical. A viewer’s impression of wholeness – the actor as a unified being in front of the camera – and of psychological and emotional continuity – the actor as character unfolding in narrative time and space – is a manufactured impression that often fails to correlate with what was. 504

With the exception of the kind of Wellesian long takes that were so highly praised by Bazin, the performances of the actors on set before the camera rarely make it through to the final edit untouched. Instead they are subject to significant alteration, recombination and rearrangement as they are edited together through the process of continuity editing. Like all other elements of cinema, the cinematic illusion by which characters appear to exist continuously in space and time is constructed by the careful arrangement of separate takes from different camera positions into a coherent whole. To clarify how this illusion is constructed, Prince makes an important distinction between acting – the ‘behavior that occurs on set to portray characters and story action’ – and performance – ‘the subsequent manipulation of that behavior by filmmakers or by actors and filmmakers,’ 505 and notes that ‘[p]erformance in cinema has always been a construction synthesized from discrete elements removed from their original contexts, rearranged, reordered, reshaped.’ 506 The digital simulation of the human form thus represents a less significant break from previous means of representing the human form than appears at first to be the case.

In drawing attention to these continuities, however, it is also important to not overlook the significant differences represented by the technology. Prince does acknowledge, for example, that digital tools ‘have given filmmakers new powers to manipulate color and to composite images’ and have ‘augmented the existing practices of completing or creating performances in post-production.’ 507 These tools include motion capture technology that tracks the bright dots attached to an actor, capturing their broad movements and allowing precise animation of the performance during post-production. It also includes digital painting techniques that can be used

504 Prince, Digital, 100.
505 Prince, Digital, 102.
to add or remove elements from an image, accurately capturing precise visual features such as the interaction of light with the surface of a character’s skin. And finally, digital facial animation allows for the precise control of a character’s facial expression, allowing the synthesis of emotional expression beyond what was achieved on set.\textsuperscript{508} All these techniques represent significant changes to the way performance is conceptualized in digital cinema, challenging more traditional understandings of the relationship of spectators to the characters that inhabit diegetic worlds. And by focusing on the continuities over the changes in the way performance is constructed, it is possible to overlook what may be an important change in our relationship to cinematic characters: a Bazinian appreciation of the human figure that existed at one point in time to be captured by the photographic medium.

However, emotional engagement with diegetic characters does not depend exclusively on the presence of flesh-and-blood human actors. In a similar way to how an emotionally compelling relationship to a literary character can develop from the sensations and images created in the reader’s mind through words, a cinematic spectator can emotionally connect with characters beyond those that can be represented by a human actor. As Lisa Bode argues, ‘it is completely possible to find points of identification and empathy with fictional characters whose bodies are not rooted in a life beyond the fiction. After all, how many gallons of tears have been shed for Woody from \textit{Toy Story} (1995), and Ellie and Carl from \textit{Up} (2009)?’\textsuperscript{509} Animators have long known of the mind’s ability to attribute personalities and emotions to diegetic characters based on just the barest outline of the character’s physical form.\textsuperscript{510} Prince suggests that this ability emerges from an evolutionary predisposition towards detecting and responding to agency in our environment: ‘character animation takes advantage of the human perceptual system’s fine-tuned propensity to scan objects and environments for signs of intention and to read these signs


\textsuperscript{510} This ability is also seen in Prince’s description of Pixar’s first production, \textit{Luxo, Jr.} (1986), a short written and directed by John Lasseter after leaving Lucasfilm to found the company: ‘The charming two-and-a-half-minute film shows a comic situation enacted by two Anglepoise (extension) desk amps, one large and one small. The animation endows these with personalities and the roles of father and child. A ball rolls onscreen. Dad bats it away with his lamp hood, and Junior chases it enthusiastically, hopping across the screen. Squash-and-stretch, timing, and exaggeration delineate the characters and their emotions. Dad moves slowly, with gravitas, Junior with quicker, chippier actions.’ Prince, \textit{Digital}, 106.
often on the basis of scant and incomplete evidence.’ Studies suggest that this understanding of intention in animated characters is largely a result of the observation and interpretation of the character’s movements. According to neuroscientists Sarah-Jayne Blakemore and Jean Decety, ‘[p]sychophysical and neurophysiological studies support the idea that the brain is a powerful simulating machine, designed to detect biological motion in order to extract intentions from the motion and to predict the future actions of other animate beings.’ The human mind is acutely sensitive to the perception of agency in others: the use of point light display stimulus in the study of motion shows that merely the suggestion of movement is necessary to invoke the simulative processes necessary for understanding and attributing emotional states to others. The sense that an animated character possesses a life beyond the outlines of the physical form suggested by an animation is thus heavily dependent on the movement conveyed by the careful inscriptions of the animator. The full process of character identification – the spectator’s attribution of an emotional state to a character as well as some degree of shared affect – is therefore largely dependent on the character’s movements. If we add to movement the other important components of characterization, such as the voice of the actor, the dialogue, their role in a plot, and their relationships to other characters, we can begin to see how character identification can occur in response to animations that sketch out a physical form that only barely corresponds to a real-life human.

Perceptual realism is therefore clearly not necessary for emotional engagement with an animated character. Emotional immersion in a fictional diegesis is possible with animated characters that barely resemble real-life humans. Nonetheless, as an animated character moves closer to human facial expressions and movements the embodied simulation mechanisms through which we understand and experience the emotions of others are intensified. The

---

511 Prince, Digital, 107. Blakemore and Decety offer a similar explanation: ‘The visual perception of motion is a particularly crucial source of sensory input. It is essential to be able to pick out the motion of biological forms from other types of motion in the natural environment in order to predict the actions of other individuals.’ Sarah-Jayne Blakemore and Jean Decety, ‘From the Perception of Action to the Understanding of Intention,’ Nature Reviews Neuroscience 2, no. 8 (2001): 561.

512 Blakemore and Decety, 566.

513 Point light display stimulus was first used by the perceptual psychologist Gunnar Johansson, who in a 1973 study attached bright lights to key points on the bodies of subjects to study the ‘visual perception of motion patterns characteristic of living organisms in locomotion.’ He found that ‘10-12 such elements in adequate motion combinations in proximal stimulus evoke a compelling impression of human walking, running, dancing, etc.’ Gunnar Johansson, ‘Visual Perception of Biological Motion and a Model for its Analysis,’ Perception & Psychophysics 14, no. 3 (1973): 201.
perceptually realistic simulation of the human form therefore remains a goal for many visual effects artists and filmmakers.\textsuperscript{514} Such attempts have so far been relatively rare, due to the enormous financial costs involved in the process, as well as its considerable technological sophistication. But decreasing costs associated with the technology as well as its increasing availability have led to a corresponding increase in the number of attempts at achieving complete transparency in digitally simulating the human form, with studios, filmmakers and visual effects artists no doubt tempted by what is perhaps the most technologically demanding and complex visual effects challenge there is. One attempt, mentioned in the opening of this chapter, were the digital recreations in \textit{Rogue One} of the actors Peter Cushing and Carrie Fisher for the characters they portrayed in the original \textit{Star Wars} film from 1977. The scenes involving the simulations attracted considerable controversy, which led to the filmmakers justifying the extensive effects work involved in both cases on aesthetic grounds. Kiri Hart, a Lucasfilm story analyst, writer, and development executive, points out that the film depicts the events leading up to those depicted in the 1977 original, which meant there was significant crossover with the characters in the two films. The recreations were deemed necessary to ensure consistency of characterization across the two productions, which were separated by nearly forty years. ‘If he’s not in the movie,’ Hart argued, speaking of the recreation of Peter Cushing, ‘we’re going to have to explain why he’s not in the movie.’\textsuperscript{515} Similarly, Carrie Fisher was recreated for a scene towards the end of the film where she delivers just one word of dialogue in response to a question about a transmission she has received. The person delivering the transmission asks: ‘What is it they’ve sent us?’ Leia’s response – ‘hope’ – functions as a narrative bridge between the two films, but also resonates as the film’s thematic and emotional climax, emphasizing the hope that has resulted from the sacrifice made by those who have died bringing the transmission to Leia. Hart justified the enormous effort and expense of recreating Fisher for this brief scene on the grounds that actually seeing her face enhances the sense of hopefulness. ‘That’s the best possible use of

\textsuperscript{514} Tom Gunning points out that there are other factors motivating this desire to create digital simulations of the human form. He discusses digital simulation in the context of Bazin’s claims in ‘The Myth of Total Cinema,’ suggesting it may be part of a more fundamental human desire to represent the world, a desire with long historical antecedents. ‘The dream of creating an artificial human,’ he writes, ‘reaches at least as far back as the Jewish mystical tradition of the Golem and its magical repetition of the original genesis of Adam.’ Gunning, ‘Special Effects,’ 325.

effects,’ Hart argued, ‘to enhance the meaning and the emotion of the experience for the viewer.’

These claims were attempts to justify the use of the recreations against the charge that they were a crass technological novelty, functioning as a form of product differentiation by which a major Hollywood studio can set itself apart from competitors by displaying its latest achievements in visual imagery. However genuine Hart’s comments, the general response from audiences and critics demonstrated indifference to these justifications, and skepticism and at times anger towards the recreations, perceiving them as both moral and aesthetic indignities. Zac Thompson in The Huffington Post argued that it is a ‘giant breach of respect for the dead,’ and in removing ‘the human element to the performance … it removes the art of the craft.’ Catherine Shoard writing in The Guardian felt that the digital recreations were done ‘remarkably well’, but that the ‘dignity of death ought to be preserved.’ These criticisms focus on the ethical concerns surrounding the use of digital technology to synthesize the human form, but other critics targeted the technological limitations of the recreations and the resulting aesthetic implications. Emily Asher-Perrin, for example, wrote that the ‘CGI use was both jarring and exceedingly creepy’ and that ‘no amount of computer advancement can change the fact that a CGI simulacrum cannot act as Cushing could.’ Finally, and most relevant for my discussion of emotional immersion, Kelly Lawler writing for USAToday emphasized the implications of the simulations for the audience’s emotional engagement:

516 Cited in Itzkoff, ‘Familiar.’
517 For more on technological novelty as product differentiation see my discussion of Life of Pi in chapter 2.
518 However, this negative response was not unanimous. For an example of a positive response to the CGI see Radio Times, ‘How a Holby City Actor Brought One of Star Wars’ Most Iconic Characters Back to Life,’ December 22, 2016. https://www.radiotimes.com/news/2016-12-22/how-a-holby-city-actor-brought-one-of-star-wars-most-iconic-characters-back-to-life/.
But while Tarkin is merely unnerving, the Leia cameo is so jarring as to take the audience completely out of the film at its most emotional moment. Leia’s appearance was meant to help the film end on a hopeful note (quite literally, as “hope” is her line), but instead it ends on a weird and unsettling one.522

These criticisms are possibly motivated by several overlapping concerns surrounding digital simulation. One of these concerns involves a reactionary response to the technology’s perceived encroachment on conventional practices and the threat represented to actors, makeup artists, costume designers, and others involved in the traditional means of cinematic characterization. A second concern, specific to this case from Rogue One, surrounds the heightened sensitivity that emerges from the intense sense of ownership felt towards the Star Wars franchise by its fans.523 The third concern is the most fundamental of the three, and it consists of a belief, usually vague and undefined, that using digital imagery to simulate the human form presents a violation of a taboo or a transgression of the natural order. According to this view, using digital technology to replicate a human crosses an ethical line that does not apply when the technology is used in other cases, such as simulating an environment, an object, or non-human characters such as animals. This third source of aversion I suspect emerges from a more fundamental anxiety surrounding representation and our relationship to digital images, concerned specifically with the dualism associated with René Descartes, an ontology that has been labelled and critiqued as the belief in a ‘Ghost in the Machine’ by the philosopher Gilbert Ryle.524 Of the three sources of aversion to the digital simulations in Rogue One it is the one

524 Ryle, Gilbert. The Concept of Mind (London: Routledge, 2009). Steven Pinker suggests that the belief first emerged as a coherent idea as a response to Thomas Hobbes, who had argued that ‘life and mind could be explained in mechanical terms.’ Pinker, Blank Slate, 9. Hobbes’ arguments were an early precursor to the idea, widely accepted today among cognitive neuroscientists and philosophers of mind, that consciousness is an emergent property resulting from neurobiological processes. This challenges the more conventional and intuitive view that consciousness exists independently of the body and functions autonomously, the inspiration for Descartes’ dualism and his challenge to the reduction of consciousness to the mechanics of neurobiology. I suggest that general cultural aversion to digital simulations of the human form may stem from their uncomfortable congruence with the
most likely to persist as a source of aversion. The first and the second concerns – the response to challenges to industry and aesthetic norms, and the sensitivity surrounding a beloved franchise – are related to this specific case, and it is easy to imagine other contexts in which the use of such simulations would not rouse such strong aversion from its critics.

However, there remain technological limitations to the processes involved in digital simulation that will continue to pose a significant barrier to its transparent use in the service of an immersive aesthetic. In this next section I’ll address this issue, approaching the digital simulation of the human form as a technological and aesthetic problem. Several questions motivate this discussion. Is it possible that digital technology could achieve perceptual realism in simulating the human form? Could the technology ever acquire the level of invisibility necessary to recede into the background of the spectator’s conscious awareness, achieving the transparency necessary for full immersion in the diegetic world? And perhaps most fundamentally, could a digitally-augmented performance ever induce in the spectator the level of emotional engagement that defines the spectator’s relationship to Clarice Starling?

The Uncanny Valley

One of the main obstacles to achieving emotional immersion in response to simulated human characters is the phenomenon known as the ‘uncanny valley’ first described in 1970 by the roboticist Masahiro Mori.525 Mori noticed that there was a powerful sense of revulsion felt towards humanlike robots as they approached but failed to achieve a lifelike appearance. He suggested that the aversion was minimal when the simulations were far from lifelike, such as with toys and industrial robots, but as the simulations came to resemble humans more and more closely there was a notable increase in the aversion felt towards them. Prince describes the phenomenon as it relates to the digital simulation of human characters in cinema:

A threshold is crossed where the imitation becomes so close and exacting that its remaining incompleteness points to its status as a surrogate, as something not real, and this results in a loss of empathy from viewers, a pulling back, as what had seemed so familiar becomes defamiliarized.\textsuperscript{526}

Such a response of course contrasts dramatically with the kind of immersed engagement with a fictional character that I’ve been describing throughout this chapter. Prince’s use of words such as ‘pulling back’ and ‘defamiliarized’ shows how a failure to achieve perceptual realism can be fatal to the processes of embodied simulation that I’ve argued are central to immersed engagement. Simulations that fall short of perceptual realism draw attention to their status as artificial recreations, presenting a serious obstacle to the spectator’s ability to appreciate them as autonomous beings with independent emotional and cognitive experiences. But the uncanny response to digitally-simulated characters can go further than this, evoking stronger responses of aversion and disgust. The use of the adjectives ‘weird’ and ‘unsettling’ to describe Leia from Rogue One in the review quoted above hint at these stronger feelings. This association with the abject was also noted by Mori, for whom the uncanny was particularly strong in instances where the simulation resembled the human form but demonstrated violations of healthy predictable behavior, such as with corpses and zombies.\textsuperscript{527}

The uncanny valley is conventionally seen as a technological problem that will eventually be overcome through increased research and development, both from research institutions with the budgets and expertise necessary for such development, and from major studios with the incentive to develop the technology to create novel cinematic experiences. Such technologically deterministic views, however, fail to account for the variety of factors that could disrupt the ongoing evolution of simulative technology towards achieving perceptually realistic human characters. These factors include opposition from actors, make-up artists, costume designers, and other industry figures with an incentive to oppose the technology, a cultural aversion to crossing

\textsuperscript{526} Prince, Digital, 121.

\textsuperscript{527} Angela Tinwell has shown, however, that the uncanny effect can be used aesthetically in the service of immersion. She shows that survival horror video games such as Left 4 Dead (2008) and Silent Hill Homecoming (2008) use the uncanny to enhance the horror of the characters, contributing to immersion, rather than detracting from it. See Angela Tinwell, The Uncanny Valley in Games and Animation (New York: AK Peters/CRC Press, 2014). DOI https://doi.org/10.1201/b17830.
the ethical line that separates the simulation of non-human and human forms, and a cultural preference for fictional characters being portrayed by human stars, who, as Robert Allen notes, are ‘social phenomena’ that ‘form an aesthetic intertext that audiences use to derive meaning and pleasure from films.’ These factors could prevent the technology from achieving the sophistication necessary for perceptually realistic human characters, or, if the technology does develop to that level, prevent it from achieving widespread industry adoption.

There may be an even more intractable barrier to the use of human-like virtual characters in mainstream cinema. Angela Tinwell, in her extensive studies on the uncanny valley in games and animation, suggests that the eerie and aversive responses associated with the uncanny may be an unavoidable part of our relationship to simulations of the human form. To demonstrate, Tinwell describes the creation of Emily, a computer-generated character that was the outcome of extensive research by Paul Debevec with the USC Institute for Creative Technologies, and Image Metrics, a 3D facial animation technology company. The creation of Emily in 2008 was heralded as a landmark in the challenge of simulating the human form and a victory in overcoming the uncanny valley. As an example of one of the more enthusiastic responses, consider the following excerpt from a review by Peter Plantec, an author and animator, who wrote about the virtual character for *VFXWorld* magazine:

> It is absolutely awesome – amazing. I’m one of the toughest critics of face capture, and even I have to admit, these guys have nailed it. This is the first virtual human animated sequence that completely bypasses all my subconscious warnings. I get the feeling of Emily as a person. All the subtlety is there. This is no hype job, it’s the real thing.

Emily is a remarkable achievement in visual simulation and a significant improvement over previous attempts to achieve perceptual realism in depicting the human form. However, the enthusiastic responses to the character, such as Plantec’s claim that ‘Image Metrics has finally

---


built a bridge across the Uncanny Valley and brought us to the other side,’ may be overstating the significance of what is admittedly an enormous technological achievement. One of the problems with comments like those from Plantec may stem from a conceptual problem associated with the uncanny valley itself. Tinwell, for example, argues against the assumption that the uncanny valley represents a fixed problem that can be solved, or in Plantec’s metaphor, an obstacle over which a bridge can be built. According to Tinwell, the assumption that there is a clearly identifiable level of realism at which the uncanny valley problem is overcome disregards the enormous complexity of our relationship to virtual human-like characters. In order to demonstrate the complexity of this relationship, Tinwell and Mark Grimshaw conducted a study involving the assessment of virtual characters using a scale based on factors such as the perceived familiarity, the perceived human-likeness, and the degree to which the characters behaved as they were expected to. The study required participants to view 14 video clips of virtual characters and 1 clip of a real person and judge each of the characters using a nine-point scale. Tinwell and Grimshaw found that the virtual characters were consistently judged to be less familiar, less human-like, and less likely to behave as expected when compared with the human subject. Emily too was judged, the authors write, ‘consistently as less familiar than the video of a human being.’

Tinwell and Grimshaw note that the participants in the study were university students from Bolton University’s School of Games Computing and Creative Technologies, as well as professionals working within the academic sector and video games industry. They are thus likely to be highly literate in the technologies of animation and digital simulation and more discerning in their assessments of realism in relation to human-like virtual characters. The authors note that the sensitivity these participants had to slight distinctions in the characters reveals important facts about the degree to which realism is culturally-determined. They argue that ‘increasing

530 Plantec, ‘Digital.’
531 The participants were asked how human-like they perceived the character to be from nonhuman-like (1) to very human-like (9). To measure the characters’ perceived eeriness, they were asked to rate how strange (eerie) or familiar they perceived the character to be from very strange (1) to very familiar (9). To measure the relationship between expected behavior and perceived familiarity for a character, participants were asked to rate from Strongly Disagree (1) to Strongly Agree (9) their response to the statement: “Based on the character's appearance, this character behaves in a way that I would expect them to.” Angela Tinwell and Mark Grimshaw, ‘Bridging the Uncanny: An Impossible Traverse?’ Games Computing and Creative Technologies, Conference Paper, (2009). http://digitalcommons.bolton.ac.uk/gcct_conferencepr/10.
532 Tinwell and Grimshaw, ‘Bridging.’
technological sophistication in the creation of realism for human-like virtual characters is matched by increasing technological discernment on the part of the viewer.' As simulative technologies develop in their sophistication, so too does the visual literacy of audiences, and they therefore, Tinwell argues, ‘become ever more discerning of imperfections in a character’s behavior from the human norm and increasingly sensitive to the uncanny in human-like virtual characters over time.’ Tinwell summarizes the implications these findings have for the relevance of the concept of the uncanny valley in understanding our relationship to human-like virtual characters:

Rather than a valley, we are faced with an uncanny wall that continues to grow in height. As time allows progress in technology to create human-like characters with an increased sophistication in graphical and behavioral fidelity, at the same time our perceptual expertise for detecting imperfections or abnormal behavior in a human-like character improves. Thus, the Uncanny Wall increases in height as a viewer’s ability to detect imperfections from normal human behavior keeps pace with technological developments in creating high-fidelity human-like characters.

Tinwell’s findings deepen our understanding of the development of simulative technology and our relationship to virtual characters, and helps to clarify the degree to which notions of realism are culturally constructed. However, the conclusion that a sense of uncanniness is inevitable in response to virtual characters rejects even the theoretical possibility of synthetic characters achieving the perceptual realism necessary to entirely avoid the uncanny. Though there remain significant technological barriers to achieving that level of perceptual realism, it is not theoretically impossible.

The positive critical reception of the visual effects work in David Fincher’s 2008 film *The Curious Case of Benjamin Button* suggests that digitally-augmented characters may already

---

533 Tinwell and Grimshaw, ‘Bridging.’
have been integrated into Hollywood’s transparent style. The film’s plot follows the story of Benjamin Button, a character born as an old man who ages backwards, losing those around him as he moves through life, eventually ending the film as an infant. Throughout his life he meets and befriends other characters, all of whom inevitably age and pass on while he continues growing younger. The most significant relationship in his life is with Daisy, a character he meets early in the film as a young girl, and with whom he will eventually have a daughter. Like the other characters, however, Daisy continues to age in the opposite direction to Benjamin, separating the characters and highlighting the tragedy inherent in Benjamin’s condition.

Extensive visual effects work was needed to depict Benjamin moving through the different stages of his life, and though anchored throughout by the performance of Brad Pitt, for large sections of the film the character is essentially a digital composite. The early scenes featuring Pitt as the aged character involved extensive combinations of digital post-production and on-set makeup effects, as well as the performances of multiple actors as stand-ins to achieve the age-appropriate look for the character at this stage of his life. It is towards the end of the film, however, where the most successful use of digital visual effects occurs. After several scenes showing the happy life Daisy and Benjamin have together, Benjamin leaves her after deciding that his condition will make it impossible for Daisy and their child to have a normal life if he remains with them. Benjamin spends the next decade of his life travelling alone, and after a long absence, he eventually returns to visit Daisy, and the couple is reunited for a brief encounter. Daisy has aged considerably since they last met, while Benjamin is younger, and their brief reunion highlights the dramatic contrast between the two characters at their different stages in life.

The scene begins with Daisy in her dance studio, closing for the end of the day and saying goodbye to her students. After the students have left the studio, she encounters Benjamin in a sequence of shots edited to emphasize the emotional toll the encounter has on her. In a

---

536 Prince praises the film’s insistence on using Pitt for the character throughout the film, arguing it is a significant improvement over other films that use different actors to portray a character at different ages. However, he criticizes the decision to use a different actor towards the end of the film: ‘Abandoning this strategy in the last act of the narrative produces a jolt that breaks the audience’s illusion that it is seeing a single character enacted by a single performer, an illusion developed seamlessly to this point.’ Prince, Digital, 141.

conventional shot/reverse-shot editing pattern, Daisy is first shown noticing someone at the door of the studio. A cut then reveals the shadowy figure she has seen (Figure 58), and another cut back to Daisy shows her recognizing that it is Benjamin, her wounded facial expression revealing the painful emotions she is experiencing (Figure 59). A cut to a close-up of Benjamin shows that he is significantly younger than when he was last seen (Figure 60).

Beginning the scene with Daisy aligns the spectator with her point-of-view, and Benjamin’s abrupt appearance as a younger man is a shock to both her and the spectator. The moment of Benjamin’s reveal in Figure 60 is heavily dependent on extensive visual effects work to achieve the younger look needed for the character at this stage of his life. An exposure of technique – a violation of the transparency needed to keep the visual effects in the background of the spectator’s awareness – would be fatal to the emotional investment required by the audience at this important moment in the film. But technological developments in face tracking and
volumetric facial capture successfully belie Benjamin’s artificial status.\textsuperscript{538} The visual effects thus received widespread acclaim from critics, including the 2009 Academy Award for Best Visual Effects. Julian Sancton described them in \textit{Vanity Fair} as ‘so perfect as to be virtually invisible, free of the usual trappings of CGI – that too-fluid, superimposed look.’\textsuperscript{539} And Stephen Prince concludes that Benjamin Button is ‘the first film to solve the uncanny valley problem.’\textsuperscript{540} Overcoming the uncanny valley problem and achieving transparency is essential to the film’s ability to produce an emotional immersion in the diegesis. It allows the processes by which a character emerges from the pixels of the cinematic image as an autonomous individual, generating genuine emotional responses in the audience. Without the distraction of visible technique, the spectator can remain focused on the drama involving these two characters separated by tragic life circumstances. The visual effects are an essential part of the film’s aesthetic, underpinning its powerful statement about the effect of time on relationships, and highlighting the painful tragedy inherent in the transience of human life.

Pitt’s performance as Benjamin is a digital composite, combining the actor’s face as captured during principal photography with additional elements introduced during the film’s extensive post-production. The creation of a fully-simulated character capable of activating the embodied simulation mechanisms that facilitate emotional immersion perhaps still eludes the technological capabilities of effects artists and filmmakers, as suggested by the predominantly negative response to characters such as Emily, or Tarkin or Leia from \textit{Rogue One}. But the successful visual effects work in \textit{Benjamin Button} proves that digital augmentation of human performances is not a barrier to emotional immersion in a cinematic world. It shows that digitally augmented characters can conform to the transparency mandated by the classical style. The film is therefore a challenge to the view that the human figure is central to emotional engagement in

\textsuperscript{538} In an interview with FXGuide, Eric Barba, the film’s visual effects supervisor, describes the system used to capture Pitt’s facial expressions: ‘The Mova Contour Capture rig is designed to hold 28 cameras in an array around the actor. They are mounted on a speed rail-like structure that surrounds about 150 degrees of the actor… The cameras are all aimed at the actor’s face which is covered with phosphorescent make-up. This allows for frame by frame tracking of patterns and each point can be tracked in 3D space. This is the first system to truly capture someone’s face moving in realtime and provide a moving mesh that can be subdivided and rebuilt and then retargeted to another mesh to drive a CGI performance.’ Eric Barba cited in Mike Seymour, ‘The Curious Case of Aging Visual Effects,’ \textit{FXGuide}, January 1, 2009, https://www.fxsociety.com/featured/the_curious_case_of_aging_visual_effects/.


\textsuperscript{540} Prince, \textit{Digital}, 136.
the cinema, and offers further evidence of the extension of classical realist aesthetics through cinema’s transition from a photographic basis and into its contemporary digital form.\(^{541}\)

Tom Gunning, who has consistently challenged the primacy of the photographic in traditional theoretical accounts of cinematic realism, suggests that the increasing application of digital visual effects technology to the simulation of the human form further problematizes the neat distinction often drawn between a ‘realistic’ photographic cinema and a ‘synthetic’ digital one.\(^{542}\) In an essay on digital effects and synthetic characters, Gunning discusses simulation in relation to Bazin’s realist phenomenology, referring explicitly to Bazin’s analysis of Susan’s suicide attempt in *Citizen Kane*, an analysis Gunning describes as ‘perhaps his most extensive discussion of the effect of deep-focus photography and its ability to preserve the dramatic unity of space and time.’\(^{543}\) But the famous shot is of course a composite, and Bazin’s embrace of it as evidence of cinema’s ability to create an impression of reality destroys, for Gunning, ‘a naïve identification of the dramatic effect he is describing with the ‘realism’ of the photographic.’\(^{544}\) A similar dramatic effect should not be excluded from our understanding of relationships with simulated characters. ‘We might wonder whether a sense of presence needs to depend entirely on the photographic,’ Gunning writes, ‘when cinema has always had a larger bag of tricks than a putative photographic ontology to convince us that we are watching living beings.’\(^{545}\) The powerful emotional response to *Benjamin Button* suggests that as the uncanny valley continues to be traversed, digitally-augmented characters will increasingly be an option for filmmakers seeking to elicit the kind of immersed engagement with diegetic characters that characterizes a classical realist aesthetic. Hollywood storytelling is likely to continue embracing the benefits of digital technology, with emotional responses to diegetic characters remaining central to the aesthetics of popular cinema.

\(^{541}\) I would also argue that the film challenges the view that there is something inherent in the human form that eludes digital representation, a view Lisa Bode describes as the idea that ‘actors could never be replicated by a machine because of some vague unquantifiable human “magic” that would always elude technology.’ Bode, *Making*, 9.


\(^{543}\) Gunning, ‘Special Effects,’ 347.

\(^{544}\) Gunning, ‘Special Effects,’ 347.

\(^{545}\) Gunning, ‘Special Effects,’ 347.
Conclusion – The Future of Digital Immersion

This thesis has been an attempt to conceptualize the powerful sense of immersion that so often accompanies our encounters with cinematic worlds. This is a mode of engagement characterized by the displacement of the real world of everyday experience by a virtual world, an engagement constructed and maintained through the formal elements of cinema. It is a conscious experience that results from a heightened attention towards the contents of a diegetic world, a diminished awareness of the real world, and a deep sense of involvement with the characters and events being depicted on screen. Cinema’s construction of an illusion of reality is essentially, as Hoberman points out in his review of Gravity, ‘an act of technological prestidigitation,’ and my discussion has therefore focused on the technologies of visual illusion that are used to achieve this displacement. I’ve divided my discussion into three chapters, each exploring a different formal element of cinema and the type of immersive response with which it is most readily associated. In chapter 1 I examined compositing as a means of constructing the virtual spaces that facilitate the spectator’s spatial immersion in the world depicted through the frame. In chapter 2 I examined the use of camera movement to move the spectator through cinema’s virtual spaces, activating spatial depth and extending duration, facilitating a spatio-temporal immersion in the diegetic world. And in chapter 3 I examined the ability of fictional characters to encourage emotional immersion by providing a point-of-view through which a spectator can experience a diegetic world. In each of these three cases, I attempted to show the ways in which digital visual effects have expanded the immersive capabilities of cinema.

I’ve positioned my discussion within a Bazinian realist phenomenology, arguing for the continued relevance of Bazinian aesthetics to the cinema of the digital era. In The Evolution of the Language of Cinema, Bazin describes the addition of sound in the late 1920s as a step forward for the cinematic tradition exemplified by directors such as Erich von Stroheim, F. W. Murnau, and Robert Flaherty. These are directors that Bazin believed avoided the expressionism inherent in montage, and who instead ‘put their faith in reality.’ For the filmmakers working within this tradition, sound was an additional element with which to capture the reality they

---

546 Hoberman, ‘Drowning.’
sought to portray on screen. Bazin thus saw the addition of sound as consistent with cinema’s other technological innovations in fulfilling the medium’s enduring myth, the ‘reconstruction of a perfect illusion of the outside world in sound, color, and relief.’ I’ve argued that the digital visual effects of modern cinema can be understood within the same aesthetic paradigm, offering an expansion of the means with which a perceptually realistic depiction of a diegetic world can be brought to the screen. The case studies I have examined generally use the imaging capabilities of a digital cinema to hide the presence of technique, but there are exceptions to this – the commercial incentives driving Hollywood cinema encourage studios to draw attention to their technological innovations in displays of spectacular excess. But Bazinian realism remains strong as an aesthetic principle, and contemporary filmmakers continue to use digital visual effects to create experiences in the cinema in line with a classical realist aesthetic.

This is an argument for digital visual effects as an expansion of the medium’s aesthetic horizons, broadening the representational tools available to filmmakers and providing spectators with new and sometimes challenging ways of encountering diegetic worlds. In this sense, I may be arguing against a critical and popular orthodoxy that views immersive spectatorship with suspicion. The perfection of illusionist representation, a goal towards which I have argued popular cinema aims, is anathema to the critics and theorists that have long warned of the dangers of the displacement of the real through simulation. In my introduction, I referred to Aldous Huxley’s dystopian vision of a future medium so overwhelming in its illusionary power that it dulls the ability for critical engagement and rational thought. A similar medium is described in Ray Bradbury’s *Fahrenheit 451* (1953), another 20th Century warning of the dangers of increasingly immersive media. In this novel’s dystopian future, books have been banned and replaced by large wall-sized screens that surround the viewer, a medium that hides the banality of its story and characterization with an overwhelming display of sensory excess. Like other dystopian visions of illusionist media, it links immersive spectatorship with diminished critical thinking and independence of thought.

Similarly, cinematic depictions of virtual reality have conventionally treated the technology as a problematic distortion of the line separating the real from the virtual. Films such

as The Lawnmower Man (1992), The Matrix (1999), eXistenZ (1999), and Total Recall (1990 and 2012), all explicitly link illusionism, simulation, and immersion with oppression, conformity, and a loss of independence and autonomy. And this anxiety surrounding immersive media is not confined to popular culture. It underlies Baudrillard’s pessimistic account of the media outlined in his theory of the ‘precession of the simulacra,’ and it features prominently in the ideological critiques of popular cinema put forward by apparatus theory.550

This vision of an overwhelming and oppressive technological apparatus reaches back to cinema’s emergence as a medium of mass spectatorship, with the embellished account of terrified spectators reacting to the projection of an oncoming train providing a potent image of cinema’s illusionary power. As Tom Gunning points out in An Aesthetic of Astonishment, an essay on early cinematic illusionism and spectatorship:

[T]his primal scene at the cinema underpins certain contemporary theorisations of spectatorship. The terrorised spectator of the Grand Café still stalks the imagination of film theorists who envision audiences submitting passively to an all-dominating apparatus, hypnotised and transfixed by its illusionist power.551

In seeking to account for the ideological implications of immersive spectatorship these critics may be exaggerating the power of immersive media, while simultaneously misjudging the spectator’s capacity for rational and critical thought. However, it is not my intention to engage in this debate here. Gunning was writing in 1989, and the ‘contemporary theorisations’ to which he was referring have been thoroughly critiqued elsewhere.552 Instead, I want to conclude with a look towards the future of immersive aesthetics.

Two of the most compelling visions of the future of immersive narrative come from Marie-Laure Ryan and Janet Murray. In *Narrative as Virtual Reality*, perhaps the most comprehensive exploration of the phenomenon of immersion, Ryan argues that immersion in narrative worlds can be profoundly rewarding:

> At its best, immersion can be an adventurous and invigorating experience comparable to taking a swim in a cool ocean with powerful surf. The environment at first appears hostile, you enter it reluctantly, but once you get wet and entrust your body to the waves, you never want to leave.\(^{553}\)

Many spectators of classical narrative cinema, as well as readers of realist novels and enthusiastic gamers, will readily testify to the psychological benefits of an immersive experience involving their medium of choice. But for Ryan, the height of aesthetic experience involves combining the pleasure of immersion with what she calls the ‘ludic immersion’ associated with interactivity. Ryan defines ludic immersion as ‘a deep absorption in the performance of a task, comparable to the intensity with which a mathematician concentrates on proving a theorem or a soloist performs a concerto.’\(^{554}\) Such interactivity ostensibly conflicts with the removal of the spectator that is mandated by classical narrative aesthetics, but Ryan looks forward to a hypothetical future art form that combines the pleasures of immersive narrative with the pleasures of ludic interactivity. She calls this hypothetical form ‘total art,’ and concludes, writing in 2015, that ‘[i]f immersive-interactive narrative is a mountain to climb, we have gone a long way toward the top in the past thirty years.’\(^{555}\)

Murray offers a similar vision of the future of immersive media in *Hamlet on the Holodeck*. She too imagines a future storytelling medium that incorporates a degree of interactivity. Like Ryan’s vision of ‘total art,’ in Murray’s hypothetical medium the invisible spectator of traditional narrative becomes an active participant. She offers a vision of what may be possible in such a medium:

\(^{553}\) Ryan, 17  
\(^{554}\) Ryan, *Virtual Reality*, 247.  
\(^{555}\) Ryan, *Virtual Reality*, 259.
Cyberdramatists of the future could present us with a complex world of many characters (like a global Victorian novel) and allow us to change positions at any moment in order to see the same event from the viewpoint of another character. Or they could let us enter a particular town over and over again in the guise of many different individuals, enabling us to see how differently the same people present themselves to us. We might be given a compelling role within the environment that confers upon us the ability to fluidly switch between viewing the world through our own character’s eyes and viewing our character through the eyes of others.\textsuperscript{556}

Murray sees digital technology as central to this future medium, and suggests that ‘the computer looks more each day like the movie camera of the 1890s: a truly revolutionary invention humankind is just on the verge of putting to use as a spellbinding storyteller.’\textsuperscript{557}

Murray was writing in 1997, and we may still be some way from incorporating into narrative the kind of interactivity she envisions. But this should not detract from the aesthetic opportunities that are currently offered by digital technology, and the enormous expansion such technology represents for the storytelling abilities of filmmakers. I have explored three films that I argue exemplify the advantages of digital cinema – \textit{Life of Pi}, \textit{Gravity}, and \textit{The Curious Case of Benjamin Button} – each of which would not have been possible without the visualization capabilities of digital technology. The further incorporation of such technology will no doubt continue to expand the aesthetic limits of the medium.

\textsuperscript{556} Murray, \textit{Hamlet}, 261.
\textsuperscript{557} Murray, \textit{Hamlet}, 9.
References


Burns, Samuel D. ‘To Catch a Thief.’ Motion Picture Daily, July 14, 1955.


Coplan, Amy. ‘Catching Characters’ Emotions: Emotional Contagion Responses to Narrative Fiction Film.’ Film Studies 8, (Summer 2006): 26-38.


http://escholarship.org/uc/item/3jg726c2.


Johansson, Gunnar. ‘Visual Perception of Biological Motion and a Model for its Analysis.’ *Perception & Psychophysics* 14, no. 3 (1973): 201-211.


Lee, Nora. ‘Motion Control.’ American Cinematographer 64, no. 5, (May 1983): 60-61.


