FACIAL EXPRESSION, EMOTIONAL TENSION, AND FILM

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Kacie Lynn Armstrong

May 2020

FACIAL EXPRESSION, EMOTIONAL TENSION, AND FILM

Kacie Lynn Armstrong, Ph.D.

Cornell University 2020

Discussions surrounding the emotional impact of film have largely been the domain of film theory and criticism. And while psychologists have ventured into the empirical study of cinematically-induced emotion, such work has traditionally focused on the narrative elements of film, most often in the context of isolated film clips. The current work examines the influence of low-level cinematic structure on emotional response among viewers in the context of the whole film.

Specifically, this dissertation investigates the emotional impact of clutter, luminance, motion, shot density, shot scale, and sound amplitude across four short films. These low-level metrics are objectively quantifiable and provide a backdrop against which to analyze the time course of emotional response as a film unfolds. Emotional response is captured via three distinct measures of emotional tension—joystick displacement as a subjective measure, and heart rate and electrodermal activity as objective measures.

An analysis of 40 viewers reveals, first and foremost, that emotional tension is synchronized within all three emotion measures, demonstrating that people respond in remarkably similar ways to cinematic narratives. The low-level structure of film at least partially explains this synchrony, as all six cinematic variables predicted emotional tension in at least half of the short films utilized in this study. An omnibus analysis reveals that film structure exerted

the most robust effect on subjective emotional tension, with all variables but clutter explaining a substantial portion of variance in joystick displacement. Subjective emotional tension was positively predicted by motion, shot density, shot scale, and sound amplitude, and negatively predicted by luminance.

Given that these findings are limited to films featuring White characters, and that group-based perceptual biases might disrupt emotional engagement with more diverse characters on screen, a second study was carried out to enable expansion of this work on a corpus of racially diverse facial stimuli. A collection of 2048 images was digitally rendered and validated for racial representativeness across four geographical regions.

Future directions and social implications of this work are discussed, with a focus on applications to intercultural communication and museum science, as well as the translation of the current findings to broader contexts.

BIOGRAPHICAL SKETCH

Kacie Armstrong was born in Rockville, Maryland and raised in Frederick, Maryland. She attended Frederick High School, where she graduated as valedictorian of her class. Kacie then attended the University of Maryland, Baltimore County, where she earned a Bachelor of Science degree in Psychology (under the mentorship of Anne Brodsky) and a Bachelor of Arts degree in Media and Communication Studies (under the mentorship of Jason Loviglio). She then earned a Master of Arts degree in Film and Visual Studies from Queen's University Belfast, where she explored the mental health of refugees through the art of filmmaking. Her work in documentary and narrative film led to her being named a finalist for the British Council's International Student of the Year Awards in 2010.

In the fall of 2014, Kacie began graduate study at Cornell University, where she worked with James Cutting to study how the visual structure of film drives emotional engagement among viewers, with a special focus on the role of faces. She also worked with Katherine Kinzler to investigate tools to mitigate racial biases in face perception among children. She is currently collaborating with Catalina Iricinschi at the University of the Arts to launch a project investigating cross-cultural differences in the experience of nostalgia. Presently, Kacie works as an Affiliate Professor of Psychology at Regis University and the Metropolitan State University of Denver. Following graduation, she plans to continue her work in academia, with a focus on teaching, writing, and advocacy.

For my husband, Stephen. Thank you for your ongoing love and unwavering support.

ACKNOWLEDGMENTS

First and foremost, I would like to express my gratitude to my advisor, James Cutting, for his incredible mentorship. Throughout my years at Cornell, James led me with an admirable balance of patience, encouragement, and trust. He pushed me to learn and achieve more than I thought was possible, and he supported my every endeavor. I owe much of my academic growth to James, and I will always be grateful for his guidance.

I am also grateful for the invaluable mentorship provided by my committee. I want to thank Katherine Kinzler for the opportunity to be involved in her lab, and for her support as I navigate the job market. I am grateful to David Pizarro for his insight into my work surrounding emotion—particularly for the enlightening conversations regarding nostalgia. Finally, I would like to thank Alexander Ophir for his willingness to both encourage me and challenge me to consider my work from different perspectives.

In addition, I owe many thanks to the rest of the Department of Psychology faculty at Cornell, with special thanks to Michael Goldstein for providing me with outreach experience and career guidance. I am also indebted to the departmental staff for their support, kindness, and encouragement; to my research assistants, Sosna Gellaw, Noah Schweizer, and Sally Liu, for study support; to Olivia Wen and S. K. for assistance with stimuli creation; and to Françoise Vermeylan for statistical consultation.

I would also like to thank Derina Samuel for her teaching mentorship. Derina has provided countless hours of training and guidance, and she has gone beyond the call of duty to ensure that I am equipped to pursue every opportunity presented to me.

I would not have made it to this point without the enduring friendships that I gained during my time at Cornell. In particular, my success would not have been possible without the support of Ayse, Carmen, Catalina, Gandalf, Isobel, Jesse, Jessica, Julia, Katerina, Kedar, both Laurens, Melissa, Olivia, Radhika, and Rajen. It has been wonderful to work with such a brilliant and fun group of people. In addition, I want to thank Lisa, Rafael, Mike, Ritza, Charlotte, Ben, George, Niko, Aubrey, and Brad for the years of friendship, for providing me an outlet beyond Cornell’s walls, and for making our little haunted house on Cayuga Street a true home.

Finally, I am infinitely grateful for my family, particularly my husband Stephen, who has continually inspired me, amused me, and challenged me over the past ten years; my parents, Theresa and Carl, who, from very little, gave me everything; my mother-in-law, Linda, and my father-in-law, Leslie, who treat me as their own and who provide endless love and support from across the pond; my brother, Carl, who is nothing less than my hero; my sisters-in-law, Katrina and Amy, who provide love and laughter when I need it most; my niece, Hazel, who is three years old and smarter than me; and my Aunt Margaret and Uncle Vance, who have served as my surrogate grandparents, and without whom I would not have made it this far. And last but not least, I would like to thank Louis for keeping my lap warm through it all.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Faces are ubiquitous. Most of us see dozens of them on a daily basis, and the average person can recognize approximately 5,000 (Jenkins, Dowsett, & Burton, 2018). It is through faces that we detect cues to a person's identity, attributes, and emotional state. And we are programmed to do so. From an evolutionary standpoint, the ability to automatically process faces allows us to identity friends and foes, predict an individual's next move, and communicate effectively with social partners (see Bruce & Young, 2012).

This nativist account is bolstered by neural circuitry known to support face processing early in life. A network of brain regions selectively responsive to faces includes the occipital face area, which is involved in early perceptual analysis of facial features and branches off into two neural pathways, one projecting to the fusiform face area, where unique identity is processed, and the other projecting to the posterior superior temporal sulcus, where changeable aspects of an individual's face are processed, including gaze, emotional expression, and lip movement (Haxby, Hoffman, & Gobbini, 2000; Kanwisher, 1997). Collectively, this core network supports automatic processing of facial stimuli to facilitate the identification of others and the interpretation of social cues, including facial expressions of emotion.

While the cognitive architecture supporting face processing is widely thought to be evolutionarily endowed, its specific programming depends on environmental exposure over the course of one's lifetime. To explain the human ability to quickly and accurately identify the myriad faces one encounters on a daily basis, Valentine (1991) proposed that faces are stored in a multidimensional cognitive *face space*, with each dimension representing a perceptually relevant facial feature. Facial stimuli are then matched to nodes within the space representing previously observed faces, and facial recognition is facilitated when a percept strongly activates one node at the exclusion of others. Critically, the face space shifts and changes based on exposure to particular faces, thereby allowing one to adapt to their current environment. In addition to Valentine's notion of the face space, holistic processing is also thought to facilitate individuation of faces—which, at their core, are incredibly similar, consisting of the same features in the same general configuration. A multitude of evidence suggests that faces are processed as "gestalts" (rather than in a part-based manner), overcoming the issue of similarity by appealing to the specific geometry of facial features (for a review, see Richler & Gauthier, 2014).

In addition to individual recognition, the processing of facial gestures plays a prominent role in social interaction—and their study enjoys a long and rich history. Darwin proposed in 1872 that facial expressions are adaptive responses to the environment and are therefore universally recognizable. In visits to preliterate cultures in New Guinea, Ekman (1971) sought to verify these claims by recording facial expressions produced in response to a range of story vignettes and by testing recognition of facial expressions depicted in photographic images. His data revealed reasonably strong universal recognition for six basic emotion categories, including anger, disgust, fear, happiness, sadness, and surprise. Ekman's account has not gone unchallenged, however. Most recently, a meta-analysis by Barrett et al. (2019) revealed that a categorical approach to face perception fails to adequately account for cultural differences in the production and interpretation of facial expressions and the various contextual cues that influence the reading of facial gestures.

Dimensional models of facial expression processing may better capture this nuance. Several of these models depict emotional expressions as existing within a two dimensional space, comprised of axes representing, for example, *pleasant/unpleasant* and *attention/rejection*

(Woodworth & Schlosberg, 1960) or *valence* and *arousal*, as in Russell's (1980) circumplex model. Given these parameters, the discrete emotional category of a given facial expression cannot be appropriately identified without contextual cues. For instance, a facial gesture that is high in arousal and low in valence could be identified as either anger or fear, depending on the events that incited the facial gesture, knowledge of the individual's traits and dispositions, and cultural accents that imbue facial expressions across geographical regions (e.g., Crivelli et al., 2016).

The debate over the nature of facial expression processing is particularly relevant to film. Indeed, given the importance of faces in everyday life, it is not surprising that faces feature prominently in film, and that film viewers tend to gaze primarily at the faces of characters on screen (Smith & Henderson, 2008), which most often occur in conversation (Cutting, 2016a). And as films are largely built around the emotional responses of characters to various situations (Plantinga, 2009; Tan, 1996), these conversations are rife with facial expressions of emotion (see Cutting & Armstrong, 2018).

Critically, the way faces are framed on screen influences how they are perceived by viewers. Framing is determined by shot scale, where the closer the shot, the larger the character's face on screen, and the longer the shot, the smaller the face on screen. A collective analysis of nearly 300 films reveals that, over the last century, mean shot scale has shifted from approximately a medium long shot (which cuts a character off at the knees) to a medium close-up shot (which cuts a character off at the chest) (Cutting & Armstrong, 2018; Salt, 1992, 2006, n.d.). The latter reduces the variability in gaze fixation across viewers, as they are able to simultaneously gather information from the eyes and the mouth of a character in a single glance

(Smith, 2012, 2013). In other words, implementation of medium close-up shots minimizes the effort required of the viewer to read emotional information from a character's face.

And this makes sense, considering the vast evidence for holistic face processing, which extends to facial expressions of emotion across both static and dynamic stimuli (e.g., Tanaka et al., 2011; Favelle, et al., 2015). Thus, given the applications of holistic processing to shot scale in film, what happens to the reading of facial expressions in comparatively longer and closer shots?

In terms of the former, non-facial cues could be used to guide emotion perception. Indeed, Aviezer, Trope, and Todorov (2012) found that, in moments of peak emotional intensity, perceived emotional valence aligned with body posture over and above facial gestures (and the converse is true as well, such that facial context can alter the categorization of bodily emotion; Lecker et al., 2019). However, given that longer shots (those which reveal the entire body of a character) typically open and close scenes in Hollywood film (Bordwell, 2006; Cutting & Iricinschi, 2015; Smith, Levin, & Cutting, 2012)—when the emotional intensity of characters is not likely to be at its pinnacle—body posture may not be a robust cue to emotional state. But the surrounding environment may also provide cues. Indeed, filmmakers are known to manipulate the scenery of a shot to drive a particular mood and to illustrate a character's state of mind (Barsam & Monahan, 2010). Building on this theory, Smith (2003) proposed that the visual structure of film predisposes viewers to interpret cinematic narratives through a prescribed emotional lens.

Empirical evidence supports these theories. Several studies have revealed mood congruency in facial expression processing, such that an observer's present emotional state heightens their sensitivity to facial stimuli conveying corresponding emotions, as well as their tendency to perceive mood-congruent emotion in neutral or ambiguous stimuli (e.g., Niedenthal et al., 2000; Voelkle et al., 2014). In the realm of film, mood congruency may at least partially explain a phenomenon known as the Kuleshov Effect. In the early twentieth century, Soviet filmmaker Lev Kuleshov edited together a sequence of shots which depicted an expressionless man alternated with images of a bowl of soup, a child in a coffin, and a woman lying suggestively on a sofa. Viewers perceived the man's facial expression as hungry, grief-stricken, and lustful, respectively—suggesting that the preceding shots led viewers to infer emotion that reflected the mood of the scene (Kuleshov, 1974). The striking power of context may point to ways in which film viewers perceive emotion in extreme closeup shots as well. Such shots typically cut off a character's mouth or eyebrows, potentially disrupting holistic processing of individual identity as well as facial expressions of emotion. Indeed, the mouth appears to be the most salient feature involved in the processing of happy expressions (e.g., Calvo et al., 2014), while the eyebrows play a prominent role in the processing of sad expressions (Beaudry et al., 2013). Without these direct facial cues, viewers likely rely on narrative context to interpret a character's emotional state.

In sum, context arguably influences emotion perception on the cinema screen just as it does in the real world. However, a few caveats must be noted. The first involves transparency. It is not always the goal of the filmmaker to make a character's emotions known. For instance, reaction shots—which show a character reacting silently to an event—typically depict facial gestures that reflect a slightly negative and slightly aroused emotional state that is difficult to categorize, thus leaving viewers in a state of speculation about what the character is thinking (Cutting & Armstrong, 2018). Thus, by obscuring a character's emotions through the carefully crafted reaction shot, filmmakers enhance viewers' emotional involvement with characters on screen.

Second, cultural differences may influence contextual effects on face perception in film. For instance, Hall (1976) outlined a distinction between low-context and high-context cultures, proposing that individuals in low-context cultures rely more on the explicit wording of messages (due to being embedded in individualistic communities and therefore less familiar with the backgrounds and habits of their social partners), while individuals in high-context cultures focus more on subtle, unspoken cues such as gestures and underlying meaning (due to being embedded in collectivist communities defined by intimate interpersonal relationships and a focus on social harmony). This model suggests that individuals across cultures may be differentially sensitive to the manipulation of shot scale in film, which can expand or constrict the viewer's access to a character's emotional gestures.

Finally, shot scale is not the only visual tool of cinema that influences emotion perception among viewers. For instance, luminance is known to affect the visual mood of a scene, such that darker scenes are perceived as more negative in valence and brighter scenes are perceived as more positive in valence (Tarvainen, Westman, & Oittinen, 2015), and filmmakers put this dichotomy to practice by aligning luminance with narrative structure (Cutting, 2016a, 2016b). Additionally, motion influences the perceived intensity of emotion. Paterson, Pollick, and Sanford (2001) found that when the speed of angry bodily movements performed by point light figures was manipulated on screen, viewers perceived faster movements as more emotionally intense than slower movements. The opposite pattern was observed for sad bodily movements. Thus, by manipulating motion, filmmakers might orient viewers to perceive a character's level of emotional arousal in a particular way.

In other cases, the visual tools of cinema might interfere with emotion perception. Cutting & Armstrong (2016) found that increased clutter (i.e., the amount of "stuff" in a scene) impedes the processing of facial expression valence in longer shots (but not closer shots). However, filmmakers seem to hold implicit knowledge of this effect, as they accommodate longer visual processing times by extending the duration of longer shots that contain a relatively high degree of clutter.

In sum, filmmakers appear to influence facial expression processing (and emotion perception more generally) by manipulating low-level contextual information (e.g., through luminance, shot scale, and shot duration)—but sometimes, they obstruct emotion perception by withholding visual cues (e.g., through the reaction shot). Though these strategies may seem contradictory in nature, they serve a single purpose: to promote emotional engagement with cinematic narratives as they unfold on screen, thus cultivating empathy for a film's characters.

The science underlying face perception feeds directly into cinematic engagement in a number of ways. For example, perhaps one of the most well-documented mechanisms by which empathy arises is through facial expression mimicry, which occurs via rapid facial reactions (RFRs), in which individuals quickly and briefly mirror the facial musculature of a social partner (Dimberg, 1982; Dimberg & Thunberg, 1998; Dimberg, Thunberg, & Elmehed, 2000). This process results in afferent feedback projected from the facial muscles to the right somatosensory cortex (which represents bodily states associated with distinct emotions), providing the mimicker with cues to the other's internal emotional state, ultimately leading to a shared subjective emotional experience (Adelmann & Zajonc, 1989; Kragel & LaBar, 2016; McIntosh, 1996, 2006; Niedenthal, Mermillod, Maringer, & Hess, 2010). In applying facial mimicry to film, Plantinga (2009) described how filmmakers focus viewer attention on characters' facial expressions through increasingly closer shots as a scene progresses, thus facilitating embodied simulation of a character's emotions.

Central to the discussion of facial mimicry is the discovery of mirror neurons, described as motor neurons that are activated during both the observation and the performance of a particular action (e.g., Likowski et al., 2012; Rizzolatti & Craighero, 2004). Mirror neurons have been directly recorded in the frontal cortex of macaque monkeys, and in humans, neural activity consistent with mirror neurons has been observed in the premotor cortex, supplementary motor area, primary somatosensory cortex, and inferior parietal cortex (Keysers & Gazzola, 2010; Molenberghs, Cunnington, & Mattingley, 2009; Rizzolatti & Craighero, 2004). The mirror system is thought to represent a mechanism for action understanding, imitation, and empathy that requires no cognitive mediation (Iacoboni, 2009; Kaplan & Iacoboni, 2006; Rizzolatti & Craighero, 2004).

However, there has been much debate among scientists regarding both the existence of mirror neurons in humans, as well as their role in empathic understanding of another's emotional state (e.g., Hickok, 2009; Kilner & Lemon, 2013). Specifically, questions have surfaced regarding whether mirror neurons *serve* emotional understanding or merely *reflect* such understanding (Lamm & Majdandžić, 2015). Related to this line of questioning, Moody et al. (2007) presented EMG data revealing that RFRs do not indicate motor mimicry per se, but rather, may be the product of one's prior emotional state, as participants who were induced into a fearful mood showed increased fearful RFRs to both fearful *and* angry faces. These findings suggest that emotional experience may be a cause rather than a consequence of the RFRs that are produced during social interaction. In addition, mirror system activity appears to be modulated by the reward value and overall goal of a specific movement (see Kilner & Lemon, 2013), as well as the perceiver's interpretation of the preceding series of events (Yeshurun et al., 2017). These findings indicate that the mirror system, rather than representing a cognitive shortcut, is an intricate system imbedded in complex neural circuitry.

While the mechanisms underlying face perception are hotly debated, and scientific understanding of their nature is continually evolving (see Bruce & Young, 2012), one thing is certain: filmmakers implicitly rely on the human tendency to automatically attend to and interpret the information gleaned from faces, which dominate cinema screens. Thus, the first question under investigation in this dissertation is the following: How does the framing of faces on screen, in conjunction with a multitude of visual features known to affect emotion processing (including clutter, luminance, motion, and shot duration), drive emotional engagement among film viewers?

This project is intended not only to illuminate the ways in which film is structured in order to optimize emotional engagement, but also to provide insight into human perceptual processes in general. Specifically, given that film is so ubiquitous, the strategies that filmmakers employ in order to captivate audiences may lead to predictions regarding mechanisms underlying face perception in the real world—predictions that might not otherwise be generated.

However, one glaring flaw in Hollywood film limits the applicability of this research to the real world—that is, its historical lack of diversity. From 2016 to 2017, people of color made up just 19.8% of lead roles in Hollywood film (despite representing nearly 40% of the U.S. population), while women made up just 32.9% (despite representing just over half of the U.S. population) (Hunt, Ramón, & Tran, 2019).

Fortunately, the tide is changing. While minority actors continue to be passed over for Academy Award nominations, their representation in Hollywood film is growing. In 2018, 26.6% of lead roles went to people of color and 41.0% to women. This shift is thought to be the result of loud calls for structural change in Hollywood coupled with the fact that films marked by greater minority representation tend to perform better at the box office—likely because diverse casts resonate more deeply with increasingly diverse audiences (Hunt & Ramón, 2020). Despite this positive shift, the corpus of Hollywood films available for psychological study is still heavily weighted toward white characters and ethnocentric narratives. And given that modern society is composed of a rich diversity not reflected in cinema, the extent to which Hollywood film can teach us about face perception in the real world is constrained by its own racial biases.

This deficiency is due in large part to fundamentally different mechanisms underlying ingroup versus outgroup face perception. Although race is widely believed to be a social construction (e.g., Gelman & Hirschfeld, 1999; Sternberg, Grigorenko, & Kidd, 2005), humans generally perceive robust racial categories (e.g., Haslam, Rothschild, & Ernst, 2000) determined by a constellation of facial cues (including skin pigmentation and physiognomy; see Bar-Haim, Saidel, & Yovel, 2009; Dunham, Dotsch, Clark, & Stepanova, 2016; Dunham, Stepanova, Dotsch, & Todorov, 2015; Hill, Bruce, & Akamatsu, 1995). Furthermore, facial processing is qualitatively different across racial categories, as demonstrated by two well-replicated phenomena in face perception research: own- and other-race effects. These effects describe, respectively, the tendency to individuate own-race faces faster and more accurately than faces of other races (for a review, see Meissner & Brigham, 2001), and the tendency to more quickly categorize other-race faces into a single group (Levin, 1996, 2000).

To account for these effects, the *perceptual expertise hypothesis* posits that asymmetrical experience in processing own- vs. other-race faces leads to differential recognition abilities, such that increased exposure to own-race faces fosters greater skill in extracting individuating facial cues (see Meissner & Brigham, 2001). Specifically, in terms of Valentine's (1991) cognitive face space, faces are thought to be spatially distributed according to degree of exposure, such that familiar own-race faces are represented more precisely along multiple dimensions and therefore spaced further apart, while other-race faces are more densely represented along the single dimension of skin color, and therefore more difficult to individuate (Byatt & Rhodes, 2004).

Accordingly, the use of Hollywood film as a medium for the psychological study of face perception is limited in that it does not extrapolate to the various ways in which we process faces across social groups. In order to generalize this work to more diverse populations, the findings that emerge must be replicated on stimuli covering a wider racial spectrum. Thus, the second goal of this dissertation involves the creation and cross-cultural validation of a large corpus of facial stimuli that represent numerous racial categories, gender categories, and emotional expressions. These stimuli were created with two main goals in mind: to provide convincing exemplars of people of color, and to ensure that the racial categories of these exemplars are perceived similarly across different geographical regions. These goals, respectively, map onto the growing diversity of film characters and the growing diversity of film audiences. Therefore, these stimuli can subsequently be utilized in studies that seek to understand how racial diversity translates to emotional engagement with cinematic storylines. For example, one might apply low-level cinematic tools (e.g., luminance, motion, and shot scale) to these stimuli to observe how the variation of such tools affects facial expression processing among viewers.

In sum, the questions investigated in this dissertation are pertinent to both the future of Hollywood (in terms of its efforts to engage viewers through the visual medium of film) and to the use of film as a tool to broadly understand the mechanisms underlying face perception in our increasingly globalized world.

CHAPTER 1: DYNAMIC EMOTIONAL TENSION AS A FUNCTION OF THE LOW-LEVEL STRUCTURE OF FILM

> Fiction is a powerful and ancient virtual reality technology that simulates the big dilemmas of human life. […] We identify so closely with the struggles of the protagonists that we don't just sympathize with them; we strongly empathize with them. We *feel* their happiness and desire and fear; our brains rev up as though what is happening to them is actually happening to us.
>
> -Jonathan Gottschall (2012, p. 67)

In 2018, 1.3 billion movie tickets were sold in the U.S. and Canada (Motion Picture Association of America, 2019), illustrating a widespread captivation with Hollywood film that can arguably be attributed—at least in part—to its emotional appeal. But this raises questions, because Hollywood movies typically tell fictional stories. Viewers are aware that the narratives they see on screen are artificial. Yet, they cry when a beloved character dies, and their hearts race as the protagonist narrowly escapes an attack. How can it be that intelligent humans feel very real emotions in response to imaginary characters written into fictional situations?

This phenomenon—known as the Paradox of Fiction—is rooted in the concept of narrative transportation, which describes viewer immersion in a world evoked by a fictional story (see Gerrig, 1993; Green & Brock, 2002). The mechanisms underlying narrative transportation have been explored in a number of ways, ranging from neuro-psychoanalytic perspectives (e.g., Holland, 2003) to the intersection between art and phenomenology (e.g., Paskow, 2008; Walton, 1993). However, the model that is arguably most relevant to the psychological study of film is that put forth by Busselle and Bilandzic (2009), which distinguishes four interrelated dimensions of narrative engagement:

1) Attentional focus (sustained engagement with the storyline)

2) Narrative understanding (ease in comprehending the narrative)

3) Emotional engagement (feeling for and with the characters)

4) Narrative presence (entering the story world)

This framework posits that consumers of fiction continually construct and update mental models encompassing the characters, locations, and plot points of a given story. As such, they are not only able to keep track of the narrative, but they also trade awareness of their immediate surroundings for the sensation of entering the story world and taking on the perspective of its characters (Busselle & Bilandzic, 2009).

Others have considered the psychological mechanisms underlying narrative engagement in the specific context of film viewing. For instance, building on cognitive theories of emotion put forth by Frijda (1986, 2007) and Oatley (1999, 2012, 2013), Tan (2018) proposed that film viewers engage in *playful simulation*. As a narrative unfolds on screen, the viewer imagines that they are present in the fictional world occupied by its characters. As such, they take on the role of witness, which involves both embodied perception of the action taking place within the fictional world and an imaginary participation in such action. The viewer does not act on these events, but they accept them as real for the sake of enjoyment. Entertaining such events as possibly true leads to genuine emotional responses, as humans are equipped with the capacity to experience palpable emotion in response to mental representations of imaginary events (see Tan, 1996).

While grounded in cognitive models of emotion, playful simulation relies heavily on perceptual input. For instance, Tan (2018) described how the *apparent realism* of film depends on the visual emphasis of emotional triggers (e.g., editing technology can strengthen the abruptness of an enemy's appearance), even if further cognitive processing of the stimulus leads film viewers to conclude that the perceived event is artificial.

Indeed, a vast body of research supports the role of perceptual mechanisms in fostering narrative engagement. This research largely focuses on the low-level visual features of film—defined as the quantitative aspects of visual film structure (Brunick, Cutting, & Delong, 2013), including shot scale (the relative size of a character within the frame of a shot), shot duration (the length of time that a shot lasts, as determined by the number of successive frames in between two cuts), luminance (the amount of light present in an image), clutter (a measure of edge density; see Cutting & Armstrong, 2016; Rosenholtz, Li, & Nakano, 2007), and motion (the change in the intensity of pixels across two adjacent frames; see Cutting, Brunick, DeLong, et al., 2011; Cutting, DeLong, & Brunick, 2011). Investigation of these features provides insight into the ways in which filmmakers capitalize on human perceptual processes to foster various aspects of narrative engagement—particularly attentional focus and narrative understanding, and they seem to have built a likely foundation for emotional engagement as well (see Busselle & Bilandzic, 2009). A review of this evidence follows; however, before considering the psychological impact of the low-level visual structure of film, it is worthwhile to describe its fundamental building block—the cut.

The Cut

A shot is defined as a series of frames, or still photographic images, that runs for an uninterrupted period of time. When these images are presented in succession at a sufficiently rapid frame rate (around 16 frames per second for early silent films and 24 frames per second or higher for contemporary films), viewers perceive fluid motion across images rather than a staccato display of visually distinct pictures. In other words, the “seams” between images disappear when the frequency of frame presentation exceeds the critical flicker fusion rate of the human visual system (approximately 60 Hz; Hecht & Shlaer, 1936).

This illusion of fluid movement is related to the phenomenon of apparent motion: when there is a fleeting period of darkness between two briefly flashed images, the visual system fills in the darkness with the perception of visual stimuli that joins the first image to the second (Wertheimer, 1912; Goldstein, 2013). In other words, when static images are sequentially displayed faster than the eye can see, they create the illusion of a smoothly flowing shot.

Of course, contemporary films are not comprised of a single continuous shot. Over the course of a film, the visual flow is physically interrupted by hundreds to thousands of sharp transitions between shots, known as cuts. The cut kindled a filmmaking revolution by allowing for the precise cutting and merging of various shots to construct a complex sequence of events. To illustrate, the rumored first cut—which is said to have been a fateful accident—occurred when early filmmaker George Méliès ran out of film while recording a streetcar in France in the late 1800s. He reloaded then restarted his camera, only to discover later that two separate events had been spliced together, serendipitously transforming the streetcar into a hearse (Packer, 2007).

Contemporary filmmakers, however, are generally less interested in the illusion of magic and more interested in narrative cohesion. To immerse the viewer in a seamless story world, they have adopted well-orchestrated editing strategies that, paradoxically, revolve around increased use of the cut. Specifically, Hollywood film is marked by a process known as intensified continuity, which describes the use of rapid editing (shorter duration shots with increased motion) and the removal of extraneous narrative details to ensure a seamless visual stream of narrative information (Bordwell 2002, 2006). In other words, through efficient, careful editing, filmmakers create a sense of visual cohesion by hiding the fact that film is comprised of a disjointed sequence of visual material. As such, the sequencing of shots in a film provides a dynamic canvas for the low-level visual features that contribute to narrative engagement.

Attentional Focus

While an optimal frame rate and masking of cuts foster a perceptually smooth visual narrative, these attributes alone do not account for viewer engagement. It is—at least in part—the low-level features of film that capture and sustain the attention of viewers for 90 minutes or longer. Eye tracking studies, which measure saccades and foveal fixations as an index of visual attention, provide valuable insight into this arena (Findlay, 2004). Such studies demonstrate that Hollywood film results in a high degree of attentional synchrony: viewers look at the same things at the same time (e.g. Goldstein, Woods, & Peli, 2007; Smith & Henderson, 2008). This effect is attributed to attentional capture resulting from the pop-out effects of low-level visual features such as color, contrast, motion, and orientation. If the features of a particular region of a scene stand out from the scene as a whole, an involuntary attentional shift occurs, drawing viewer gaze to the salient area of the screen (for a review, see Mital, Smith, Hill, & Henderson, 2011; Wolfe & Horowitz, 2004).

During free viewing of single-shot film clips, gaze location is best predicted by motion, and as motion in a shot increases, so does attentional synchrony. Static features including color, luminance, orientation, and corners also predict gaze location, though to a lesser degree (Carmi & Itti, 2006; Itti, 2005; Mital et al., 2011). In viewing static scenes, fixations tend to cluster at the same locations, but not at the same time (Mannan, Ruddock, & Wooding, 1997), suggesting that the motion that characterizes dynamic stimuli is key to guiding attention in real time. Thus, motion is a feature that allows film, by its very nature, to capitalize on early visual processing to guide the viewer's gaze.

But what happens across shots, where cuts interrupt the visual flow? Based on the evidence discussed above, one might assume an involuntary, bottom-up mechanism of gaze control determined by exogenous cinematic features such as motion (Carmi & Itti, 2006). However, if such a mechanism were responsible for attentional synchrony, this pattern of results would persist across cuts, due to the fact that the salience of bottom-up information should be impervious to narrative structure (unlike top-down information regarding a shot's content, which is impoverished in the beginning as compared to the middle or end of a shot; Mital et al., 2011). However, Mital et al. (2011) found no relationship between exogenous visual features and gaze clustering directly following a cut. Instead, clustering appears to be the result of a general tendency to saccade to the center of the screen (also see Tatler, 2007; Tseng, Carmi, Cameron, Munoz, & Itti, 2009). Furthermore, Loschky, Larson, Magliano, and Smith (2015) found that, among participants who viewed a 12 second *James Bond* film clip, those who were shown the preceding 2.5 minutes of the film showed greater attentional synchrony across the first shot of the 12 second clip than those who were provided with no narrative context. These findings suggest that, in addition to bottom-up features, top-down factors regarding the broader visual narrative play a role in gaze location.

Filmmakers seem to implicitly know this. Mital et al. (2011) described a trend in filmmaking that centers focal objects on screen after a cut—a stylistic choice that aligns with viewers' tendency to saccade to screen center to reorient to scene changes. In other words, filmmakers may have ascertained, though probably not consciously, that disruptions to top-down factors that contribute to motion tracking need to be resolved by reliably guiding the viewer to the center of the screen once one shot has ended and another has begun.

In addition to visually structuring film to optimize gaze control both within and across shots, filmmakers also demonstrate an implicit knowledge of the dynamic rhythm of attentional systems. Over the course of a cognitive task, attention ebbs and flows, resulting in a distinct pattern of variance in reaction time data. These vacillations follow a mathematical power law known as pink noise, or $1/f$ noise, indicating an inverse relationship between reaction time and frequency (Gilden, 2001). This pattern is also referred to as fractal noise due its self-similarity at different scales (see Delong, Brunick, & Cutting, 2013). Fractal structures are seen in a wide variety of natural phenomena—for example, in the spatial layout of coastline segments and in the temporal rhythms of avalanches and earthquakes (Newman, 2013). Likewise, since around 1960, Hollywood film has increasingly adhered to a $1/f$ pattern in terms of shot duration (whereby longer shots are less frequent; Cutting, DeLong, & Nothelfer, 2010), suggesting that, through trial and error over the decades, filmmakers have internalized the rhythmic patterns of film that best harness viewer attention to the narrative as it unfolds on screen (Brunick et al., 2013; Cutting et al., 2010).

Neural evidence provides further support for attentional synchrony among film viewers. Hasson, Nir, Levy, Fuhrmann, & Malach (2004) found that viewing a segment of *The Good, the Bad and the Ugly* (Leone, 1966) evoked highly correlated neural activity within various processing levels of the auditory cortex, the visual cortex, and multisensory and language areas located in the temporal and parietal lobes. Overall, approximately 45% of the cortex was synchronized among viewers and time-locked to the content of the film. In contrast, an unedited video of a concert recorded from a single viewpoint resulted in synchrony across only 5% of the cortex. These findings demonstrate that film has the distinct capacity to induce harmonized responses throughout viewers' brains—a capacity that is not reflected in natural stimuli (also see Hasson, Malach, & Heeger, 2010).

Having established the role of low-level cinematic features in capturing and sustaining attention during film viewing, I will now consider how the visual structure of film contributes to the second dimension of narrative engagement described by Busselle and Bilandzic (2009)—that is, the higher-order cognitive process of narrative understanding—a process that, in film, is largely achieved through event segmentation.

Narrative Understanding

An event is the building block of any given narrative. Defined by Zacks and Tversky (2001, p. 17) as "a segment of time at a given location that is conceived by an observer to have a beginning and an end," an event, by definition, necessitates boundaries. The ability to parse an ongoing narrative into smaller, more manageable events is linked to better memory for narrative content as well as stronger predictions about what will happen in the future (see Richmond, Gold, & Zacks, 2017).

Cinematic events in contemporary film are represented by scenes comprised of an average of seven shots (Cutting, Brunick, & Candan, 2012), and, like theater, are defined by shifts in character, location, and/or time (Polking, 1990; Zacks & Tversky, 2001). Viewers generally agree on where to segment whole films into events, suggesting that filmmakers successfully cue viewers to event boundaries (Cutting, Brunick, & Candan, 2012). Indeed, several low-level visual features explain a substantial amount of segmentation variance; shot scale accounts for 13% and shot duration 9% (Cutting, Brunick, & Candan, 2012), as scenes generally begin and sometimes end with longer-scaled, longer-duration shots in order to provide viewers with optimal visual information (Bordwell, 2006; Cutting & Iricinschi, 2015; Smith, Levin, & Cutting, 2012). Color, explaining 4% of segmentation variance, is manipulated to mark event boundaries in a way that cleverly replicates location and time changes in the real world: urban scenes are often grayish-brown while suburban scenes are green, and night scenes are imbued with hints of blue while day scenes are yellow or orange (Cutting, Brunick, & Candan, 2012). Finally, although motion is not nearly as robust a cue as shot scale and shot duration, it does signal event boundaries in the real world, and to some extent, in film (Magliano & Zacks, 2011; Richmond et al., 2017; Smith, 2012; Speer, Swallow, & Zacks, 2003; Zacks, 2004; Zacks, Speer, Swallow, & Maley, 2010), where scene changes tend to be marked by heightened motion (Cutting, Brunick, & Candan, 2012).

Evidence suggests that event segmentation, like shot duration, follows a fractal pattern, regardless of whether viewers are instructed to attend to small- or large-scale events. Like avalanches and earthquakes, the perceived small-scale events of film seem to have the same $1/f$ temporal structure as the perceived large-scale events (Blau, Petrusz, & Carello, 2013). Collectively, these findings illustrate that filmmakers facilitate viewer comprehension of the narrative by carefully constructing film to appeal to the parsing mechanisms of the human mind. And once again, neural evidence corroborates this notion. Hasson, Furman, Clark, Dudai, & Davachi (2008) revealed numerous brain regions where neural activity is substantially more correlated across viewers during portions of a film that are later successfully (versus unsuccessfully) recalled in terms of their narrative content. These brain regions include the parahippocampal gyrus (associated with episodic encoding) and the superior temporal gyrus (associated with visuomotor integration and narrative comprehension). This widespread synchrony attests to the power of film to engage viewers by triggering a particular pattern of neural processing that promotes narrative encoding.

Given substantial evidence that filmmakers implicitly manipulate low-level visual features of film to harness visual attention and facilitate event segmentation among viewers, it follows that such features might also influence the third component of narrative engagement described by Busselle and Bilandzic (2009)—that is, emotional engagement.

Emotional Engagement

Let us return for a moment to the Paradox of Fiction, which describes film viewers' tendency to feel real emotions in response to fictional scenarios (Paskow, 2008; Walton, 1993). While this paradox pertains to a number of literary formats, filmmakers are privileged to have at their disposal an arsenal of visual tools that drives emotional engagement among viewers. Indeed, previous research indicates that luminance, shot scale, and motion are instrumental in setting the emotional tone of a scene.

The horror genre takes particular advantage of luminance. Films such as *Apollo 18* (López-Gallego, 2011), *The Blair Witch Project* (Myrick & Sanchez, 1999), *The Descent* (Marshall, 2005), and *Buried* (Cortés, 2010) are relatively dark, presumably heightening the viewer's sense of helplessness as they are, in a sense, blinded during the film (Brunick et al., 2013). The sense of helplessness felt by the viewer may mirror the inferred helplessness of the film's fictional victims, thus building an emotional connection between the viewer and the character.

Furthermore, Cutting (2016b, 2016a) demonstrated that the luminance of scenes in Hollywood film aligns with the emotional tone of the narrative: the literal darkest moment of the film occurs at a point in plot development when the protagonist seems least likely to achieve their goal, and the literal brightest moment occurs once that goal has been achieved. These findings suggest that filmmakers use luminance to cue audiences to the emotional content of scenes. Similarly, Smith's (2003) mood-cue approach to film stipulates that cinematic structure (in terms of narrative, visual, and audio content) promotes a broad predisposition among viewers to interpret cinematic events through a specific emotional lens. In other words, by cultivating mood through film structure, filmmakers orient viewers to experience emotions in a prescribed fashion. Thus, perhaps by decreasing luminance at low points in the narrative and increasing it at high points, filmmakers guide viewers to feel helplessness and elation, respectively, in response to a character's losses and victories.

Additionally, as face perception is integral to cultivating empathy with characters, shot scale is a vital ingredient for emotional engagement. Indeed, shot scale is known to facilitate viewer processing of facial expressions on screen. As discussed earlier, a trend toward medium-closeup shots in contemporary Hollywood film enables viewers to more easily extract facial expression cues via holistic processing (see Smith, 2012, 2013), and where clutter interferes with the processing of facial expression valence, filmmakers implicitly extend shot duration to account for longer visual processing times (Cutting & Armstrong, 2016). Thus, filmmakers seem to manipulate the size of a character's face on screen (in conjunction with shot duration) in order to drive emotional engagement through the human tendency to automatically attend to and interpret facial stimuli.

Finally, the degree of motion in a film exerts a strong influence on emotional engagement in three distinct ways. First, like shot scale, motion also appears to be beneficial to facial expression processing (e.g., Ambadar, Schooler, & Cohn, 2005). In an fMRI study, Trautmann, Fehr, and Herrmann (2009) found that, among women, exposure to dynamic vs. static stimuli depicting neutral, happy, and disgusted expressions enhanced neural activity in brain regions associated with emotional memory encoding and emotional arousal. These findings suggest that film, by its dynamic nature, is a medium rich in opportunity for emotional engagement with characters as they outwardly process the cinematic world in which they are embedded.

Second, as previously discussed, the speed of bodily movements performed by point light figures on screen influences the perceived intensity of emotion, such that faster angry movements are perceived as more emotionally intense that slower angry movements, while slower sad movements are perceived as more emotionally intense than faster sad movements (Paterson et al., 2001). In the realm of film, these findings apply particularly well to the jump cut—an editing technique in which two sequential shots of the same action are recorded from camera positions that vary only slightly (if at all), giving the impression of jumping slightly ahead in time. Hogan (2007) proposed that, by altering the temporal aspect of motion, jump cuts may serve to increase the density of gestures, and consequently, their emotional impact.

Third, motion induces positive affect. Pronin and Jacobs (2008) found that participants who viewed film clips with a rapid shot pace reported a more positive mood than those who viewed slower paced clips. Indeed, shot duration has decreased over time due to conventions of continuity editing, resulting in a quicker cinematic pace in Hollywood film (Delong et al., 2013). Pronin and Wegner (2006) attributed the resulting positive affect among contemporary film viewers to quicker speed of thought. Likewise, Grodal (2009) proposed that the rapid activity of contemporary films may be linked to dopamine release in the viewer, which is associated with reward and positive affect (e.g., Fields & Margolis, 2015; Wise & Rompre, 1989).

It is important to note that the evidence presented here is largely relevant to viewers' *perception* of emotion portrayed through film rather than *felt* emotional experience within viewers themselves. While these findings do suggest ways in which filmmakers might manipulate the visual structure of film to drive emotional comprehension of cinematic narratives, any conclusions regarding emotional engagement remain speculative—particularly given that emotional engagement, by definition, involves empathic *feeling* for and with characters (Busselle & Bilandzic, 2009; Gerrig, 1993; Green & Sestir, 2017; van Laer, de Ruyter, Visconti, & Wetzels, 2014).

While a robust empirical link between the visual structure of film and true emotional response among viewers has yet to be established, music is a domain of cinema that enjoys greater empirical support in this realm. Indeed, music is known to incite physiological change (e.g., Kelley, Andrick, Benzenbower, & Devia, 2014)—particularly when perceived as emotionally powerful (Rickard, 2004). Additionally, when paired with film, music appears to have an additive effect on emotion, such that more positive music results in higher valence ratings of film clips (Ellis & Simons, 2005). The emotional impact of music in film is likely due to its contribution to the mood of a scene, priming effects related to the perception of characters' facial expressions (Tan, Spackman, & Bezdek, 2007), and the foreshadowing of narrative events (Boltz, Schulkind, & Kantra, 1991; Magliano, Dijkstra, & Zwaan, 1996).

However, music is not the only palpable source of sound in film. Conversations between characters occur in 72% of all shots in dramas, 60% of shots in comedies, and 43% of shots in action films (Cutting, 2016a), thus lending another source of auditory stimuli that may influence emotional response among film viewers. Given that listeners perceive louder music as more emotionally arousing (Dean, Bailes, & Schubert, 2011), the overall sound amplitude of a film—

including both non-diagetic and diagetic sound—is yet another low-level feature of film worth investigating in terms of its emotion-eliciting properties.

Regardless of sensory modality, the extant work surrounding the low-level structure of film does not include an investigation of the time course of emotional engagement in the context of complete cinematic narratives. Given the dynamic nature of emotion, this piece of the puzzle is invaluable to understanding the relationship between film structure and cinematically-induced emotion.

The Dynamic Nature of Emotional Experience

As described above, filmmakers are known to utilize a range of low-level features to drive narrative engagement through perceptual, cognitive, and emotional mechanisms. They use motion and shot duration to capture and sustain attentional focus; they use shot scale, shot duration, color, and—to a lesser extent—motion to drive event segmentation in service of narrative comprehension; and they use luminance, motion, and shot scale (as well as the interaction between shot scale and clutter) to, respectively, set the emotional tone of a scene, influence perceived emotional intensity of characters' actions, and facilitate processing of characters' facial expressions—all of which are mechanisms by which filmmakers may drive emotional engagement by facilitating emotional comprehension of the narrative.

However, previous work investigating the relationship between the low-level visual structure of film and emotional engagement among viewers is largely based on retrospective emotion ratings, or limited to brief instances of film viewing confined to a single shot or isolated film clip. While the latter provides a rich foundation of knowledge surrounding the instantaneous influence of visual cinematic stimuli on emotional aspects of narrative engagement, it leaves untouched the accumulating emotional impact of low-level visual stimuli as a narrative gradually unfolds. Given that filmmakers' implementation of such features varies as a function of plot development (Cutting, 2016a), the broader dynamic structure of cinematic narratives is crucial to understanding the visual landscape of film and its emotional consequence on viewers.

Critically, like film, emotion is defined by its dynamic nature. While no consensus has been reached regarding a psychological definition of emotion, most researchers agree that it generally involves a positive or negative experience accompanied by a change in physiological activity (Schacter, Gilbert, Wegner, & Nock, 2014). As such, emotion appears to be comprised of both subjective (i.e., cognitively experienced) and objective (i.e, physiological) components, but the interplay between the two is theorized in a number of competing ways. For instance, William James (1884) claimed that emotion results not from the cognitive processing of salient stimuli, but from the perception of changes that occur within the body in response to such stimuli. In other words, rather than following physiological change, emotion *is* physiological change (Critchley & Nagai, 2012). This explanation gives rise to the James-Lange theory, which states that a stimulus triggers bodily activity, which in turn produces an emotional experience in the brain (Lange, 1887). In contrast, the Cannon-Bard theory proposes that emotional experience in the brain occurs alongside (rather than as a consequence of) physiological activity (Bard, 1934; Cannon, 1929). And in the spirit of Aristotle's (350 BCE/1991) philosophical position that emotional experiences are shaped by our judgments and evaluations, Schachter and Singer's two-factor theory stipulates that subjective emotional experience follows cognitive inferences surrounding the cause of physiological arousal (Schachter & Singer, 1962). More recently, Barrett (2017) expanded on this contextual thread, claiming that the brain constructs emotion by imbuing sensation with meaning, and it does so by making predictions about the environment based on previous neural activity.

While these theories differ in their explanations of the specific trajectory of emotional experience, they are all grounded in the premise that emotion does, in fact, *unfold* over time, manifesting alongside a pattern of physiological change in response to a stimulus. Thus, emotion is a dynamic experience that is best understand temporally. As such, studies that measure both subjective and objective markers of emotion *in real time* as a film progresses will better capture the relationship between low-level visual salience and emotional reactivity among viewers.

Indeed, this relationship cannot be fully captured by retrospective emotion ratings, which dominate the extant literature on cinematically-induced subjective emotion (see Lehne & Koelsch, 2015). Such ratings typically capture discrete emotions (e.g., Davydov, Zech, & Luminet, 2011; Schaefer, Nils, Sanchez, & Philippot, 2010; Visch, Tan, & Molenaar, 2010), reflecting a theoretical approach to emotion that presupposes a set of core emotional states that are categorical and distinct (e.g., disgust, fear, happiness), biologically determined, and universally produced and recognized (e.g., Darwin, 1872/2009; Ekman & Friesen, 1975; Izard, Huebner, Risser, & Dougherty, 1980; Izard, 1992). However, given their wide range and subjective nature, as well as the fact that people are generally not adept at labelling their own emotional states (e.g., Taylor & Bagby, 2012; van der Velde et al., 2013), discrete emotions are difficult to measure in real time without disrupting the viewer. Fortunately, such limitations are not pertinent to the goals of this project, which is less concerned with the discrete emotions experienced by viewers and more concerned with the *occurrence* of emotion in general. Thus, a deeper look at alternative theories of emotion is warranted.

In contrast to models of emotion that treat emotional states as discrete entities, other researchers have championed a dimensional approach, whereby emotions are conceptualized as existing within a hypothetical two- or three-dimensional space. For instance, Russell's (1980) circumplex model suggests that emotions fall within a circular space comprised of two intersecting axes representing valence (i.e., positive vs. negative) and arousal (i.e., emotional intensity). The Positive Activation – Negative Activation (PANA) model similarly incorporates dimensions of valence and arousal, but assumes that positive affect and negative affect operate as two separate systems (Watson & Tellegen, 1985), and the Pleasure Arousal Dominance (PAD) model conceptualizes emotion as existing along the three aforementioned dimensions (Mehrabian, 1995). Other models offer a hybrid approach to emotion, combining aspects of discrete emotional theories and dimensional theories (e.g., Plutchik, 1991). A common thread among the wide array of dimensional models is that—rather than treating emotional states as discrete categories with rigid boundaries—they assume a degree of contextual influence on the subjective experience of emotion. For instance, in looking closer at the circumplex model (Russell, 1980), an emotional response that is situated high on the arousal axis and low on the valence axis could *objectively* be classified as a number of different emotions, including anger or fear. However, the *subjective* experience of such an emotional state as, for example, fear is the result of an individual's cognitive interpretation of their physiological state within the context of multiple factors, including the eliciting stimuli, prior experience, and semantic knowledge (Posner, Russell, & Peterson, 2005; Russell, 2003).

While dimensional models may, theoretically, better capture cinematically-induced emotion (particularly since this class of emotion appears to depend on such contextual influences as perspective taking and mental models; Busselle & Bilandzic, 2009), the difficulty in capturing multiple dimensions of emotional experience in real-time remains. This difficulty arises from the fact that the subjective experience of these emotional components would be difficult to ascertain over a continuous stretch of time.

Given that both discrete and dimensional models of emotion are not amenable to ongoing subjective measurement during film viewing, this project sought to operationalize emotion as existing along a single dimension that would:

a) Collapse multiple components of emotional experience into a single measure
b) Be relatively easy to capture in real-time with little cognitive load placed on the viewer

Here, Busselle and Bilandzic's (2009) approach to narrative engagement is informative. The authors claim that emotional engagement with a story is rooted in the arousal component of emotional experience rather than valence, and that its occurrence is not specific to any discrete emotion.

Thus, a suitable measure of dynamic emotional response is that of emotional tension—described broadly as a state of emotional arousal resulting from states of conflict, instability, dissonance, or uncertainty (Lehne & Koelsch, 2015). Emotional tension is independent of emotional valence, and as such, it captures a wide array of discrete emotional states. Furthermore, emotional tension necessarily unfolds over time, rendering it an optimal construct to reflect the time-varying nature of emotional experience in response to fluctuating cinematic stimuli. Methodologically, emotional tension can be captured in real time via the movement of a joystick or virtual slider. Paired with physiological markers of emotional arousal, a comprehensive picture of dynamic emotional response emerges—one that can be analyzed against a backdrop of the low-level visual features that comprise a whole film.

Study Overview

This project addresses a gap in the current psychological understanding of narrative engagement with cinema—that is, the influence of low-level visual features on emotional engagement in the context of the whole film. In other words, what role does the film as visual stimulus play in cinematically-induced emotion?

To answer this question, I measured viewers' emotional responses—in real time—to a series of short films that vary in genre. By capturing subjective emotional tension (via continuous movement of a joystick), as well as two objective measures of emotional arousal (via physiological markers), I was able to capture a time course of cinematically-induced emotion that can be mapped against the dynamic visual structure of each film as it unfolds. In isolating the effects of clutter, luminance, motion, shot density (as a proxy for shot duration), shot scale, and sound amplitude on all three measures of emotion, this project not only explores the extent to which each low-level feature influences emotional reactivity among viewers, but it also provides insight into the degree of emotional synchrony between viewers, as well as the degree of dissociation between subjective and objective measures of emotion during film viewing (with implications for understanding the nature of cinematically-induced emotion and its role in narrative engagement).

Methods

Study Design Considerations for Time-Series Data

The analysis of physiological data against eliciting stimuli comes with inherent difficulties due to the onset latency of physiological response systems (which typically range, in a positively skewed manner, from 1 to 4 seconds). This latency is attributed to the time required for stimulus processing, for automatic nervous system (ANS) nerve conduction, and—in the case of galvanic skin response (GSR)—for penetration of sweat through the epidermis (see Benedek & Kaernbach, 2010; Boucsein, et al., 2012; Glass, Beuter, & Larocque, 1988). Similarly, one can expect a brief lag in subjective emotional response measures as the viewer indicates their emotional state only after the stimulus has been processed. In the context of the current study, such asynchronicity is further complicated by the dynamic nature of film as a stimulus, with its multitude of continually fluctuating low-level variables. As such, classic techniques employing trough-to-peak analysis within a given response window are inappropriate (and often flawed in their usage; see Benedek & Kaernbach, 2010), since it is not feasible to attribute isolated peaks to specific instances of cinematic variation.

While not a perfect solution, this study utilized a binning method, in which both independent and dependent variables were averaged over 5-second intervals. The resulting datapoints captured the mean values among low-level cinematic features over each time window, as well as their cumulative influence on emotional response variables, thus building some room for physiological lag and enabling a more straightforward and interpretable analysis of the data. The statistical handling of physiological lag is further discussed below (see "Data Analysis").

Instruments and Materials

Subjective emotional tension

For the purposes of this study, subjective emotional tension was defined as any felt emotional reaction that viewers experienced in response to a film as it unfolded on screen. It was measured using a ThrustMaster USB Flight Stick with three movement axes (although only one axis was utilized for this study; see Procedures) and a weighted base with rubber pads to prevent shifting. As this joystick is intended primarily for use in flight gaming, it is designed to register a high degree of mechanical precision, and it is built with an enlarged hand rest to ensure secure and comfortable movement. The joystick also includes four buttons and one trigger, which were neither programmed nor utilized for the purposes of this study.

The joystick was connected via USB to a desktop computer enabled with LabVIEW software (National Instruments, 2016). LabVIEW is a visual programming platform commonly used to collect data from hardware devices. Data acquisition in LabVIEW operates through visual instruments, which are graphical interfaces that utilize a series of interconnected controls to program the monitoring and collection of data from specified devices. To allow for communication between LabVIEW and the ThrustMaster joystick, I programmed a visual instrument (VI) consisting of two primary components: a block diagram (containing the graphical source code commanding LabVIEW to monitor the joystick) and a front panel (a user interface consisting of button controls and a display graph). The graphical source code was constructed with a series of built-in "Input Device" controls created specifically for data acquisition from attached devices. These controls include the "Initialize Joystick" control, which signaled the start of data collection when a mouse was used to click the "run" button on the front panel; the "Acquire Input Data" control, which, in conjunction with a timing control, programmed the recording of data points every 5 seconds; and the "Close Input Device" control, which signaled the termination of data collection when a mouse was used to click the "stop" button on the front panel.

The VI recorded the amplitude of joystick displacement on an arbitrary scale from 0 (no displacement) to 32,768 (maximum displacement) with a resolution of 256 units, resulting in 128 possible values. For monitoring purposes, a line graph on the front panel was programmed to visually display joystick displacement in real time. Upon termination of data collection, LabVIEW generated an Excel file containing two variables: timepoint and amplitude.

Objective emotional tension

Objective indices of emotion are typically captured by measuring bodily changes that signal activity within the ANS. The ANS operates between the central nervous system and the visceral organs and, as such, is integral in producing physiological reactivity to external stimuli. Within the ANS, the sympathetic nervous system (SNS) acts to quickly mobilize bodily responses to emotionally-charged stimuli (i.e., “fight-or-flight” responses), while the parasympathetic nervous system (PNS) slowly dampens physiological reactivity as it prepares the body for tasks that do not require immediate action (i.e., “rest-and-digest” or “feed-and-breed” responses). Specifically, the SNS triggers a release in stress hormones—namely, cortisol and adrenaline—which produce changes in heart rate, sweat gland activity, pupil dilation, digestive processes, and respiratory rate (among other physiological changes), and these responses generally subside as emotionally-charged stimuli dissipate. The SNS is involuntarily activated, and it operates largely below the level of consciousness, rendering it an ideal candidate for the objective measure of emotional tension (for a review, see Keltner, Oatley, & Jenkins, 2013; Schmidt & Thews, 1989).

For the purposes of this study, I focused on two particularly non-intrusive markers of physiological reactivity: heart rate and electrodermal activity. These measures capture, respectively, the number of heart beats per minute and the degree of electrical conductance of the skin (which fluctuates in response to sweat production). Both are known to reflect SNS arousal across a broad range of emotional experiences; these measures generally increase when individuals experience basic emotions such as anger, disgust, fear, happiness, and surprise, as well as more complex emotional states including frustration, amusement, and engagement (for reviews, see Bergstrom et al., 2014; Kreibig, 2010; Lisetti & Nasoz, 2004; Neumann & Westbury, 2011; Shu et al., 2018). However, research linking heart rate and electrodermal activity to sadness yields mixed results, where states of sadness associated with crying generally show an increase in physiological arousal, while sadness defined by an approach component (e.g., empathy or tenderness) shows a decrease in physiological arousal (Gross, Frederickson, & Levenson, 1994; Marsh, Beauchaine, & Williams, 2008; Sternbach, 1962). Nonetheless, meta-analyses suggest a generalized sympathetic activation of heart rate and electrodermal activity in response to emotionally-charged stimuli (Kreibig, 2010), where both measures reflect intensity, but not valence, of emotional arousal (Bergstrom et al., 2014).

It is important to note that, while heart rate is often utilized as a physiological measure of emotional arousal (particularly when capturing empathy; see Neumann & Westbury, 2011), many researchers suggest the use of heart rate variability as a more accurate (or additional) measure of emotional response (see Appelhans & Luecken, 2006; Brouwer et al., 2018). However, others question the validity of heart rate variability as an objective measure of emotion at milder levels of physiological arousal (Choi et al., 2017). Nonetheless, limitations imposed by its accurate recording and measurement (i.e., via electrocardiogram; Appelhans & Luecken, 2006) precluded its use in this study.

Heart rate. Pulse rate (PR) was measured as a proxy for heart rate via a Contec pulse oximeter (model CMS50DA+), which attaches to an individual's index finger. The device features a small LED screen with a blood oxygen saturation (SpO2) value display, a PR value display, and a PR waveform display. The oximeter is accompanied by software that allows for the real-time transmission of PR data to a desktop computer via USB connection. The software includes a recording function that stores data in a line graph displaying the user's PR once every second (with time indicated according to the desktop computer clock). The oximeter accommodates a PR between 30 and 250 beats per minute (bpm) with a resolution of 1 bpm and an accuracy of ±2 bpm.

Electrodermal activity. Electrodermal activity was measured via a NeuLog galvanic skin response (GSR) sensor (model NUL217). GSR describes the variation in electrical conductance of skin tissue due to moisture on the surface of the skin resulting from SNS activation of sweat glands. GSR sensors capture this variation by applying a small voltage to two electrodes attached to two separate fingers and measuring the resulting resistance (by calculating the voltage difference between the two electrodes). The skin on the hands is particularly sensitive to this measure, given the high density of sweat glands located on the palms and fingers.

The NeuLog GSR sensor consists of two silver probes attached to Velcro straps that wrap around the base of the fingers. The device operates by applying a continuous 0.5 voltage current to each of the electrodes. Dry skin produces high resistance and low conductance, while moist skin produces low resistance and high conductance; thus, the easier the skin conducts the electrical current, the greater the GSR reading, indicating greater emotional arousal (see Balters & Steinert, 2017; Bergstrom et al., 2014).

For data collection purposes, the GSR sensor was connected to a NeuLog USB module, which sends digital data from the sensor to a desktop computer via USB connection. The sensor is accompanied by software that allows for the monitoring of GSR in real-time (via a continuously updated line graph). The software also includes an experiment function that records and stores data according to the user's specifications. I programmed the software to record GSR values once every second (the slowest sampling rate available) for up to one hour after initiation of the recording process. The data were recorded in arbitrary units with a range from 0 (zero conductance) to 65,279 (maximum conductance), with a resolution of 1 unit. At the end of the recording period, the software generated an Excel file containing two variables: timepoint and GSR.

Cinematic stimuli

The short films shown to participants include the following:

- A 2 minute and 26 second trailer for *The Light Between Oceans* (Cianfrance, 2016), a romantic drama that depicts the events that transpire after a lighthouse keeper and his wife discover a baby in a rowboat and decide to raise a her as their own.
- *Borrowed Time* (Coats & Hamou-Lhadj, 2015), an animated action adventure of 6 minutes and 45 seconds that depicts a Sheriff who revisits the site of an accident, only to be confronted with memories that force him to relive his role in a tragedy that occurred decades ago.
- *Paperman* (Kahrs, 2012), an animated romantic comedy of 6 minutes and 33 seconds that depicts a day in the life of an office worker after he meets an attractive woman and proceeds to use a fleet of paper airplanes to get her attention.
- *Stutterer* (Cleary, 2015), a live-action drama of 13 minutes and 4 seconds that depicts a man whose crippling stutter not only leads him to feel isolated from the world, but also compromises a flourishing online relationship.

- *The Voorman Problem* (Gill, 2013), a live-action science fiction of 12 minutes and 27 seconds that depicts a psychiatrist's experience dealing with an incarcerated patient who claims to be a god.

Thus, the four films span a wide range of genres, representing both live action and animated formats. Additionally, the films were reasonably short to accommodate the study task while still containing traditional narrative structures in their entirety. While a range of theories exist regarding narrative structure (for a review see Cutting, 2016a), collectively, the selected short films best align with a structure put forth by Thompson (1999) and Bordwell (2006), which stipulate four acts of roughly equal duration: the *setup* (in which the characters and their goals are introduced, along with an inciting incident that drives the dramatic progression of the film), the *complicating action* (in which the goals of the protagonist are derailed and further developed), the *development* (in which the story broadens, the protagonist struggles to achieve their goals, and secondary characters are developed), and the *climax* (in which the protagonist launches into action, and the story is resolved and/or a new social order established)[1].

The four films and the trailer were screened on a 13-inch MacBook Air, and participants wore Sony ZX Series On-Ear Headphones while viewing the films. These headphones block a

---

[1] While *Paperman* and *The Voorman Problem* neatly fit this model, *Borrowed Time* and *Stutterer* each consist of a comparatively uneven act structure. Respectively, the complicating action and development are considerably shorter in duration than the other three acts. One could argue for combining the second and third acts of both films to reflect the traditional three act structure promoted by Aristotle (335 BCE/2013), thus achieving a more even narrative division. For consistency, however, I have applied the four act structure across all four films (see Appendix A for a narrative breakdown of each film).

moderate amount of noise, but are not strictly noise-cancelling, thus allowing for ongoing communication with participants in between film viewings.

Participants and Randomization

Participants were recruited via SONA, an online platform used by the Cornell Department of Psychology to sign up undergraduate students for experimental studies. This study was advertised as a project investigating viewer's emotional responses to film that would take about 45 minutes to complete.

In the first phase of the study, I intended to recruit 20 participants to view *Paperman* and *Stutterer.* However, data for four participants were excluded due to either participants' failure to follow instructions (i.e., moving the joystick in the wrong direction) or equipment failure (i.e., power outages) that disrupted data collection. Thus, a total of 24 participants were recruited in order to obtain 20 complete sets of data. In the second phase of the study, I again intended to recruit 20 participants to view *Borrowed Time* and *The Voorman Problem*. However, two participants were excluded due to equipment failure. Thus, a total of 22 participants were recruited in order to obtain 20 complete sets of data. Of the final 40 participants from both phases of the study, 25 were female and 15 were male, with a mean age of 19.3 years. Twenty-four participants self-identified as White, 10 as Asian, four as Black, and two as Mixed Race. All participants received two SONA credits for their participation in the study (which are generally used as extra credit for psychology courses in which students are enrolled).

All participants viewed one live action and one animated film (either *Paperman* and *Stutterer*, or *Borrowed Time* and *The Voorman Problem*) in randomized order, such that half of the participants in the first phase of the study viewed *Paperman* followed by *Stutterer*, while the other half viewed *Stutterer* followed by *Paperman*. Likewise, in the second phase of the study, half of the participants viewed *Borrowed Time* followed by *The Voorman Problem*, while the other half viewed *The Voorman Problem* followed by *Borrowed Time*. All participants viewed the trailer for *The Light Between Oceans* during a practice trial preceding their viewing of the two short films.

Procedures

One participant at a time completed the study. Each participant arrived at the lab at a predetermined time (as scheduled on SONA), at which point they were asked to read and sign a consent form and given the opportunity to ask questions about the study. Once their questions were answered, the participant was led into a quiet room within the lab space and seated in front of the laptop computer, which was pre-adjusted with the screen at maximum brightness and with volume at the midpoint of the available range.

The participant was first asked to remove all jewelry from their hands and to clean the fingers of their non-dominant hand with an alcohol wipe. They were then instructed to attach the GSR sensor probes to the third (middle) and fourth (ring) fingers of their non-dominant hand by wrapping the Velcro straps snugly at the base of each finger (a diagram was provided as guidance). Then, the participant was asked to insert the second (index) finger of their non-dominant hand into the pulse oximeter. Once both devices were attached to the participant's fingers, they were instructed to place their hand in a relaxed position on the desk while avoiding pressure on the GSR sensor probes (given that pressure from hard surfaces can disrupt readings; Bergstrom et al., 2014).

Next, the participant was informed of the definition of emotional tension (i.e., a broad state of felt emotional arousal, regardless of the category of emotion experienced) and instructed on how to use the joystick to indicate their emotional tension as they viewed the short films. Specifically, they were instructed to keep their dominant hand resting on the joystick throughout the duration of each film, and to move the joystick in the sagittal direction away from their body when they felt any degree of emotional tension. The participant was informed that the central position of the joystick indicated a neutral emotional state, and that the distance they moved the joystick from the central position should correspond to the degree of emotional tension that they felt at any given moment. They were also instructed to move the joystick back toward the central position as their emotional tension dissipated. The participant was then asked to move the joystick to get a feel for its range. They were instructed not to move the joystick in the horizontal direction or in the sagittal direction toward their body (they were assured that the joystick "catches" at the central position, thus preventing unintentional backward motion). The participant was then given the opportunity to ask questions about the task. Once their questions were answered, they were informed that they would practice the task while watching the trailer for *The Light Between Oceans*.

The participant then put on the headphones and returned their non-dominant hand to its resting position. They were instructed to keep their hand as still as possible throughout the duration of the study. The pulse oximeter and GSR sensor were then turned on, and the corresponding software programs were started to ensure that the devices were running properly (the screen of the desktop computer that ran the software programs was angled away from the participant to prevent distraction). Then, the recording functions for the pulse oximeter, the GSR sensor, and LabVIEW were initiated. Given that LabVIEW recorded joystick data every 5 seconds, all programs were initiated at exact 5-second intervals (e.g., 10:32:40 or 10:32:45, as indicated by the desktop computer clock) to allow for precise merging of the data. Exact start times were manually recorded for all programs.

Once the participant indicated that they were ready to begin, they were instructed to adjust the volume on the laptop and/or the angle of the screen if needed, but they were encouraged to make these adjustments only at the beginning of each film to minimize disruptions during the joystick task. Then, the trailer was started in full-screen mode, and the participant began the task. Upon completion of the trailer, the participant's performance was corrected if necessary (i.e., if they moved the joystick in the wrong direction, as indicated by the graph generated by LabVIEW), in which case they were asked to demonstrate the correct movement of the joystick before beginning the test trials. The participant was also given the opportunity to ask questions about the task. Once their questions were answered, the participant rested for a period of three minutes to allow their physiological measures to return to baseline. Then, the participant completed the joystick task for the first short film. After the first film, the participant again rested for three minutes before completing the joystick task for the second short film. The start time for each film was manually recorded (again, at 5-second intervals as indicated by the desktop computer clock).

At the end of the second film, the recording functions for all three software programs were halted, and the participant was instructed to remove the headphones, pulse oximeter, and GSR sensor probes. Then, they were asked whether they had previously seen either of the two short films, and their responses were manually recorded. Finally, the participant was asked if they had any questions about the research. After their questions were answered, they were thanked for their time and released from the study. The data files generated by the three software programs were exported, saved and labelled with the participant's study ID, and then backed up on a USB drive. Finally, the joystick, pulse oximeter, and GSR sensor probes were sterilized with alcohol wipes in preparation for the next participant.

Hypotheses

While this project was largely exploratory, I did put forth a number of tentative hypotheses based on past research. I predicted that, overall, emotional tension would increase with greater motion, greater shot density, greater sound amplitude, closer scaled shots, and lower levels of luminance. However, I held no expectations regarding the influence of clutter. Additionally, I held no expectations regarding how these predictions would manifest differently across genre or across live action vs. animated films. Thus, by incorporating a wide range of genres within the cinematic stimuli, I sought to not only detect overarching patterns, but to identify, through exploratory analyses, specific conventions of genre that might be the driving force behind these patterns. Finally, I held no expectations regarding how emotional reactivity would vary across the three different measures of emotion.

Additionally, while it is highly probable that numerous interactions between the cinematic variables would be reflected in the data (e.g., hypothetically, luminance might interact with shot scale to affect GSR, and the interaction between shot density, motion, and clutter might affect joystick displacement, and so on), I did not attempt these analyses for pragmatic reasons.

Data Coding and Cleaning

While data collection was ongoing, a research assistant manually entered the PR data into an Excel file for each participant (entering only the values that occurred every 5 seconds following the start of each film). A spot check of the research assistant's work revealed a 99.5% accuracy rate. Likewise, the GSR files were modified to retain only the values that were recorded every 5 seconds following the start of each film.

Once data collection was complete, the data files for all 40 participants were split by film (based on the recorded timestamps). The data were trimmed such that the first row of data corresponded to the opening shot of the film and the last row of data corresponded to the last shot of the film. Values that exceeded the maximum range for each variable (as defined by the corresponding software programs) were excluded, as such values were likely the result of technical glitches (e.g., joystick displacement values greater than 32,768 were excluded, as were PR values greater than 250 and GSR values greater than 65,279). Then, the data for all three variables were standardized (via z-score transformation) within each participant to allow for easier comparison between emotion measures. Finally, a separate file was created for each emotion measure (joystick displacement, PR, and GSR) for each film, resulting in twelve files, and the data for all participants were merged into the corresponding files.

To capture the low-level visual structure of each film, a MATLAB script first divided the films into 5-second bins to align the measurement of cinematic variables with the measurement of emotion variables. Then, all frames were downsampled to a 256 x 256 array of pixels and converted to 8-bit grayscale, yielding pixel values ranging from 0 (black) to 255 (white). These values were gamma-transformed to reflect the nonlinear sensitivity of the human eye to brightness. Then, a series of MATLAB scripts generated the following variables:

- *Luminance* was measured by averaging pixels within each downsampled frame, and then across frames within each 5-second interval, yielding a single luminance value for each timepoint (ranging from 0 to 255, such that lower values indicated darker frames and higher values indicated brighter frames).

- *Motion* was operationalized as flicker motion, defined as the change in pixel values across adjacent frames of a film (without regard to the structural patterns of pixel change). Pixels in successive downsampled frames were correlated, and correlations were averaged within 5-second bins, resulting in values ranging from 0-1. For ease of data interpretation, each value was transformed to 1-r, such that lower values indicated less motion and higher values indicated more motion. This yielded a single value for each timepoint.
- *Clutter* was operationalized as edge density, captured using a Laplacian of Gaussian edge-detecting algorithm, which defines edges as zero crossings within each frame (i.e., regions where there is a sharp change in luminance; Marr, 1982). The algorithm generated images containing a series of jagged lines (of a single pixel in width) where edges occurred. The proportional clutter in each frame was assessed by measuring the ratio of black to white pixels. Proportional clutter was averaged within each 5-second interval, yielding a single edge density value for each timepoint.
- *Sound amplitude* captured the mean volume (including both non-diegetic and diegetic sound) across each 5-second interval. Amplitude values were scaled to range between 0 and 40.

The following variables required manual coding of the original (non-binned, non-downscaled) films:

- *Shot density*, defined here as the number of shot transitions (i.e., cuts) occurring within each 5-second interval, was determined by timestamping the first frame of every shot.

The number of first frames in each 5-second interval was then summed to compute shot density for each timepoint. Thus, lower values indicated a slower shot pace (i.e., fewer shots of longer duration within the preceding interval), and higher values indicated a quicker shot pace (i.e., more shots of shorter duration). In other words, shot density serves as a proxy for shot duration, as the two are directly and negatively related.

- *Shot scale*, defined as the relative size of a character within the frame of a shot, was coded on a scale from 1-7 (see Fig. 1), with lower values indicating longer, more distant shots (where the character's face is relatively small within the frame) and higher values indicating closer shots (where the character's face is relatively large within the frame). Only the fifth frame of each shot was coded (given that the first few frames of a shot are occasionally blurred or the characters' faces otherwise obstructed). A research assistant and I coded *Paperman* together to ensure that we agreed on the implementation of the coding system (resulting in an inter-rater reliability of 94%). Then, we split the remainder of the films and spot checked each other's work. Shot scale values were then computed by averaging fifth-frame shot scale within each 5-second interval, yielding a single shot scale value for each timepoint.

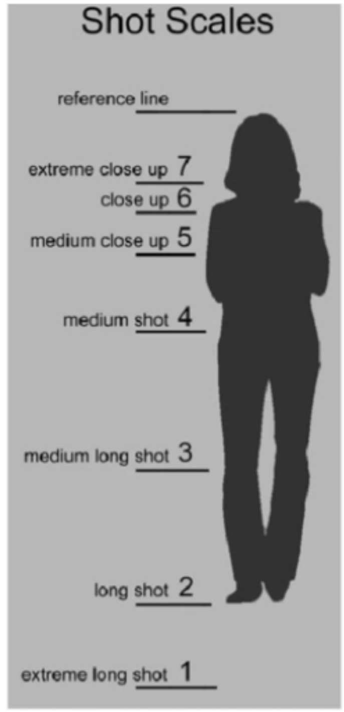

Fig. 1. A visualization depicting how shot scale was determined and coded (adapted from Cutting & Armstrong, 2016).

The cinematic variables described above were added to each of the emotion data files for all four films (aligned according to timepoint), along with variables that coded for participant ID, gender, race, film, and whether or not the participant had previously seen the film. Then, the data from all twelve files were merged, long format, into a single file.

Data Analysis

To assess the influence of the visual structure of each film on both subjective and objective emotional tension, the data were analyzed via multi-level regression modelling. A single model was run for each emotion measure across each of the short films, resulting in twelve distinct analyses. Each model treated the emotion measure of interest (joystick displacement, PR, or GSR) as a repeated measure nested within participant ID (to control for random subject variance). Clutter, luminance, motion, shot density, shot scale, and sound amplitude were treated as fixed effects, along with gender and race. To accommodate the time-series nature of the data, timepoint was added as a polynomial fixed effect reflecting the polynomial function that best fit the corresponding repeated measure.

An omnibus analysis was also performed to assess the aggregate effects of each of the low-level cinematic features across all four films. The regression model described above was modified such that timepoint was removed and film added as a fixed effect. A separate model was run for each of the three emotion variables. The resulting analyses were then split by live action vs. animated films to further assess distinct patterns across styles of filmmaking.

Additionally, to assess the extent to which subjective and objective measures of emotion track each other during film viewing, the three emotion measures were correlated within each film. Cross-correlation techniques are commonly used to assess correlation among time-series variables—particularly physiological measures—that may be subject to onset latency. This strategy determines the similarity between two time series when they are shifted incrementally in relation to each other, thus identifying the time lag resulting in maximal correlation among a given pair of variables (see Yavuz, Claassen, & Kleinberg, 2019). As such, a cross-correlation was performed to capture potential physiological lag that may be present in the data.

Finally, *Paperman* was isolated as a film of interest, in that it was the only film that a portion of participants had previously seen. Thus, each of the three emotion measures were compared between groups (familiar versus naïve viewers) to assess differences in emotional tension as a function of prior knowledge regarding the narrative.

All regression analyses were interpreted with a Bonferroni corrected alpha criterion of 0.008 (to account for multiple comparisons within each emotional response variable, i.e. 0.05/6).

Results

Data Cleaning

After each film was trimmed to begin and end with the opening and closing shots, *Borrowed Time* consisted of 67 timepoints (including a start time of 0) for a length of 5 minutes and 30 seconds, *Paperman* consisted of 76 timepoints (including a start time of 0) for a length of 6 minutes and 15 seconds, *Stutterer* consisted of 147 timepoints (including a start time of 0) for a length of 12 minutes and 10 seconds, and *The Voorman Problem* consisted of 133 timepoints (including a start time of 0) for a length of 11 minutes and 0 seconds.

Due to erratic measurement in the equipment, 49 joystick datapoints of values greater than 32,768 were excluded. Specifically, 13 datapoints for *Borrowed Time* were excluded, including one each at timepoints 24, 38, 39, and 53, two each at timepoints 25, 40, 41, and three at timepoint 42. Eight datapoints for *Paperman* were excluded, including one each at timepoints 43, 44, 63-67, and 69. Twenty-six datapoints for *Stutterer* were excluded, including one each at timepoints 3-6, 10, 127, 130, 131, 133-135, 138, 139, and 142-146, and two each at timepoints 128, 136, 140, and 141. Two datapoints for *The Voorman Problem* were excluded, including one each at timepoints 96 and 106. In addition, nine PR datapoints were missing (again, due to erratic measurment)—including seven datapoints for *Paperman* at timepoints 69-75 and two datapoints for *The Voorman Problem* at timepoints 130 and 131. All GSR datapoints were recorded and retained. Thus, of the 8,460 possible datapoints for each emotion measure, 99.42% were included for the joystick analyses, 99.89% for the PR analyses, and 100% for the GSR analyses.

To ensure a reasonably normal distribution for each of the three emotion variables, I followed standard recommendations to avoid formal normality testing for sample sizes greater than 300 (as such tests are overly sensitive to small variations in large datasets) and to instead examine raw skewness and kurtosis values (with an absolute skew value greater than 2.00 and an absolute kurtosis value greater than 7.00 serving as reference points for a substantial departure from normality; George & Mallery, 2010; Kim, 2013; West, Finch, & Curran, 1995). Given these guidelines, all three emotion variables exhibited a reasonably normal distribution: the joystick data were distributed with a skew of 1.714 (SE = .027) and a kurtosis of 5.320 (SE = .053), the PR data were distributed with a skew of .496 (SE = .027) and a kurtosis of .704 (SE = .053), and the GSR data were distributed with a skew of .396 (SE = .027) and a kurtosis of -.263 (SE = .053). As such, no further transformations were performed on the dataset.

Finally, a sixth-order polynomial function was found to be the best fit for all three emotion variables across each of the four films. Thus, to accommodate the time-series nature of the data, timepoint was added as a sixth-order polynomial fixed effect to each of the twelve regression models.

Gender and Race Effects

In an omnibus analysis, gender did not exert an influence on joystick displacement ($t(8402) = .03$, $p = .973$), PR ($t(8442) = .01$, $p = .999$), or GSR ($t(8451) = .02$, $p = .983$) measures. Similarly, race exerted no influence on joystick displacement ($t(8402) = -.03$, $p = .975$), PR ($t(8442) = -.01$, $p = .994$), or GSR ($t(1, 8451) = -.05$, $p = .963$). Additionally, neither factor was significant for joystick displacement, PR, or GSR when the data were split by film. Thus, gender and race were removed as fixed factors from subsequent analyses.

Emotion Variables

Before investigating the effects of cinematic structure on emotional response among viewers, I will first describe the time course of the emotion measures across each of the four films, indicating plot points that correspond to prominent peaks in the data (Figures 2-5). This depiction will lay a foundation for subsequent analyses, which repeatedly make use of the following graphs. (See Appendix A for a complete plot breakdown of each film.)

Borrowed Time

Viewers of *Borrowed Time* exhibited six prominent peaks across the three emotion measures (Fig. 2). These peaks correspond to the following narrative plot points:

1) In the flashback, the boy has just been handed the reins of the horse-drawn carriage that he and his father are riding in as they flee from a gunman. The father leans over the back of the carriage, rifle in hand.
2) The boy and his father are fleeing from a gunman in a horse-drawn carriage, which has just toppled over, flinging the father from its interior. The father rolls over a cliff
3) While struggling to pull his father back up over the cliff, the boy accidentally shoots his father with a rifle. The father's pocket watch flies through the air and lands on the ground nearby, splattered with blood.
4) In the present day, the sheriff slips over the edge of the same cliff that his father fell from decades age. He clings to the rockface.
5) The sheriff has just wiped dirt from the face of his father's old pocket watch to reveal a portrait of himself (as a young boy) with his father, both with content facial expressions.
6) The sheriff sobs as he grasps his father's old pocket watch.

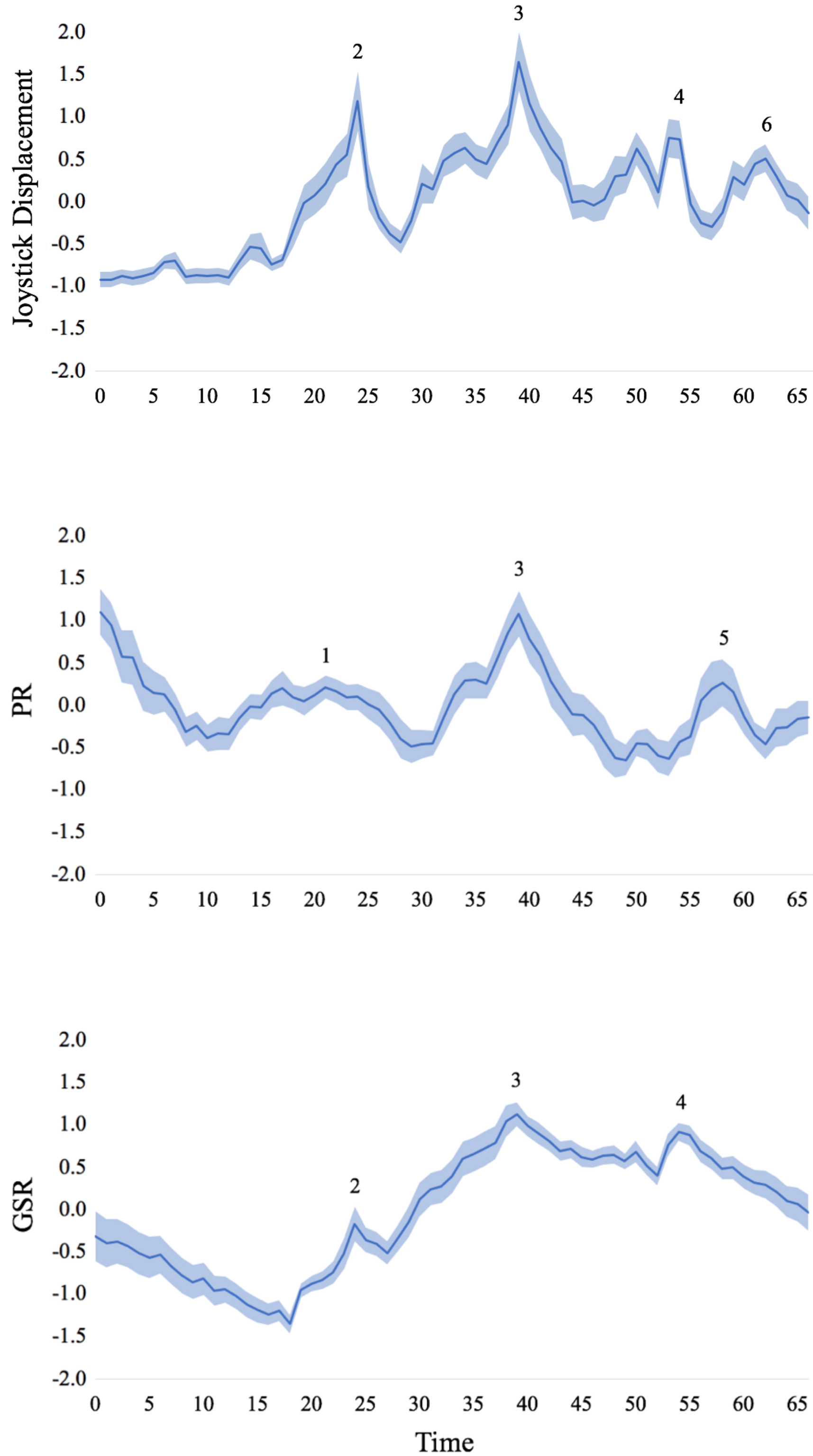

Fig. 2. Mean z-standardized emotion measures over time for *Borrowed Time*. Darker bands indicate 95% confidence intervals.

Paperman

Viewers of *Paperman* exhibited seven prominent peaks across the three emotion measures (Fig. 3). These peaks correspond to the following narrative plot points:

1) George stands on a train platform; a train rushes by.
2) At work, George has folded his last piece of paper (marked with Meg's lipstick) into a paper airplane. He is poised to throw the airplane from the office window when it is sucked away by a gust of wind. George leans out of the (very high) open window, fruitlessly reaching for the airplane.
3) Having run out of his office building to catch up to Meg, George nearly collides with several vehicles as he dangerously crosses the street.
4) George notices, perched on a mailbox on a busy city street, the paper plane that he lost to the wind. George snatches the airplane and, dejected, throws it far into the air.
5) On the train, a swarm of paper airplanes keeps George pinned to his seat. A mother pulls her child away from George.
6) Meg steps from the train onto the platform, beholding the paper airplane marked with her own lipstick.
7) Alongside the closing credits, a series of snapshots show George and Meg happily chatting in a cafe, with the epochal paper airplane sitting in front of them.

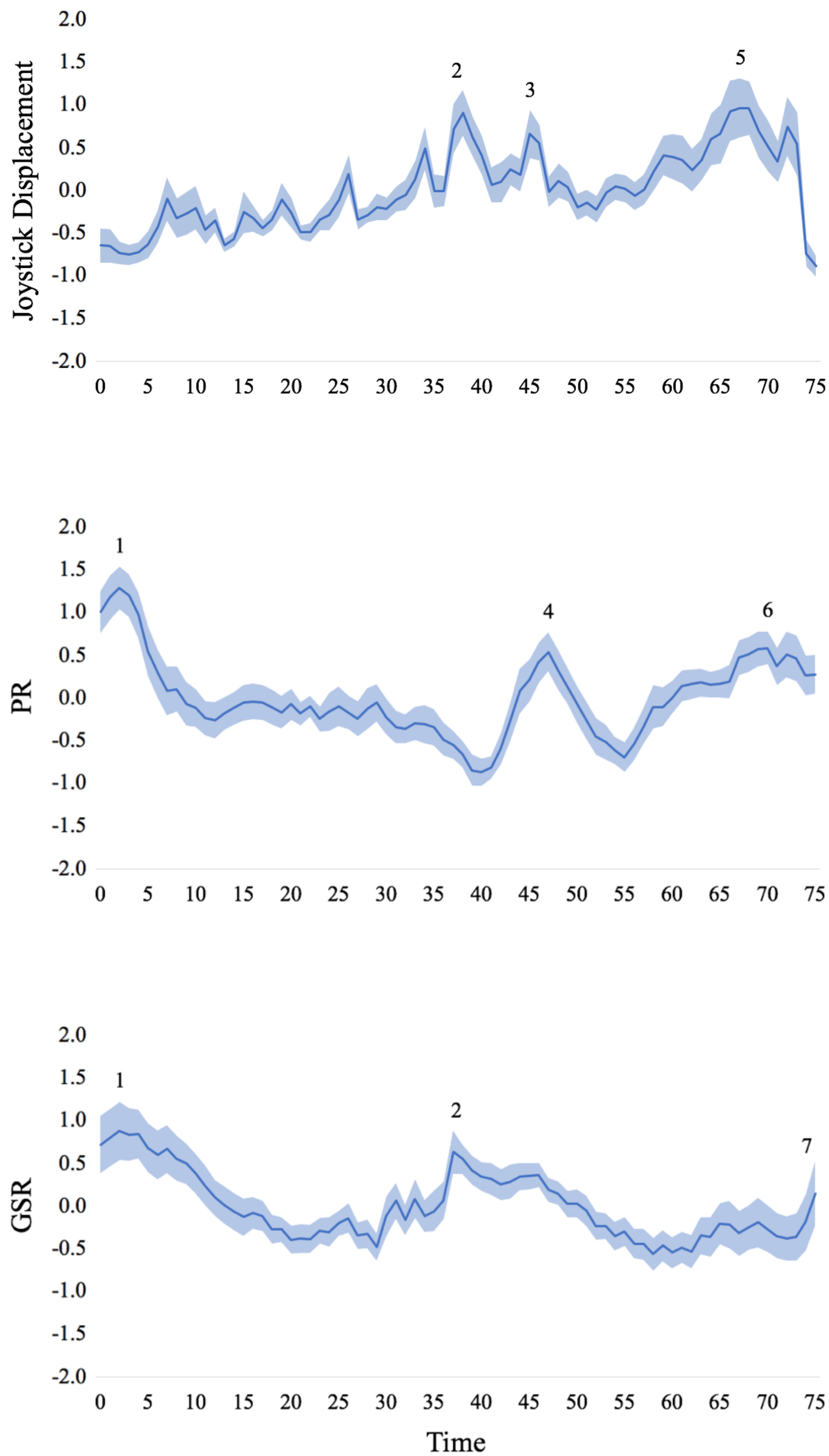

Fig. 3. Mean z-standardized emotion measures over time for *Paperman*. Darker bands indicate 95% confidence intervals.

Stutterer

Viewers of *Stutterer* exhibited seven prominent peaks across the three emotion measures (Fig. 4). These peaks correspond to the following narrative plot points:

1) Due to a severe stutter, Greenwood struggles to communicate with a customer representative when he places a call regarding his internet bill.
2) Reciting a poem out loud to his father, Greenwood falters over the words.
3) Greenwood has just received a message from Ellie, his online love interest, requesting to meet him in person for the first time. Greenwood closes his laptop.
4) Greenwood is sitting at a bus stop checking his phone when his attention is caught by a man shouting angrily at his girlfriend.
5) Greenwood confronts the angry man, who then turns to Greenwood and shouts obscenities at him. The screen goes black.
6) Having walked to their agreed meeting place, Greenwood looks across the street to see Ellie in person for the first time.
7) Ellie has just glanced across the street to see Greenwood. Tenderly, they hold each other's gaze.

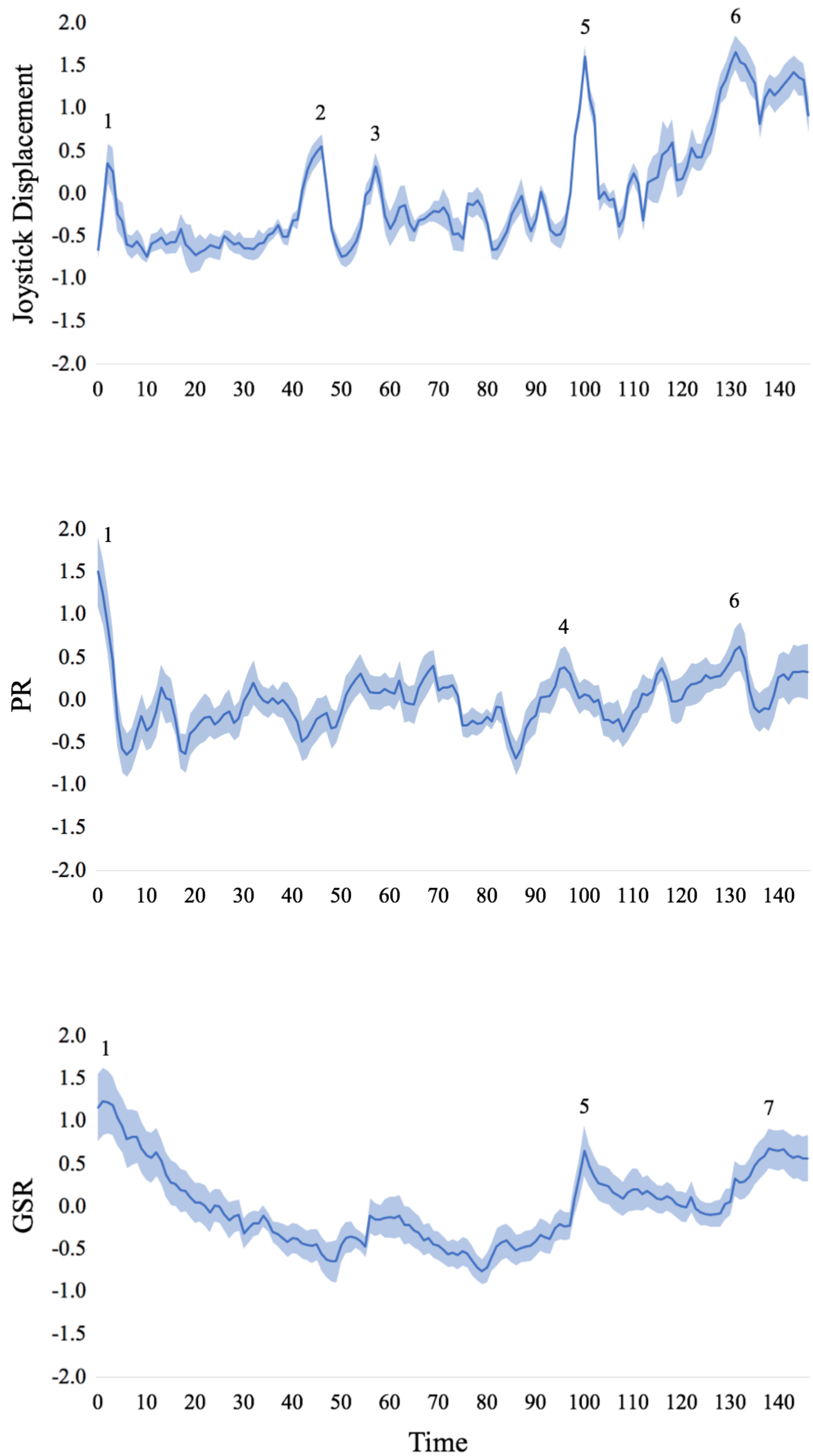

Fig. 4. Mean z-standardized emotion measures over time for *Stutterer*. Darker bands indicate 95% confidence intervals.

The Voorman Problem

Viewers of *The Voorman Problem* exhibited eight prominent peaks across the three emotion measures (Fig. 5). These peaks correspond to the following narrative plot points:

1) Loud music plays as Dr. Williams walks quickly toward walled grounds.
2) After arriving at the entrance to a prison, Dr. Williams introduces himself through a speaker on the front door. No one responds. Faint chanting can be heard through the speaker.
3) In Governor Bentley's office, the governor invites Dr. Williams to sit down, and then pours liquor into his own coffee cup.
4) Dr. Williams walks into a room in which Voorman, a delusional inmate, is seated and restrained in a straitjacket.
5) Dr. Williams, seated across from Voorman, opens a notepad and begins to speak.
6) At home, Dr. Williams consults a map to discover that Belgium has disappeared (just as Voorman promised it would).
7) Back at the prison, Voorman has just threatened to switch lives with Dr. Williams, promising to seduce his wife. Dr. Williams responds by calling Voorman a sick man.
8) Voorman has switched bodies with Dr. Williams. Voorman exits the room with Dr. Williams belongings, and Dr. Williams shouts, desperately into the empty room, that *he* is the real Dr. Williams.

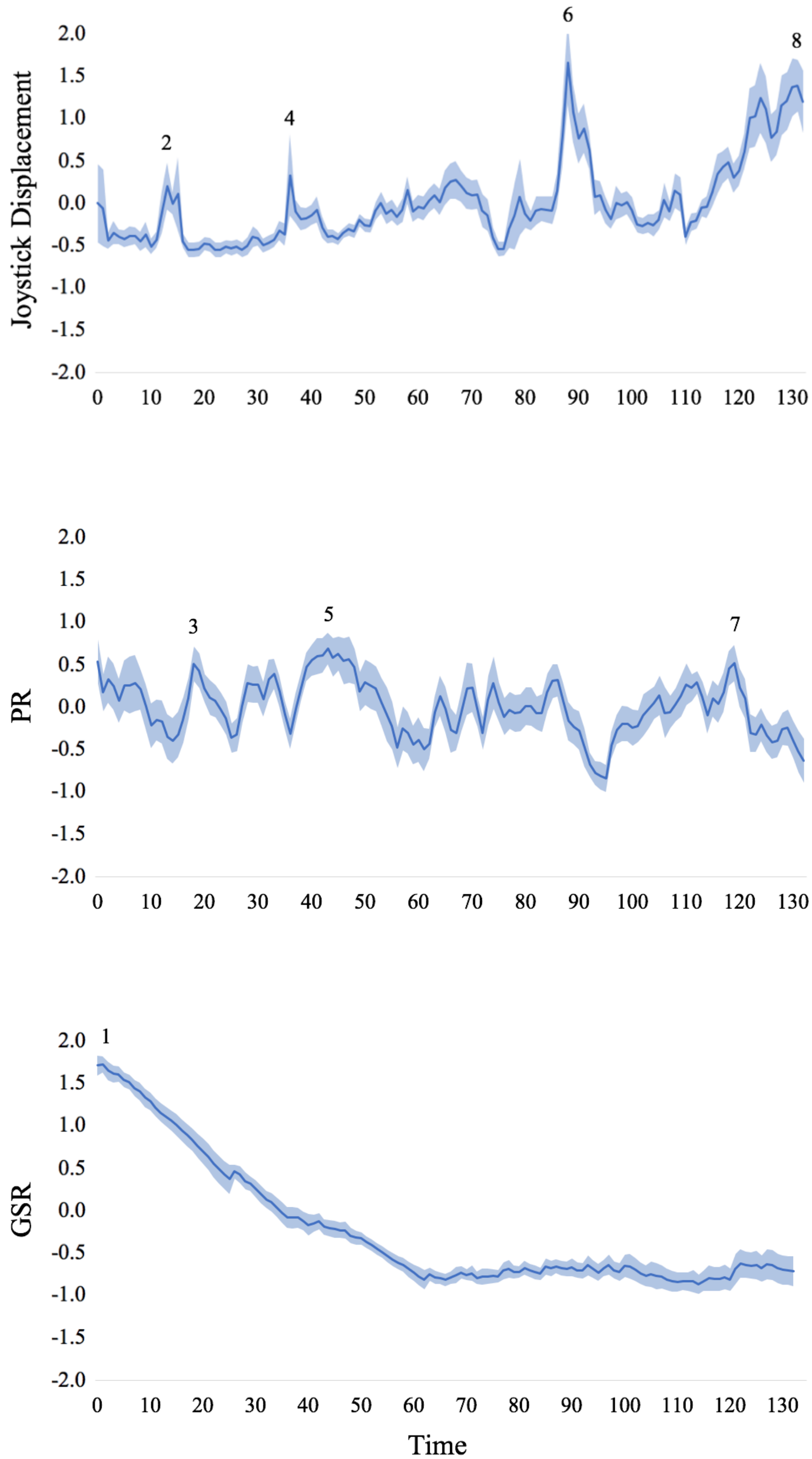

Fig. 5. Mean z-standardized emotion measures over time for *The Voorman Problem*. Darker bands indicate 95% confidence intervals.

Emotional Synchrony

Figures 2-5 not only depict the time course of emotional response as a function of the narrative of each film, but they also depict a remarkable synchrony in emotional response among viewers. The consistently narrow confidence bands suggest that viewers respond similarly to film in terms of both subjective and objective measures of emotion. This synchrony adds an important piece to the body of research already established in this area: in addition to attentional synchrony (Smith & Henderson, 2008; Smith, 2012, 2013) and neural synchrony (see Hasson et al., 2010), we now have evidence that viewers' physiological activity and their felt emotional states are synchronized as well. In an attempt to forge an understanding of this emotional synchrony in the context of film structure, the following sections delve deeper into the influence of each of the six low-level cinematic variables on emotional tension.

The results of all twelve regression analyses are divided by cinematic variable (clutter, luminance, motion, shot density, shot scale, and sound amplitude), and each variable is described in relation to each of the four films.

Cinematic Variables

Clutter

Fig. 6 displays the pattern of clutter over the course of each film. *Stutterer* consists of the highest degree of clutter (M = .063, SD = .017), followed by *Paperman* (M = .058, SD = .012). Comparatively, *The Voorman Problem* (M = .048, SD = .011) and *Borrowed Time* (M = .037, SD = .007) consist of less clutter, with relatively fewer peaks. The influence of clutter on emotional engagement among viewers of each film is described in greater detail below.

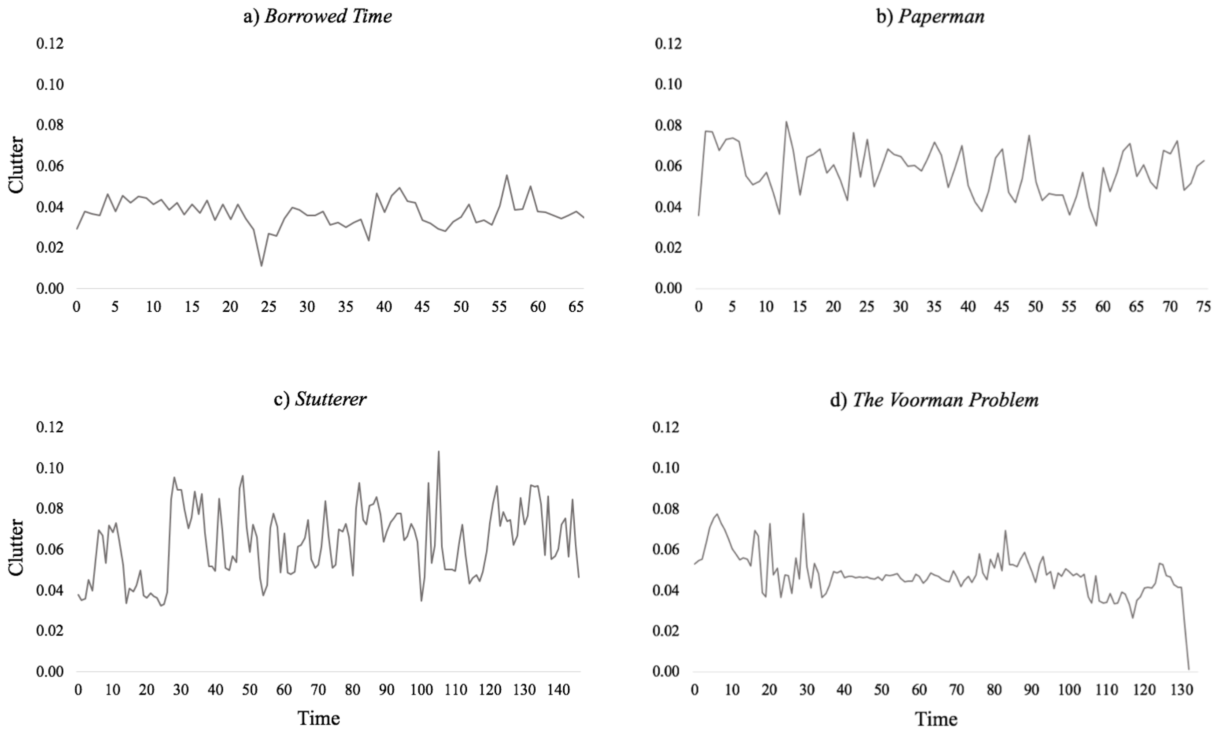

Fig. 6. Clutter over time in a) *Borrowed Time*, b) *Paperman*, c) *Stutterer*, and d) *The Voorman Problem*. Clutter units refer to the proportional edge density in each frame, such that higher values indicate more clutter across the preceding 5-second interval, while lower values indicate less clutter.

Borrowed Time. Neither PR ($t(1331) = -.04$, $p = .971$) (Fig. 7b) nor GSR ($t(1331) = .47$, $p = .636$) (Fig. 7c) were predicted by clutter. However, clutter did exert a significant influence on joystick displacement ($t(1318) = -3.74$, $p < .001$), such that as clutter increased, joystick displacement decreased, and vice versa (Fig. 7a). In the context of the narrative, this negative relationship is most prominently illustrated in four areas. First, prior to trough 1, clutter decreases for 15 seconds, accompanied by a sharp rise in joystick displacement. This portion of the film consists of several cross-cuts between the scene in which the sheriff approaches the edge of the cliff and the flashback scene depicting the carriage accident. Directly following this segment, clutter increases as joystick displacement sharply declines for 20 seconds prior to peak

2. During this segment of the film, the boy recovers from the carriage accident and then runs to the edge of the cliff from which his father has just fallen. Third, prior to peak 3, clutter increases as joystick displacement again sharply decreases for 10 seconds. This segment consists entirely of a close-up shot of the father's stopwatch as it gradually ceases to tick. Finally, prior to peak 4, clutter again increases for 10 seconds, accompanied by a decrease in joystick displacement. During this segment, the sheriff crawls toward his father's stopwatch and holds it in his hands.

In sum, objective emotional tension was unaffected by clutter in *Borrowed Time*, while subjective emotional tension was significantly predicted by clutter in a negative direction.

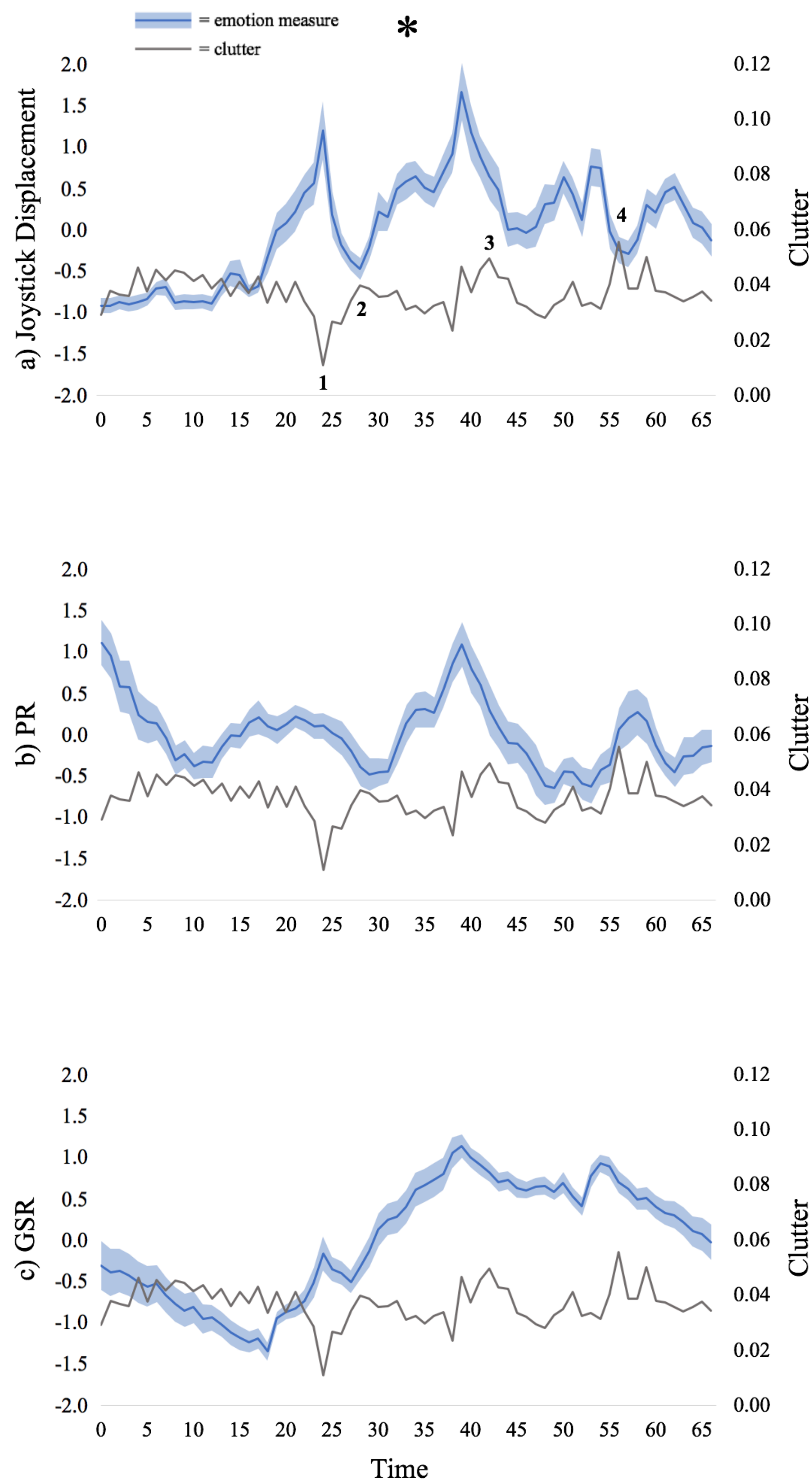

Fig. 7. Mean z-standardized emotion measures as a function of clutter for *Borrowed Time*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

Paperman. Clutter was not a significant predictor of joystick displacement ($t(1503) = -.99$, $p = .322$) (Fig. 8a), PR ($t(1504) = 1.32$, $p = .187$) (Fig. 8b), or GSR ($t(1511) = -.59$, $p = .558$) (Fig. 8c).

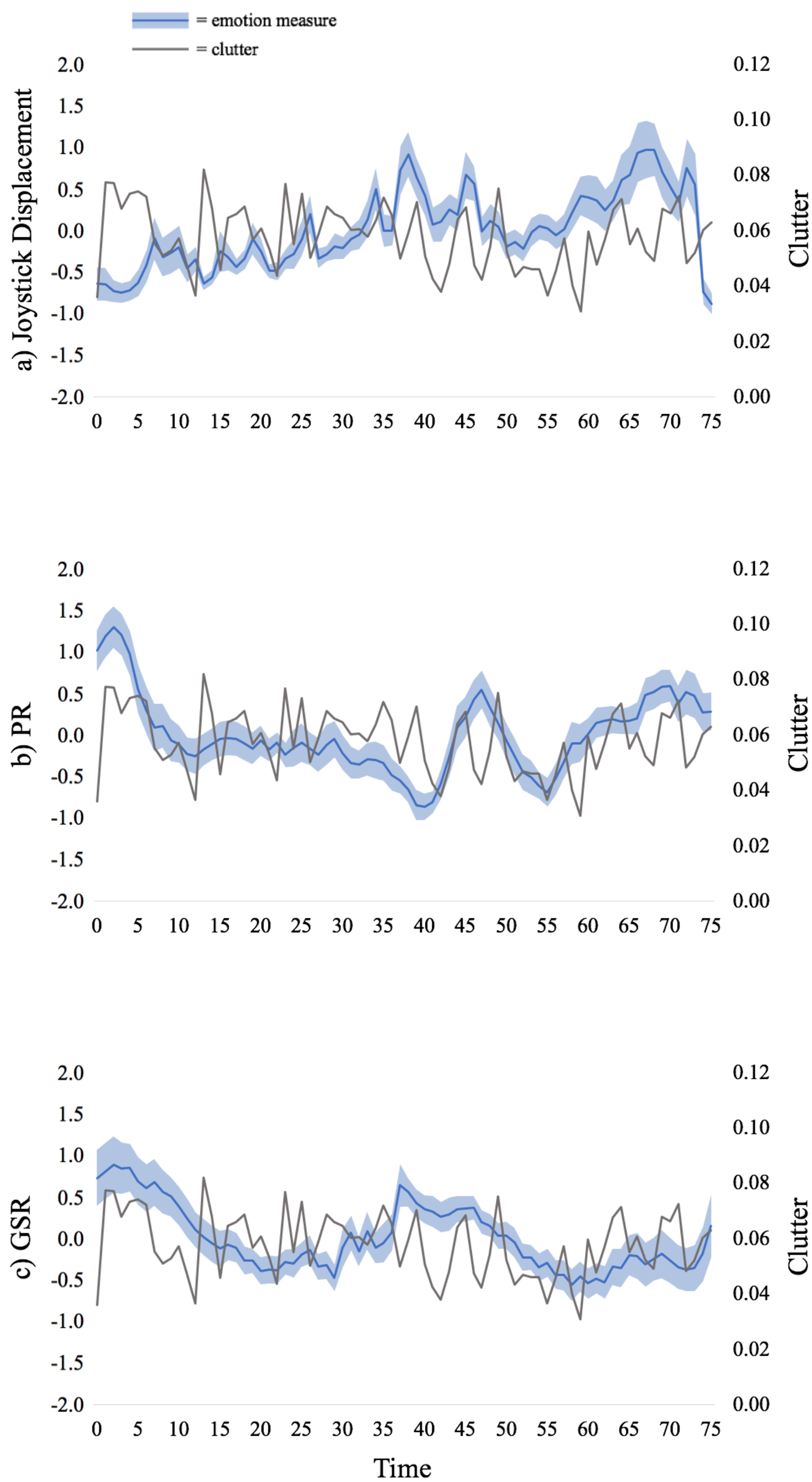

Fig. 8. Mean z-standardized emotion measures as a function of clutter for *Paperman*. Darker bands indicate 95% confidence intervals.

Stutterer. PR was predicted by clutter ($t(2931) = -3.88$, $p < .001$) (Fig. 9b), and so was GSR ($t(2931) = -5.76$, $p < .001$) (Fig. 9c). However, clutter did not exert a significant influence on joystick displacement ($t(2905) = -2.34$, $p = .020$) (Fig. 9a). A negative relationship was observed between both clutter and PR and clutter and GSR, as illustrated most prominently in three areas for each measure.

In terms of PR (Fig. 9b), clutter sharply decreases for 15 seconds prior to trough 1, accompanied by a rise in PR. During this segment of the film, Greenwood reads a message from his online love interest, Ellie, requesting to meet him in person for the first time. Second, prior to peak 2, clutter increases for 15 seconds as PR sharply declines. This portion of the film consists of Greenwood sitting at his workspace, thinking, and then getting up to retrieve his laptop. Finally, prior to trough 3, clutter decreases sharply for 10 seconds, accompanied by a rise in PR, during which Greenwood, having heard a notification on his laptop, rises from bed in the dark.

In terms of GSR (Fig. 9c), clutter sharply decreases for 10 seconds following peak 1, accompanied by a modest increase in GSR. This portion of the film consists of Greenwood working at his desk while he is on hold with his internet provider, culminating in a pan to his laptop following the sound of a notification. Second, prior to trough 2, clutter sharply decreases for 15 seconds as GSR rises. During this segment of the film, the screen goes black after Greenwood confronts a man at a bus stop who is assaulting his girlfriend. Finally, prior to trough 3, clutter again sharply decreases for 10 seconds as GSR rises. During this segment, Ellie has just glanced across the street to see Greenwood for the first time, and they tenderly hold each other's gaze.

In sum, subjective emotional tension was unaffected by clutter in *Stutterer*, while objective emotional tension was significantly predicted by clutter in a negative direction.

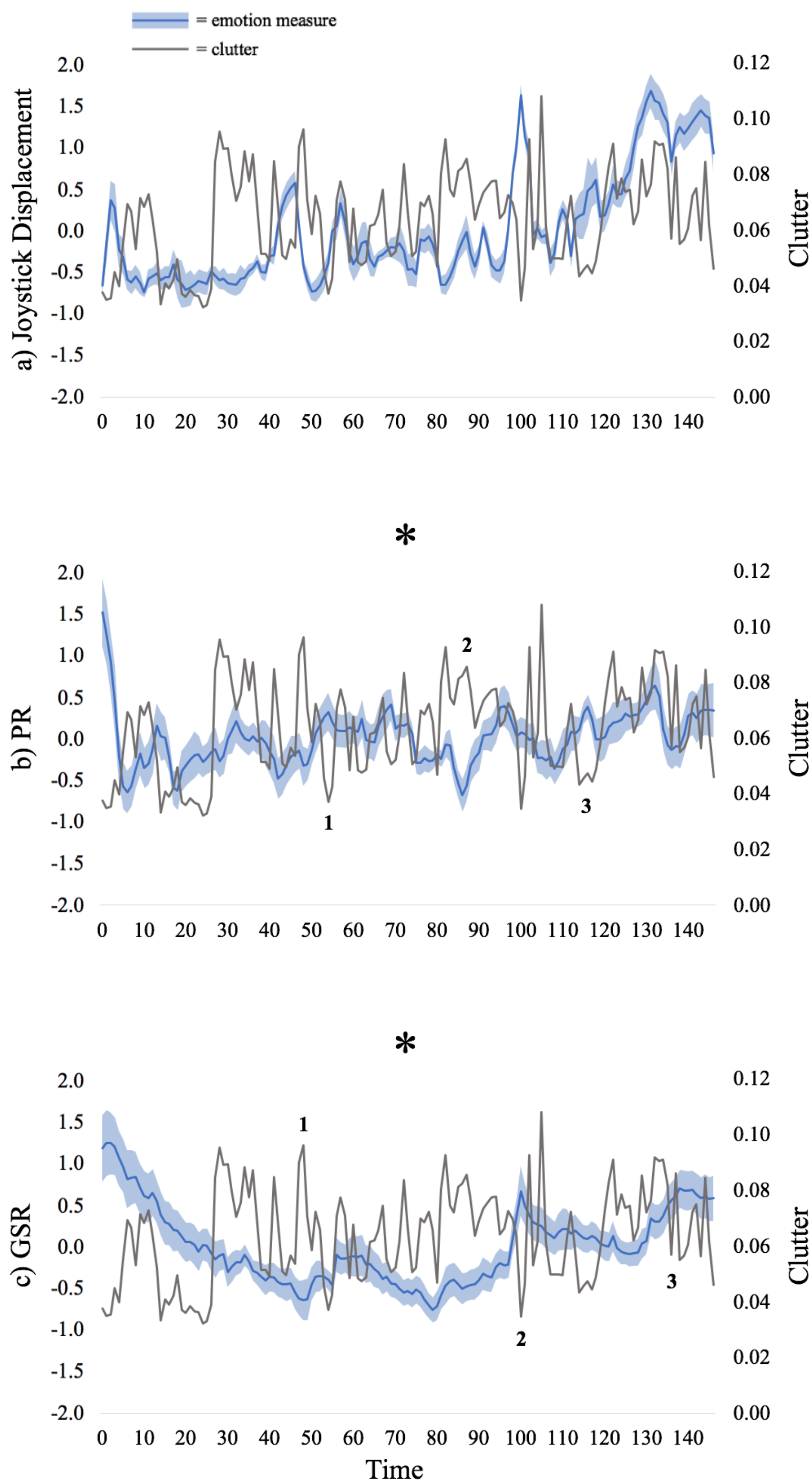

Fig. 9. Mean z-standardized emotion measures as a function of clutter for *Stutterer*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

The Voorman Problem. Magnitude of joystick displacement was not predicted by clutter ($t(2649) = 1.50$, $p = .133$) (Fig. 10a), and neither was PR ($t(2649) = -2.59$, $p = .010$) (Fig. 10b). However, clutter did exert a significant positive influence on GSR ($t(2651) = 9.31$, $p < .001$) (Fig. 10c). In the context of the narrative, this relationship is most prominently illustrated in two areas. First, prior to trough 1, clutter decreases for 30 seconds, accompanied by a steady decline in GSR. This portion of the film depicts Dr. Williams arriving at the prison and listening to faint chanting through the speaker on the front door. Second, prior to peak 2, clutter rises for 40 seconds, accompanied by a modest rise in GSR. During this segment of the film, Voorman has just switched places with Dr. Williams, who is now fighting against the straitjacket that holds him in place.

In sum, subjective emotional tension was unaffected by clutter in *The Voorman Problem*. Of the objective measures of emotion, GSR—but not PR—was significantly predicted by clutter in a positive direction.

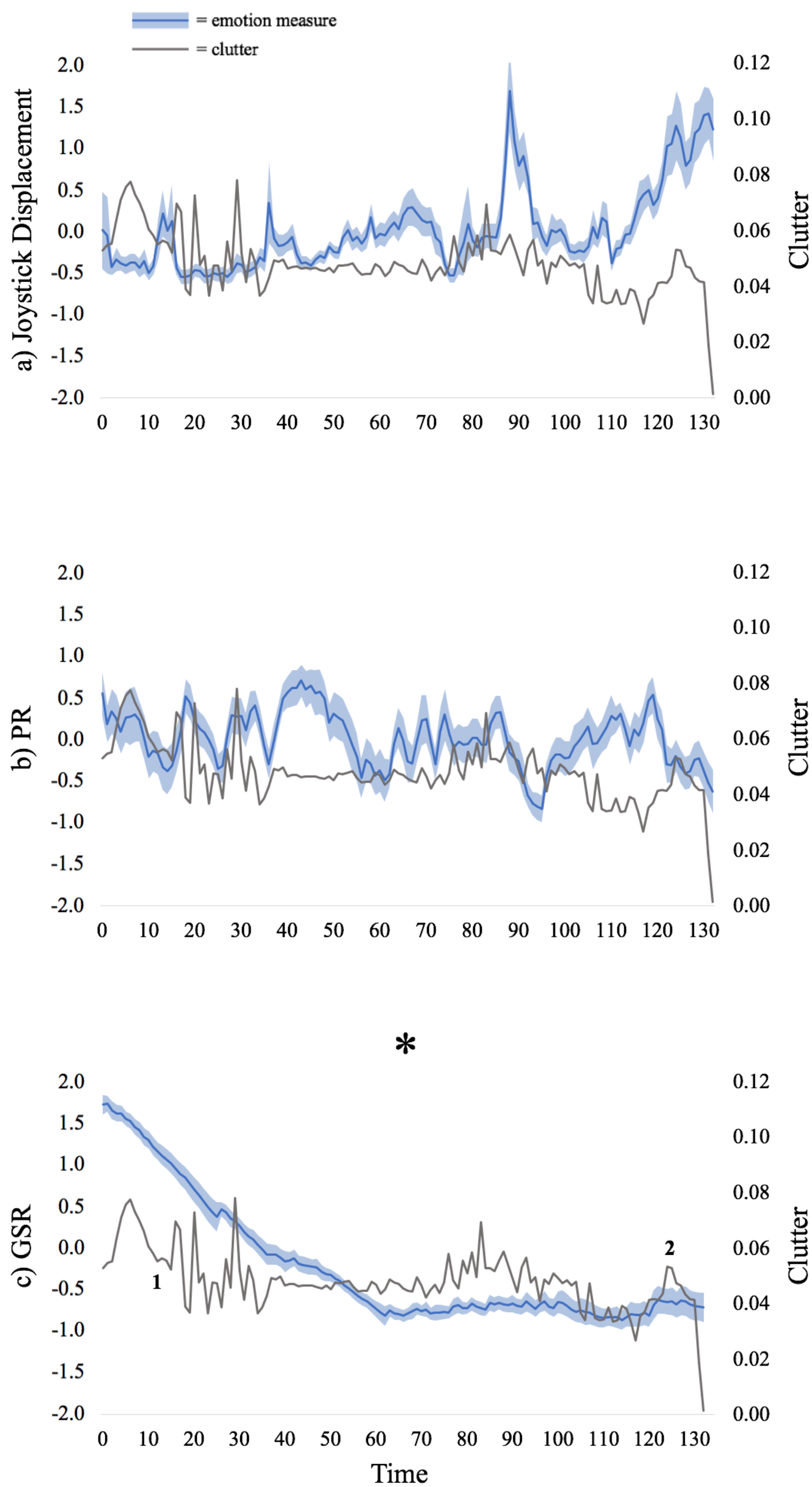

Fig. 10. Mean z-standardized emotion measures as a function of clutter for *The Voorman Problem*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

Interim discussion. Clutter did not uniformly predict emotional tension across films, and in *Paperman*, it did not predict emotional response at all (despite a relatively high degree of clutter in the film). Speculatively, it could be the case that the competing low-level features of *Paperman* (or elements of the narrative itself) diminished the potential effects of clutter on emotional response among viewers.

In contrast, clutter negatively predicted at least one emotion measure in both *Borrowed Time* and *Stutterer* (joystick displacement in the former; PR and GSR in the latter), such that as clutter increased, emotional tension decreased, and vice versa. These patterns may be the result of the perceptual effects of clutter: since clutter is known to interfere with visual search (e.g., Henderson, Chanceaux, & Smith, 2009), perhaps more cluttered shots distract viewers from extracting relevant emotional content from a given scene (e.g., a character's facial expressions), thus limiting emotional engagement. Indeed, Cutting and Armstrong (2016) found that greater clutter impedes the processing of facial expression valence in longer-scaled shots. However, an analysis of Hollywood film revealed that filmmakers accommodate the need for longer visual processing times by prolonging the duration of longer-scaled shots that contain a relatively high degree of clutter (Cutting & Armstrong, 2016). While this convention of filmmaking suggests that clutter should not interfere with emotional comprehension of cinematic storylines, short films may not benefit from the same techniques that have been cultivated in Hollywood feature films over the last several decades.

To corroborate this notion, I performed a linear regression across the four short films to assess how the interaction between shot scale and clutter predicted shot density (which I treated as a proxy for shot duration, given that lower values indicate shots of longer duration, while higher values indicate shots of shorter duration). Applying a stringent alpha criterion of .01, the expected interaction emerged for *Paperman* ($t(1519) = -7.60$, $p < .001$) and *The Voorman Problem* ($t(2659) = -2.71$, $p = .007$). In contrast, no significant interaction emerged for *Borrowed Time* ($t(1339) = -2.20$, $p = .028$) or *Stutterer* ($t(2939) = 1.44$, $p =.150$)—critically, these were the only two films in which clutter negatively predicted emotional tension, suggesting that a lack of accommodation for the prolonged processing times needed for longer-scaled, cluttered shots may at least partially explain the effects described above.

An alternative explanation for these effects is tied to an inherent linkage between clutter and shot scale. Due to camera mechanics, depth of focus typically decreases as shot scale increases toward close-up shots, resulting in a blurring of background details that effectively reduces clutter (see Cutting & Armstrong, 2016). As expected, shot scale negatively predicted clutter for *Paperman* ($t(1519) = -5.84$, $p < .001$), *Stutterer* ($t(2939) = -22.16$, $p < .001$), and *The Voorman Problem* ($t(2659) = -15.31$, $p < .001$), but not for *Borrowed Time* ($t(1339) = 1.97$, $p = .049$) (which could be explained by the fact that animated films are not constrained by the focus limitations that have historically plagued cameras). Thus, given that closer-scaled shots arguably increase emotional engagement (e.g., by fostering facial expression processing; Smith, 2012, 2013), what is ostensibly a negative relationship between emotional tension and clutter could instead be an artifact of a positive relationship between emotional tension and shot scale (particularly in the case of *Stutterer*, where closer-scaled shots happen to be less cluttered). However, an analysis of shot scale (described in greater detail below; see "Shot Scale") reveals that this explanation does not hold for the four films utilized in this study, thus bolstering the shot density explanation put forth above.

Finally, the *The Voorman Problem* stood out as the only film in which clutter predicted GSR in a positive direction, such that as clutter increased, so did emotional tension, and vice versa. This effect may be due to the ways in which filmmakers manipulate mise-en-scène—a term used in film theory to describe the visual theme of a story, as achieved primarily through set design and cinematography. Everything that appears before the camera (as well as their collective arrangement) drives the vision of a film by setting a particular mood and, critically, by suggesting a character's state of mind (Barsam & Monahan, 2010). For instance, a scene staged in an outrageously messy bedroom might reveal a character's mindset to be confused or incoherent, as if filming the character from the inside (McGrail, n. d.). Likewise, a scene in which a character moves through a vast pastoral landscape might imply a psychological sense of abandon, or—in a different story—an intense vulnerability, ultimately communicating something vital about the character without the need for exposition. As such, shots that contain a high degree of clutter could implicitly communicate a sense a psychological mayhem, thus driving an increase in emotional tension among viewers as they process a character's turmoil through the character's surroundings. Indeed, one might expect this theory to be most applicable to a film like *The Voorman Problem,* which specializes in mind games and psychological puzzlement.

Despite the appeal of this notion, it is important to note that the positive relationship between clutter and GSR among viewers of *The Voorman Problem* could also be caused by the highly elevated starting point of GSR, followed by a dramatic decline that persists over the first half of the film. This unusual pattern is potentially the result of competing low-level features (discussed in greater detail below; see "Sound Amplitude").

As a side note before discussing the remaining low-level features under investigation, I must acknowledge the dissociation between emotion measures that was revealed here. Clutter differentially affected opposing indices of emotion, such that subjective measures were impacted in one film, while objective measures were impacted in other films. As this phenomenon persists in a similar manner across all six low-level features, it will be broadly addressed in a separate section below (see "Correlation Among Emotion Measures").

Luminance

Fig. 11 displays the pattern of luminance over the course of each film. *Paperman* consists of the highest degree of luminance (M = 164.55, SD = 32.85), followed by *Stutterer* (M = 115.29, SD = 36.69). Comparatively, *The Voorman Problem* (M = 98.32, SD = 14.62) and *Borrowed Time* (M = 96.50, SD = 17.68) are notably darker, with roughly equal mean luminance values. The influence of luminance on emotional engagement among viewers of each film is described in greater detail below.

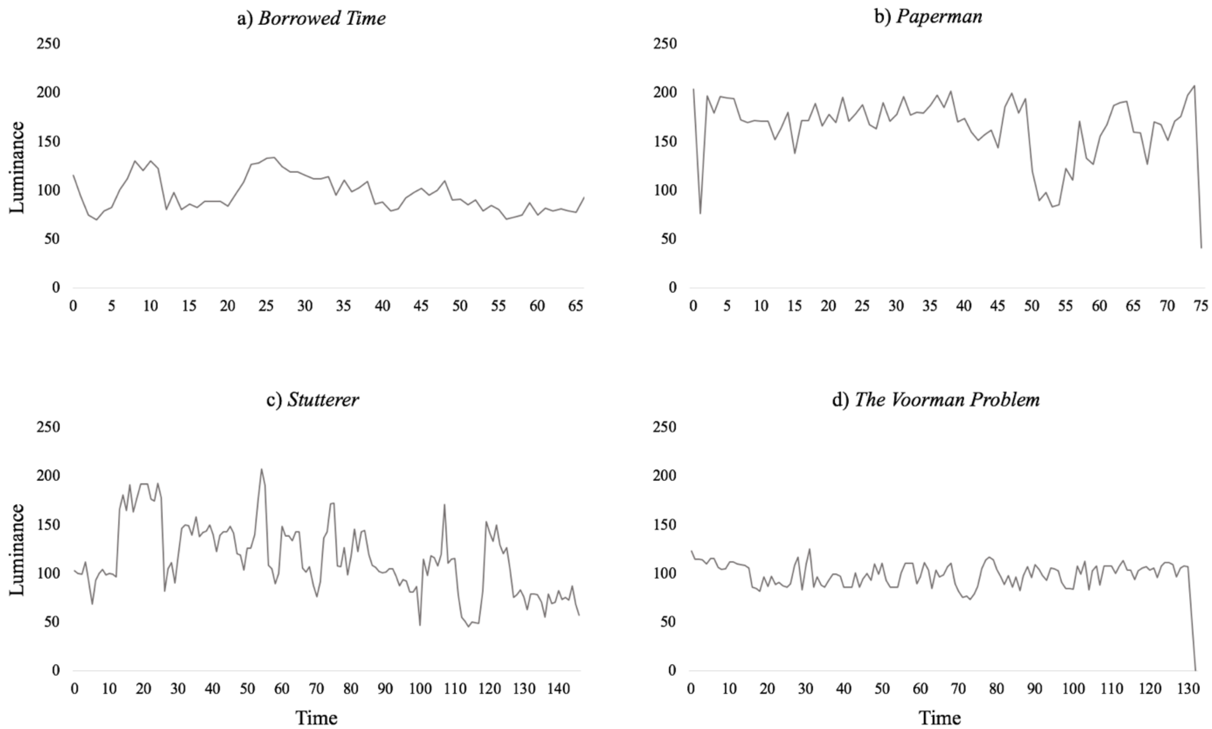

Fig. 11. Luminance over time in a) *Borrowed Time*, b) *Paperman*, c) *Stutterer*, and d) *The Voorman Problem*. Luminance units refer to mean grayscale pixel values ranging from 0 (black) to 255 (white), such that higher values indicate brighter frames across the preceding 5-second interval, while lower values indicate darker frames.

Borrowed Time. Luminance did not exert a significant influence on joystick displacement ($t(1318) = -.88$, $p = .380$) (Fig. 12a). However, PR was negatively predicted by luminance ($t(1331) = -4.83$, $p < .001$) and so was GSR ($t(1331) = -2.82$, $p = .005$), as illustrated most prominently in three areas (Fig. 12b and 12c). First, prior to peak 1, luminance increases for 35 seconds, accompanied by a steady decline in both PR and GSR. This portion of the film consists of the sheriff looking plaintively into the distance, followed by a flashback scene in which the boy's father hands him a pocket watch while they ride in a horse-drawn carriage through the desert. Second, prior to trough 2, luminance sharply declines while both PR and GSR rise for 5 seconds. This segment of the film consists of a single shot of the pocket watch after it has fallen and landed behind a rock. Finally, prior to trough 3, luminance decreases for 10 seconds while PR rises and GSR remains elevated, during which the sheriff crawls toward his father's stopwatch and holds it in his hands.

In sum, subjective emotional tension was unaffected by luminance in *Stutterer*, while objective emotional tension was negatively predicted by a change in luminance.

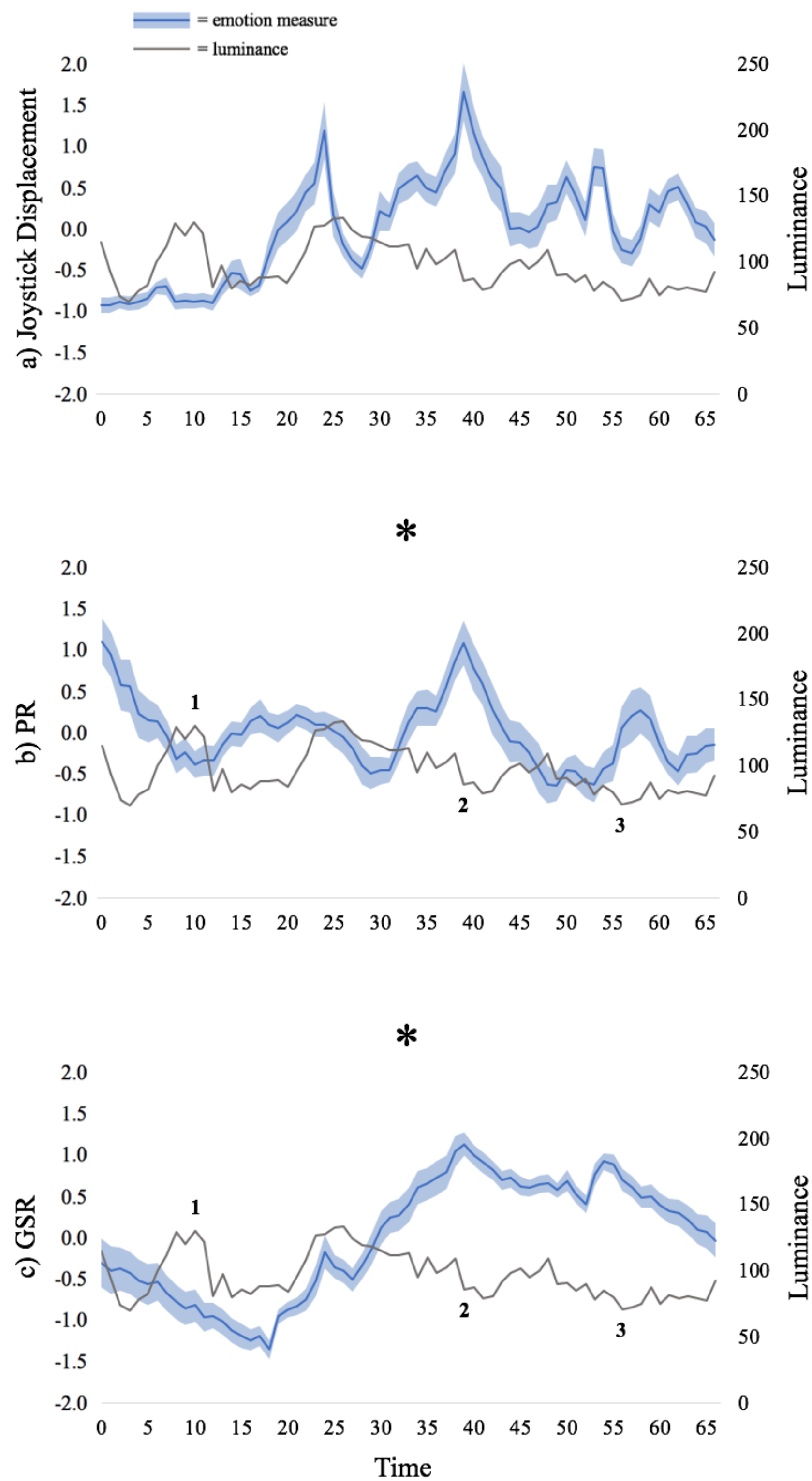

Fig. 12. Mean z-standardized emotion measures as a function of luminance for *Borrowed Time*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

Paperman. PR was not predicted by luminance (*t*(1504) = 1.61, *p* = .107) (Fig. 13b), and neither was GSR (*t*(1511) = -.04, *p* = .968) (Fig. 13c). However, luminance did exert a significant influence on joystick displacement (*t*(1503) = 3.42, *p* = .001), such that as luminance increased, so did joystick displacement, and vice versa (Fig. 13a). In the context of the narrative, this positive relationship is most prominently illustrated in three areas. First, prior to peak 1, luminance increases for 5 seconds, accompanied by a sharp rise in joystick displacement. This portion of the film consists of George's boss catching him off task, followed by George catching sight of Meg walking out of the building onto the street below. Second, prior to trough 2, luminance sharply declines as joystick displacement decreases for 10 seconds, during which George's paper airplane drifts down into a small alleyway, joining hundreds of other paper airplanes on the ground below. Finally, prior to trough 3, both luminance and joystick displacement sharply decline for 5 seconds as the film ends. This segment consists of still shots, overlaid on a black background, depicting George and Meg talking in a café.

In sum, objective emotional tension was unaffected by luminance in *Paperman*, while subjective emotional tension was positively predicted by a change in luminance.

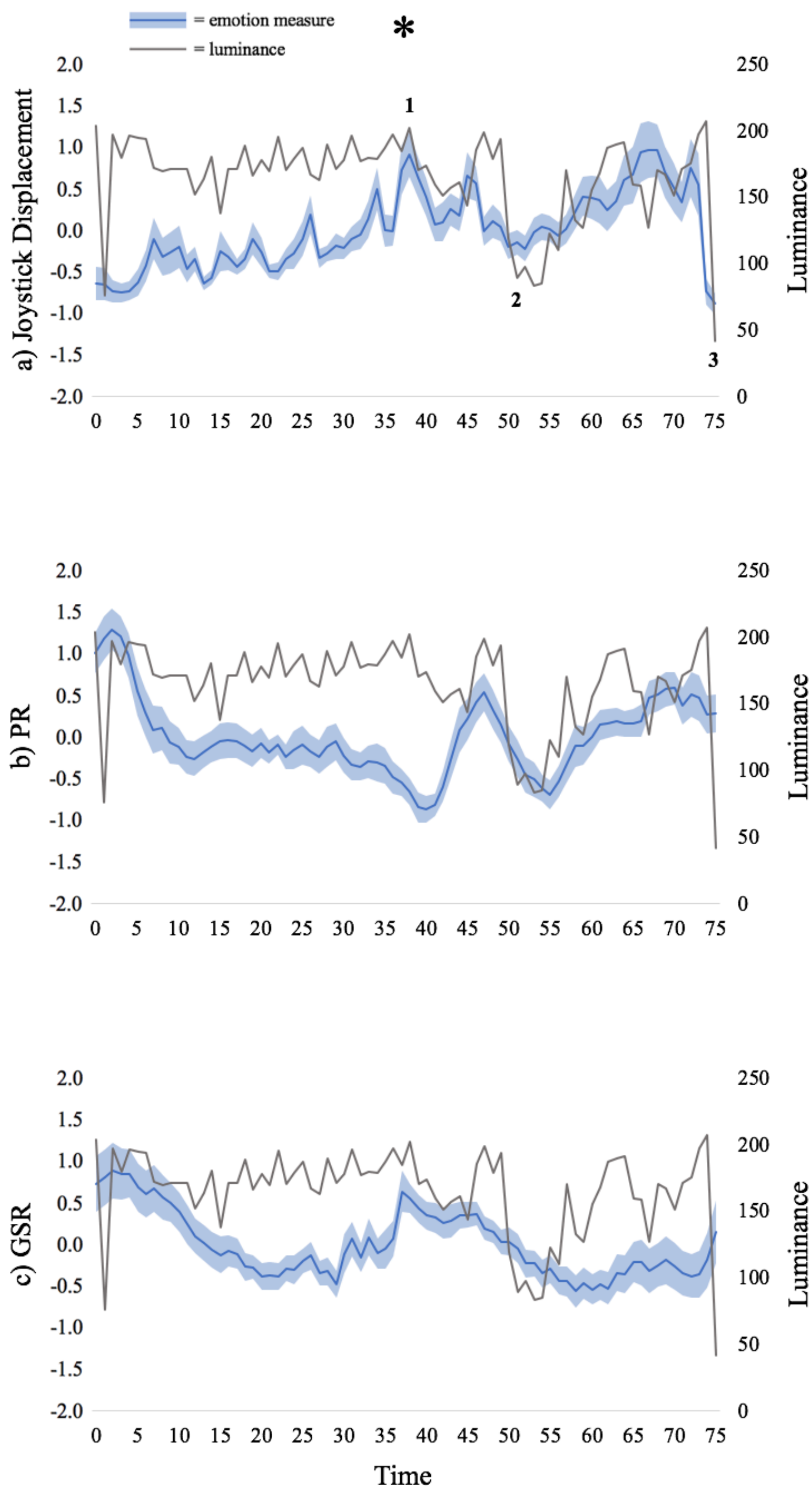

Fig. 13. Mean z-standardized emotion measures as a function of luminance for *Paperman*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

Stutterer. Luminance did not exert a significant influence on PR ($t(2931) = -2.54$, $p = .011$) (Fig. 14b). However, joystick displacement was negatively predicted by luminance ($t(2905) = -5.88$, $p < .001$) (Fig. 14a), and so was GSR ($t(2931) = -8.65$, $p < .001$) (Fig. 14c). In the context of the narrative, this relationship is most prominently illustrated in three areas (Fig. 12b and 12c). First, prior to trough 1, luminance sharply declines for 5 seconds, accompanied by a sharp increase in both joystick displacement and GSR. During this portion of the film, the screen goes black after Greenwood confronts a man at a bus stop who is assaulting his girlfriend. Second, prior to peak 2, luminance sharply increases while both joystick displacement and GSR decrease for 10 seconds. This segment consists of Greenwood sitting in his bedroom, despondent, as he checks his laptop for messages from Ellie. Finally, prior to trough 3, luminance decreases for 10 seconds while both joystick displacement and GSR increase, during which Greenwood looks across the street to spot Ellie for the first time.

In sum, subjective emotional tension was significantly predicted by a change in luminance in *Stutterer*. Of the objective measures of emotion, GSR—but not PR—was also predicted by a change in luminance. Both effects revealed a negative relationship.

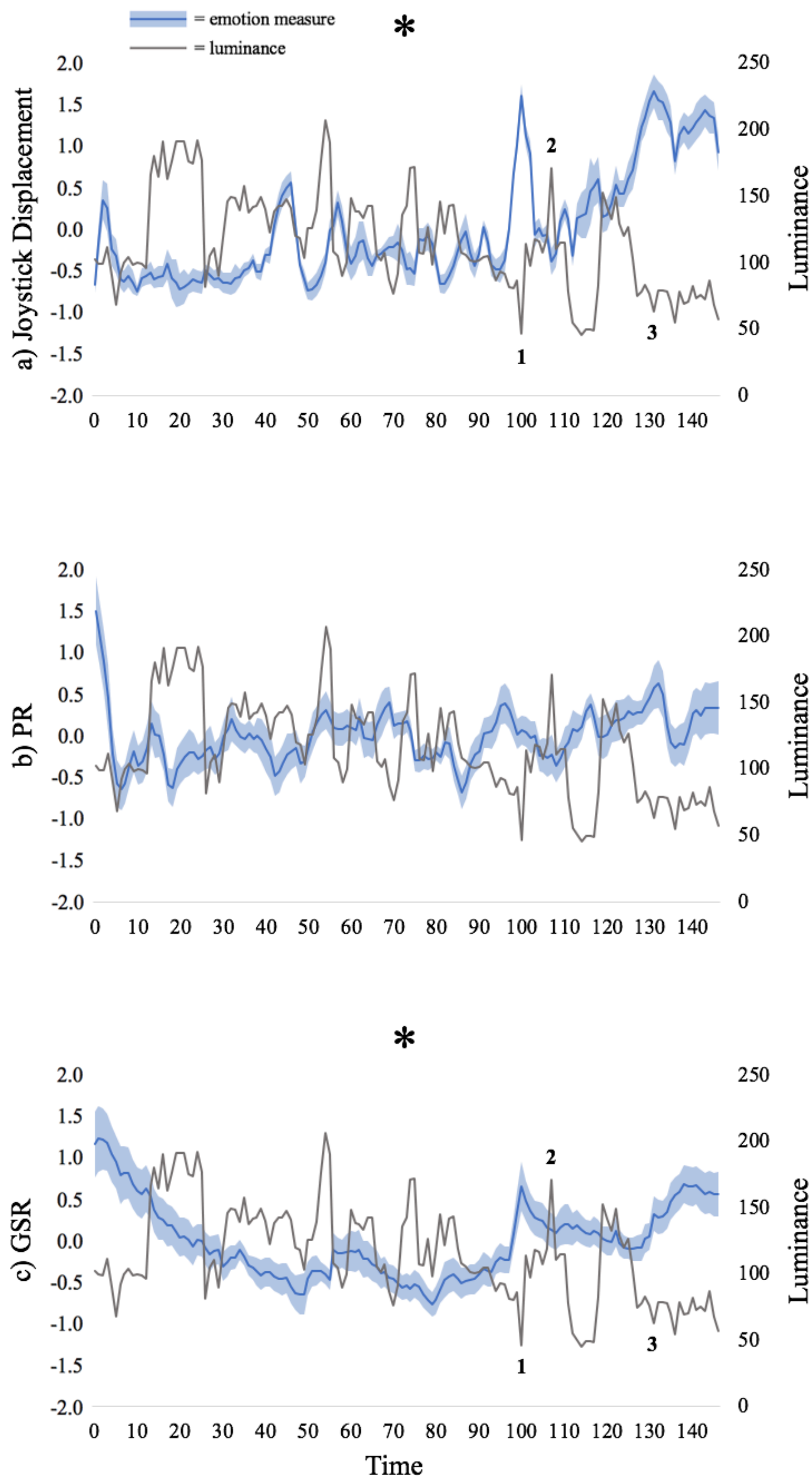

Fig. 14. Mean z-standardized emotion measures as a function of luminance for *Stutterer*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

The Voorman Problem. Magnitude of joystick displacement was not predicted by luminance ($t(2649) = .47$, $p = .637$) (Fig. 15a), and neither was PR ($t(2649) = .59$, $p = .553$) (Fig. 15b). However, luminance did exert a significant influence on GSR in a positive direction ($t(2651) = 4.86$, $p < .001$) (Fig. 15c). In the context of the narrative, this relationship is most prominently illustrated in two areas. First, prior to trough 1, luminance decreases for 15 seconds, accompanied by a steady decline in GSR. This portion of the film depicts Dr. Williams meeting with Governor Bentley in the latter's office, where the governor pours liquor into his coffee cup. Second, prior to trough 2, luminance declines for 5 seconds, accompanied again by a steady decline in GSR. During this segment of the film, Dr. Williams asks about Voorman's conviction, and Governor Bentley tells him that the computer systems crashed last year (implying that Voorman's crime is unknown).

In sum, subjective emotional tension was unaffected by luminance in *The Voorman Problem*. Of the objective measures of emotion, GSR—but not PR—was positively predicted by a change in luminance.

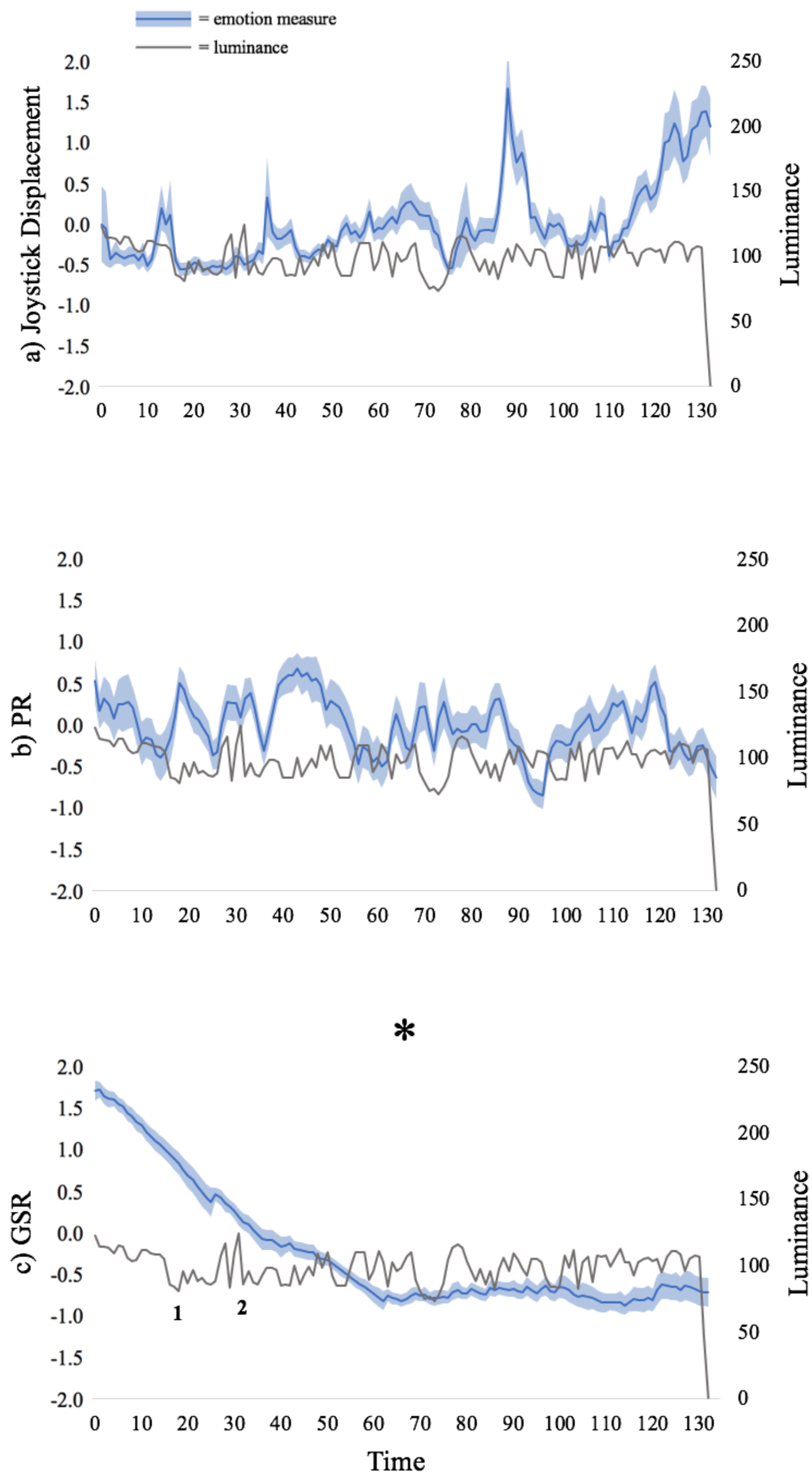

Fig. 15. Mean z-standardized emotion measures as a function of luminance for *The Voorman Problem*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

Interim discussion. As tentatively predicted, a negative relationship emerged between luminance and emotional tension in *Borrowed Time* and *Stutterer*, such that as luminance decreased, emotional tension increased, and vice versa. These findings reflect the emotion-modulating effects of brightness in both experimental and applied settings: brightness predicts pleasure when viewing a series of color swatches (Valdez and Mehrabian, 1994), as well as the perceived valence of a film scene (Tarvainen, Westman, & Oittinen, 2015), and light therapy results in improved mood among patients with depression (e.g., McEnany & Lee, 2005; Pinchasov, Shurgaja, Grischin, & Putilov, 2005).

The history of film is replete with examples of filmmakers harnessing the emotional impact of brightness. Most prominently, a lighting technique known as chiaroscuro was popularized in cinema nearly a century ago. Chiaroscuro originated with Renaissance painting and describes the use of strong contrasts between light and dark on canvas, often to emphasize solidarity of form (Hall, 1987). In cinematography, the term seamlessly translates to the contrast between light and dark on the movie screen. Works of art that make intentional use of this technique tend to be darker overall, using light conservatively to focus the viewer's attention on certain features of a scene or portions of a character's face. The purpose of this practice is to obscure the overall image, thus creating suspense ("What Is Chiaroscuro in Film?" 2019)—which may explain why modern horror films have adapted the practice (see Brunick et al., 2013).

While chiaroscuro is most prevalent in the black-and-white films of German Expressionism and Hollywood Film Noir (movements that, in different stylistic ways, emphasize a character's inner desires and fears, which are said to reflect the collective social mindset of a given era), it has persisted throughout the decades following, emerging as a defining feature of films such as *The Godfather* (Coppola, 1972), *Casino Royale* (Campbell, 2006), and *There Will*

Be Blood (Anderson, 2007) (see "What Is Chiaroscuro in Film?" 2019). While the technique may not be obviously present in the average contemporary film, its qualities are resurrected in the form of diminishing luminance: Hollywood films have gradually grown darker over the past several decades (Cutting, Brunick, DeLong, et al., 2011). By manipulating brightness in this manner, filmmakers arguably evoke the same emotional impact that sustained the popularity of past film movements aesthetically defined by chiaroscuro. In other words, by shrouding the screen in darkness, filmmakers implicitly create a mood of heightened tension.

In sum, given empirical evidence that brightness elevates mood, as well as a robust history of filmmakers forging entire cinematic movements that have been aesthetically defined by brightness and contrast, it follows that a change in luminance should predict emotional engagement among film viewers. Indeed, of the six low-level features explored in this study, luminance boasts the highest number of significant effects (six out of a possible twelve). Four of those effects occurred in the expected direction: for both *Borrowed Time* and *Stutterer*, darker images heightened emotional tension while brighter images tempered it.

However, two effects occurred in the opposite direction: relatively brighter images were associated with increased joystick displacement for *Paperman* and increased GSR for *The Voorman Problem*. The former is potentially explained by a trend in animated film. In contrast to live action films, which have grown darker over time, animated films intended for children have maintained a steady level of brightness (Brunick & Cutting, 2014). This trend is reflected in *Paperman*, which is by far the brightest of the four films utilized in this study ($M = 164.55$, $SD = 32.85$). Speculatively, it could be the case that baseline brightness exerts a differential effect on viewers, such that for relatively darker films, increased darkness drives emotional tension, and for relatively brighter films, increased brightness has the same effect.

On the other hand, the positive relationship between luminance and GSR in *The Voorman Problem* cannot be explained by a relatively high degree of luminance. In fact, *The Voorman Problem* is the second darkest of the four films (M = 98.32, SD = 14.62). It is substantially darker than *Paperman*, and it is also darker than *Stutterer,* which is marked by the expected negative relationship between luminance and emotional tension. However, as noted in the discussion regarding clutter, this positive relationship could be caused by the highly elevated starting point of GSR among viewers of *The Voorman Problem.*

An alternative explanation could account for the positive relationship observed between luminance and emotional tension for both *Paperman* and *The Voorman Problem.* Given that joystick displacement and GSR collapse emotion into a single, unidirectional dimension (thus eliminating the ability to detect the valence of emotional response), the nuance underlying the emotional impact of luminance may not be fully captured by these measures (i.e., since brightness elevates mood while darkness lowers it). Indeed, given their respective plotlines, the emotional tension experienced by viewers of *Borrowed Time* and *Stutterer* may have been more negative in valence than the emotional tension experienced by viewers of *Paperman* and *The Voorman Problem*. Specifically, viewers of the latter may have experienced the relatively positive emotions of, respectively, relief and tenderness for George as he pursues and ultimately catches up to Meg, and bemusement at the incredulity of Voorman switching bodies with Dr. Williams. Critically, emotional tension does rise around these plot points, as does luminance (though more so for *Paperman*), suggesting that increased luminance, as expected, might contribute to an elevation in emotional state. In sum, when looking more closely at the context of each film, the effects captured in these analyses might indeed reflect the expected direction of the relationship between a change in luminance and emotional response.

Motion

Fig. 16 displays the pattern of motion over the course of each film. By far, *Paperman* consists of the highest degree of motion (M = .073, SD = .073). Comparatively, *Stutterer* (M = .013, SD = .016) and *Borrowed Time* (M = .008, SD = .012) consist of very little motion, with the latter showing just a few minor peaks. *The Voorman Problem* is relatively flat (M = .003, SD = .004). The influence of motion on emotional engagement among viewers of each film is described in greater detail below.

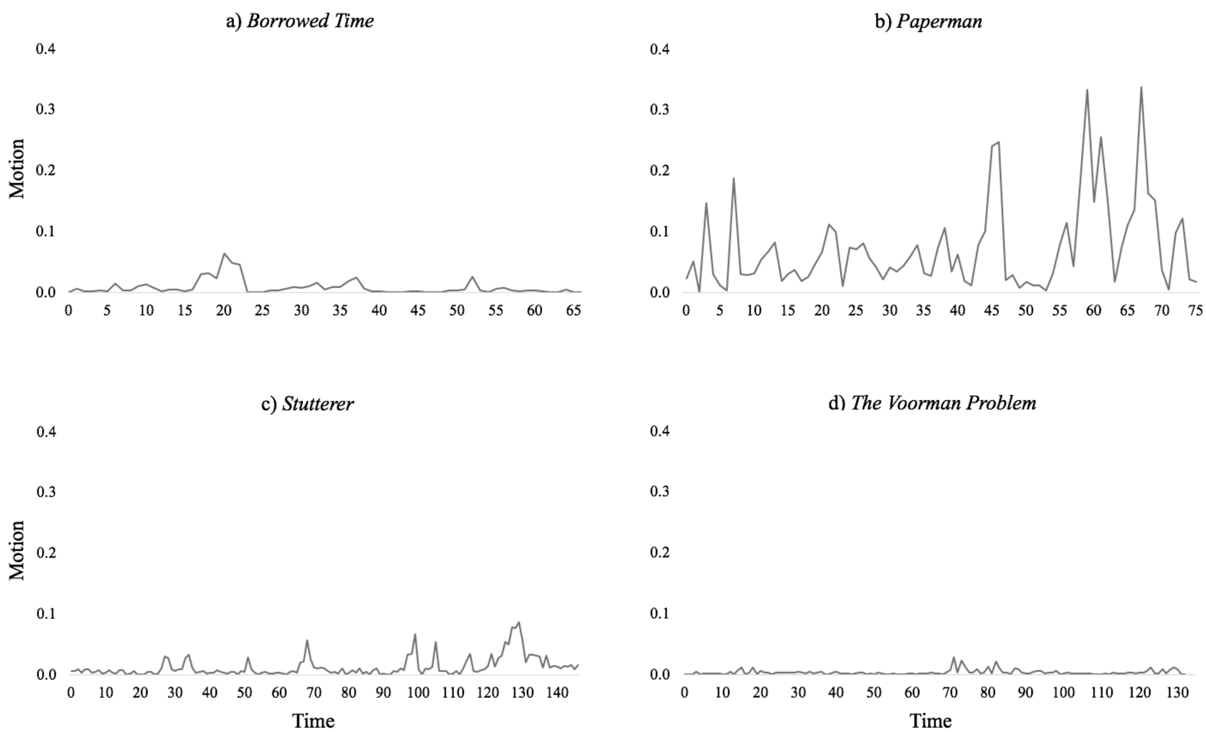

Fig. 16. Motion over time in a) *Borrowed Time*, b) *Paperman*, c) *Stutterer*, and d) *The Voorman Problem*. Motion units refer to 1-r, such that higher values indicate more motion across the preceding 5-second interval, while lower values indicate less motion.

Borrowed Time. Magnitude of joystick displacement was not predicted by motion ($t(1318) = .60$, $p = .549$) (Fig. 17a), and neither was PR ($t(1331) = -1.63$, $p = .103$) (Fig. 17b). However, motion did exert a significant influence on GSR ($t(1331) = -8.17$, $p < .001$), such that as motion decreased, GSR increased, and vice versa (Fig. 17c). In the context of the narrative, this negative relationship is most prominently illustrated in three areas. First, following peak 1, motion decreases for 15 seconds, accompanied by a sharp increase in GSR. This portion of the film consists primarily of close-up shots of the carriage accident. Second, following peak 2, motion again decreases for 20 seconds while GSR rises. This segment is comprised of close-up shots of the father's stopwatch—first, at it flies through the air and bounces on the ground, then, as it rests among rubble and a slow zoom captures the gradual cessation of its ticking. Finally, following peak 3, motion decreases for 10 seconds as GSR sharply increases, during which the sheriff struggles to climb back onto the cliff from which he has just slipped.

In sum, subjective emotional tension was unaffected by motion in *Borrowed Time*. Of the objective measures of emotion, GSR—but not PR—was significantly predicted by motion in a negative direction.

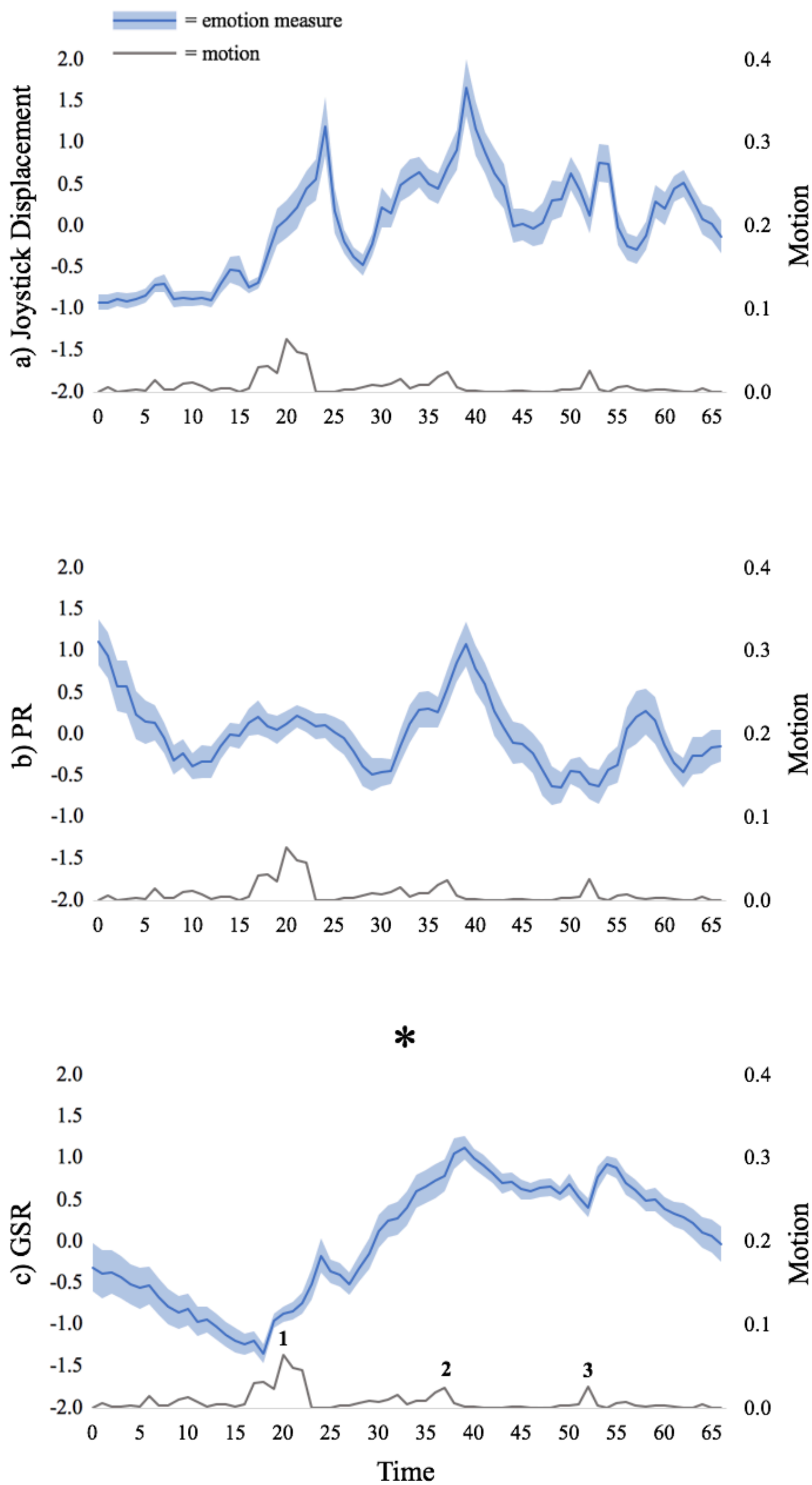

Fig. 17. Mean z-standardized emotion measures as a function of motion for *Borrowed Time*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

Paperman. Motion was not a significant predictor of joystick displacement ($t(1503) = 1.71$, $p = .087$) (Fig. 18a), PR ($t(1504) = .91$, $p = .365$) (Fig. 18b), or GSR ($t(1511) = -1.75$, $p = .080$) (Fig. 18c).

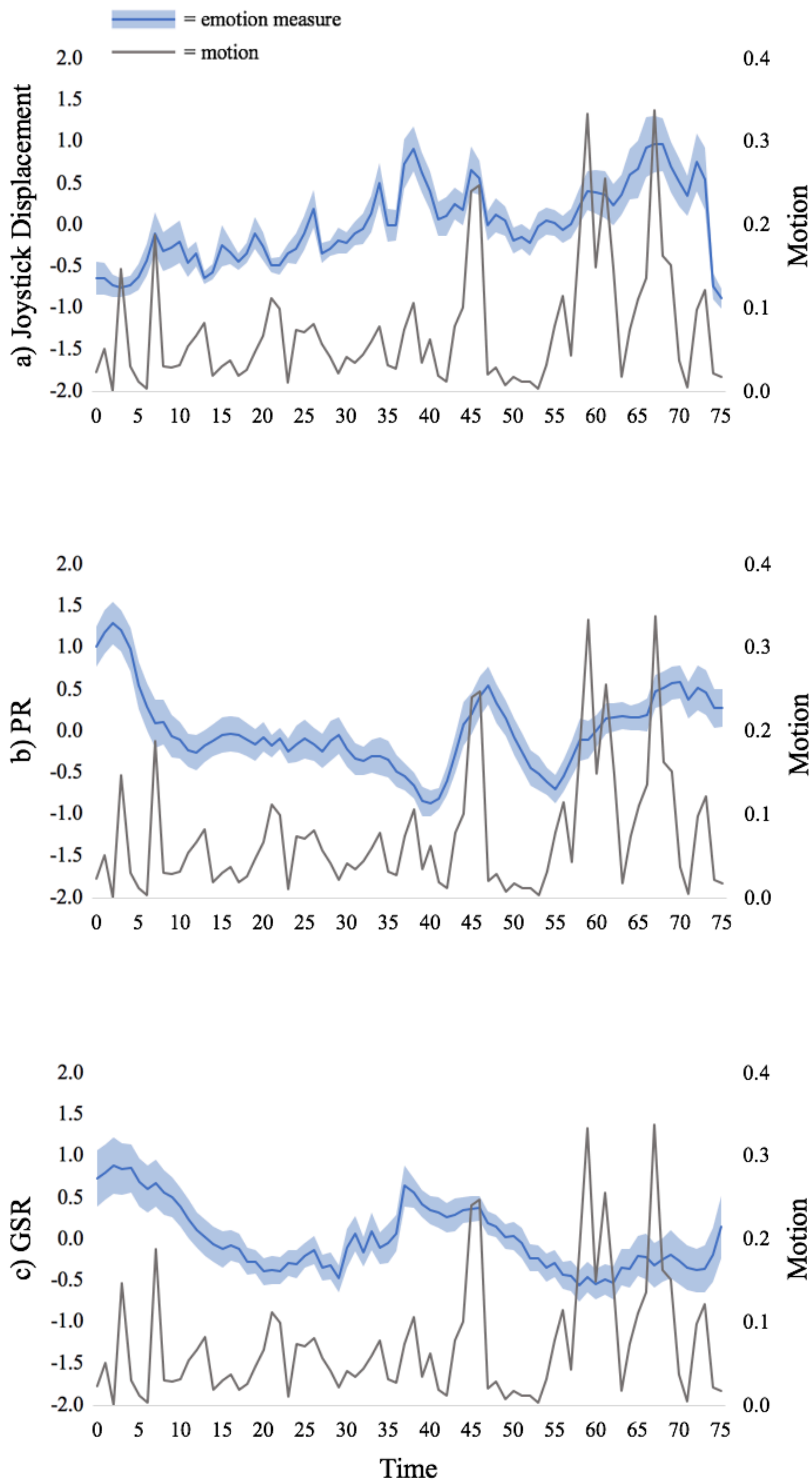

Fig. 18. Mean z-standardized emotion measures as a function of motion for *Paperman*. Darker bands indicate 95% confidence intervals.

Stutterer. Magnitude of joystick displacement was predicted by motion ($t(2905) = 3.09$, $p = .002$) (Fig. 19a), and so was PR (Fig. 19b) ($t(2931) = 2.65$, $p = .008$). However, motion did not exert a significant influence on GSR ($t(2931) = .37$, $p = .712$) (Fig. 19c). A positive relationship was observed between motion and joystick displacement and motion and PR, as illustrated most prominently in three areas (Fig. 19a and 19b). First, prior to peak 1, motion gradually increases for about 15 seconds, accompanied by a rise in both joystick displacement and PR. This portion of the film consists of the scene in which Greenwood, while riding the train, observes an attractive woman sitting nearby and mentally speculates about her life. Second, prior to peak 2, motion gradually increases for about 15 seconds, accompanied by a sharp rise in joystick displacement. During this portion of the film, Greenwood interferes with a man who is physically assaulting his girlfriend at a bus stop. While PR decreases during this segment, it continues to decrease immediately following peak 2, where motion also decreases for about 10 seconds, during which the aftermath of Greenwood's bus stop debacle is revealed—a bandaged nose. Finally, prior to peak 3, motion rises (albeit jaggedly) for about 35 seconds, accompanied by a steep rise in joystick displacement and a less steep, yet steady, rise in PR. This portion of the film consists of Greenwood's meandering and anxiety-ridden journey to the location where he has arranged to meet Ellie in person for the first time.

In sum, subjective emotional tension was significantly predicted by motion in *Stutterer*. Of the objective measures of emotion, PR—but not GSR—was also predicted by motion. Both effects revealed a positive relationship.

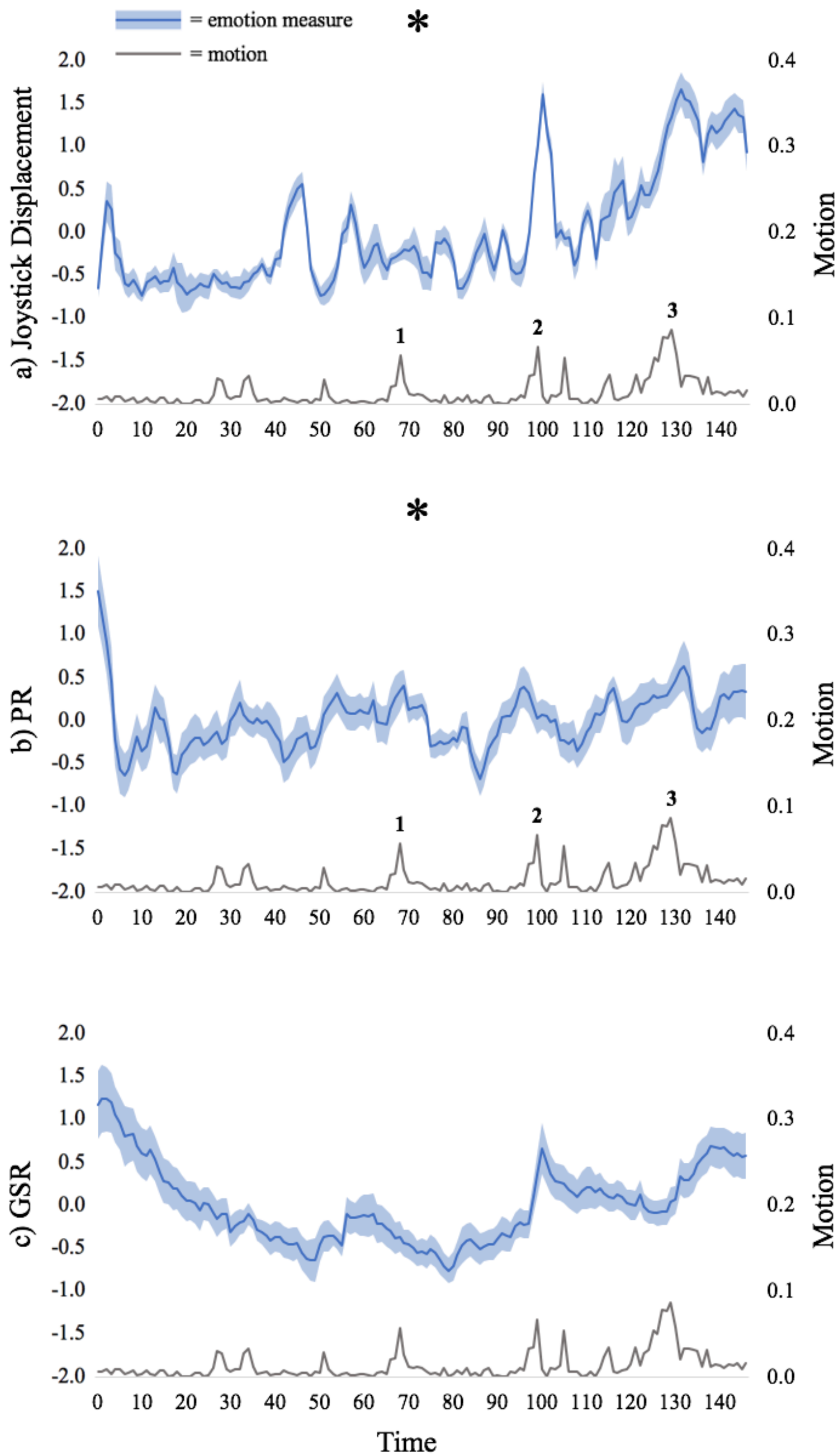

Fig. 19. Mean z-standardized emotion measures as a function of motion for *Stutterer*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

The Voorman Problem. Motion was not a significant predictor of joystick displacement ($t(2649) = 2.31$, $p = .021$) (Fig. 20a), PR ($t(2649) = .60$, $p = .547$) (Fig. 20b), or GSR ($t(2651) = .58$, $p = .565$) (Fig. 20c).

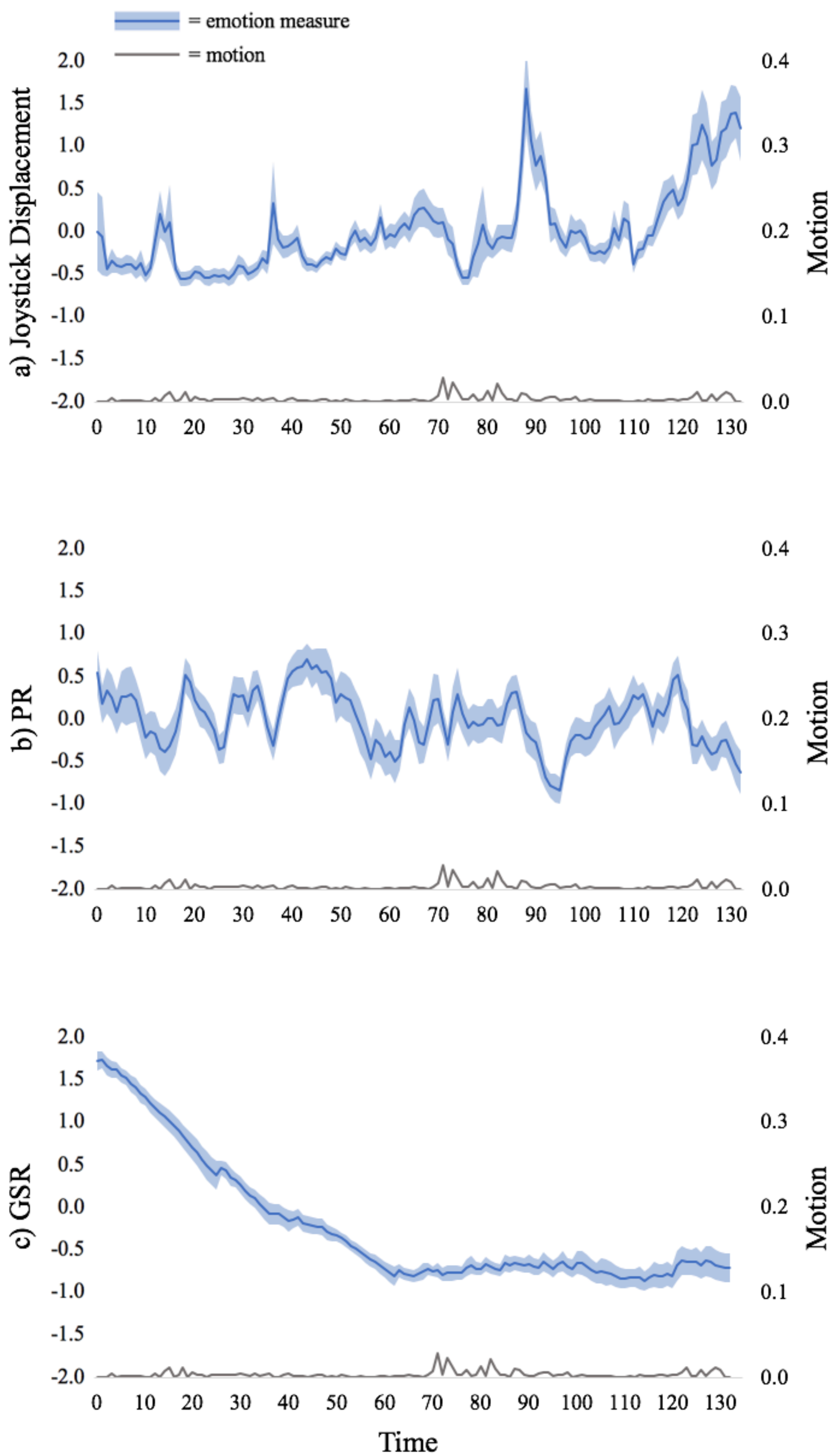

Fig. 20. Mean z-standardized emotion measures as a function of motion for *The Voorman Problem*. Darker bands indicate 95% confidence intervals.

Interim discussion. Like both clutter and luminance, the influence of motion was not uniform across films. Notably, in both *Paperman* and *The Voorman Problem,* motion did not predict emotional response at all. This is perhaps unsurprising in terms of *The Voorman Problem*, given the very low degree of motion over the course of the film. *Paperman*, on the other hand, consisted of a high degree of motion (aligning with previous findings regarding the difference in motion between animated and live action films; Cutting, 2014). Again, it could be the case that competing low-level features of *Paperman* (or elements of the narrative itself) diminished the potential effects of motion on emotional engagement.

In *Borrowed Time*, the direction of the relationship between motion and GSR was opposite my tentative prediction that higher motion would drive greater emotional response among viewers. A narrative analysis of this pattern revealed that peaks in GSR occurred during emotionally fraught scenes in which motion decreased, including the carriage accident scene, the immediate aftermath of the protagonist's father's death, and the protagonist's struggle to climb back over a cliff from which he had just slipped. Thus, perhaps the narrative elements of these scenes drove emotional response regardless of their low-level features.

On the other hand, in *Stutterer*, both joystick displacement and PR were influenced by motion, and in some cases, this relationship manifested in scenes that were less emotionally fraught. For instance, one such peak occurred during the scene in which Greenwood mentally speculates about the life of a woman sitting near him on the train. While this scene held no particular narrative punch (serving as more of a pause to provide insight into Greenwood's inner thoughts), the style in which it was filmed may have contributed to heightened emotional response. Specifically, the scene was shot using a handheld camera, which results in a "shaky" style of filmmaking, thus increasing the measure of motion. Incidentally, in the realm of film studies, this technique is known to shrink the distance between a film and its audience (in a pseudo-documentary manner) by conveying the impression of presence in cinematic action (see Pandža, 2018)—indeed, this explanation hints at the psychological mechanisms that might bridge cinematic motion and narrative engagement among viewers.

In contrast, the sequence in which Greenwood meanders toward the location where he has agreed to meet Ellie is filmed in an entirely different manner—yet, a manner which also increases the degree of cinematic motion. Specifically, this sequence consists of six jump cuts (an editing technique in which two sequential shots of the same action are recorded from camera positions that vary only slightly), which theoretically serve to increase emotional impact by amplifying the density of emotional gestures (Hogan, 2007). Thus, it appears that filmmakers employ a variety of stylistic techniques that increase cinematic motion, but that these techniques may differentially affect opposing measures of emotion.

Shot density

Fig. 21 displays the pattern of shot density over the course of each film. *Paperman* consists of the highest mean shot density (M = 1.88, SD = 1.31), followed by *Borrowed Time* (M = 1.81, SD = 1.81). Comparatively, *Stutterer* (M = 1.10, SD = .98) and *The Voorman Problem* (M = .68, SD = .70) are marked by a lower mean shot density, with the latter never exceeding two shot transitions per 5-second interval. The influence of shot density on emotional engagement among viewers of each film is described in greater detail below.

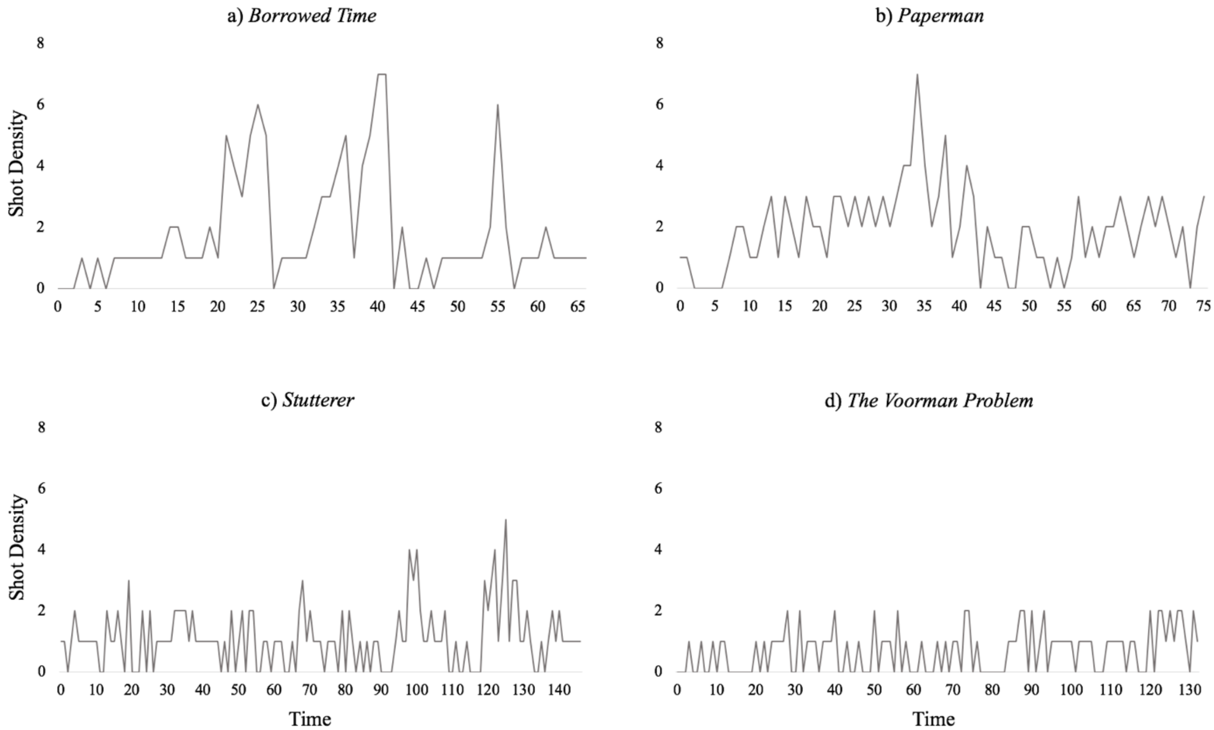

Fig. 21. Shot density over time in a) *Borrowed Time*, b) *Paperman*, c) *Stutterer*, and d) *The Voorman Problem*. Shot density units refer to the number of shot transitions (i.e., cuts) occurring within each 5-second interval, such that higher values indicate a quicker shot pace across the preceding 5-second interval, while lower values indicate a slower shot pace.

Borrowed Time. Magnitude of joystick displacement (Fig. 22a) was predicted by shot density ($t(1318) = 7.52$, $p < .001$), and so was PR ($t(1331) = 5.18$, $p < .001$) (Fig. 22b). However, shot density did not exert a significant influence on GSR (Fig. 22c) ($t(1331) = 2.29$, $p = .022$). A positive relationship was observed between shot density and joystick displacement and shot density and PR, as illustrated most prominently in three areas (Fig. 22a and 22b). First, prior to trough 1, shot density sharply decreases for 10 seconds, accompanied by a steep decline in joystick displacement and a modest decline in PR. During this portion of the film, the boy recovers from the carriage accident and then runs to the edge of the cliff from which his father has just fallen. Second, prior to peak 2, shot density sharply increases for 15 seconds, accompanied by a steep rise in both joystick displacement and PR. During this segment, the boy accidentally shoots his father with a rifle, then the father's pocket watch flies through the air and lands on the ground nearby, splattered with blood. Directly following this scene, shot density sharply decreases for 10 seconds prior to trough 3, during which time both joystick displacement and PR decrease. This portion of the film consists entirely of a close-up shot of the father's stopwatch as it gradually ceases to tick.

In sum, subjective emotional tension was significantly predicted by shot density in *Borrowed Time*. Of the objective measures of emotion, PR—but not GSR—was also predicted by shot density. Both effects revealed a positive relationship.

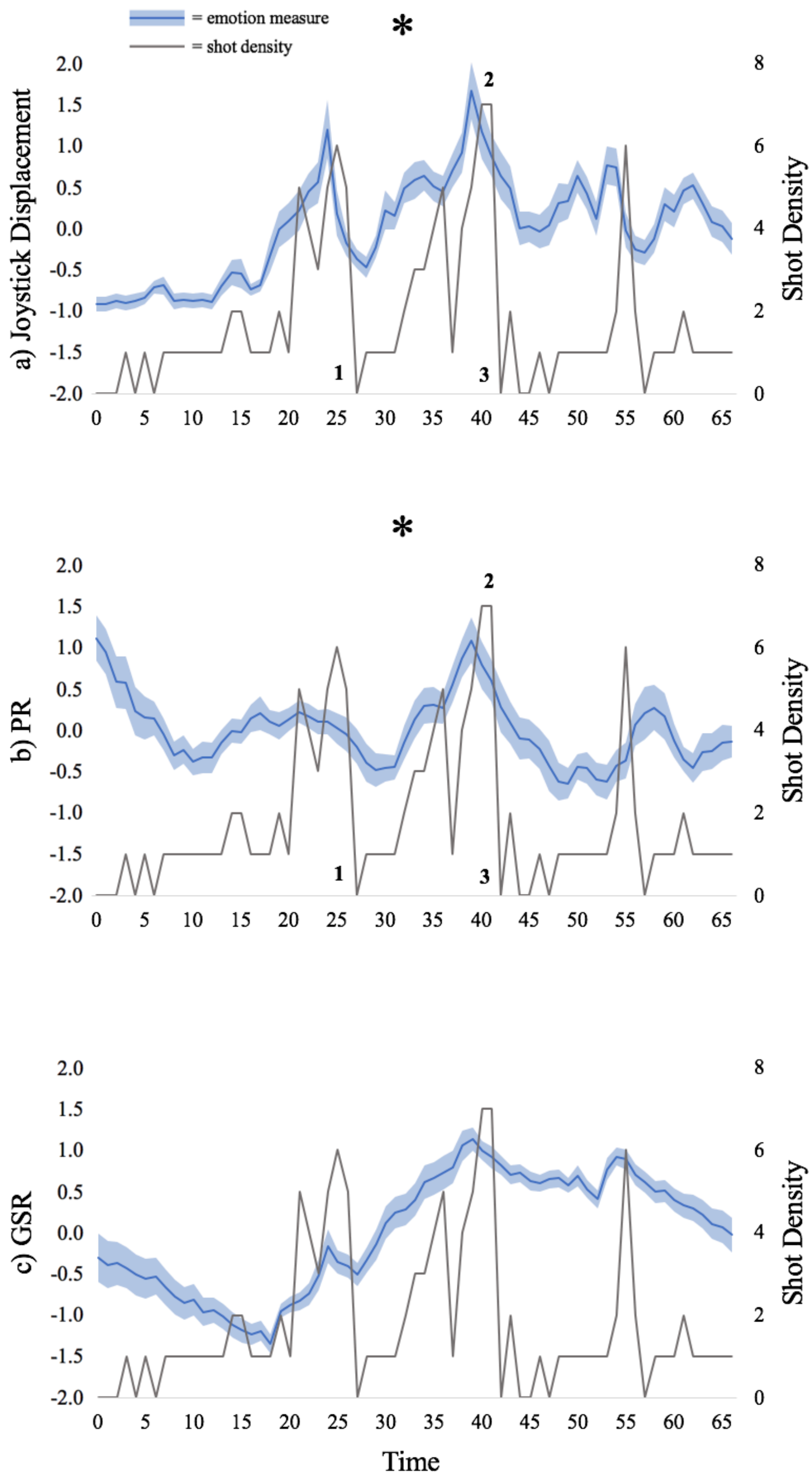

Fig. 22. Mean z-standardized emotion measures as a function of shot density for *Borrowed Time*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

Paperman. Magnitude of joystick displacement was not predicted by shot density ($t(1503) = 2.04$, $p = .042$) (Fig. 23a), and neither was GSR ($t(1511) = -2.57$, $p = .010$) (Fig. 23c). However, shot density did exert a significant influence on PR in a negative direction ($t(1504) = -5.74$, $p < .001$) (Fig. 23b). In the context of the narrative, this relationship is most prominently illustrated in three areas. First, prior to peak 1, shot density increases for 10 seconds, accompanied by a decline in PR. This portion of the film depicts George and Meg standing on an elevated train platform when a train rushes by, causing the wind to slap a piece of paper onto Meg's face, which George peels off, followed by Meg smiling at him. Second, prior to trough 2, shot density sharply decreases for 10 seconds, accompanied by a steep rise in PR. During this segment of the film, George runs out of his office to catch up to Meg. Finally, prior to peak 3, shot density increases for 5 seconds as PR decreases, during which George throws a paper airplane far into the air, sending it toward a dark alley.

In sum, subjective emotional tension was unaffected by shot density in *Paperman*. Of the objective measures of emotion, PR—but not GSR—was significantly predicted by shot density in a negative direction.

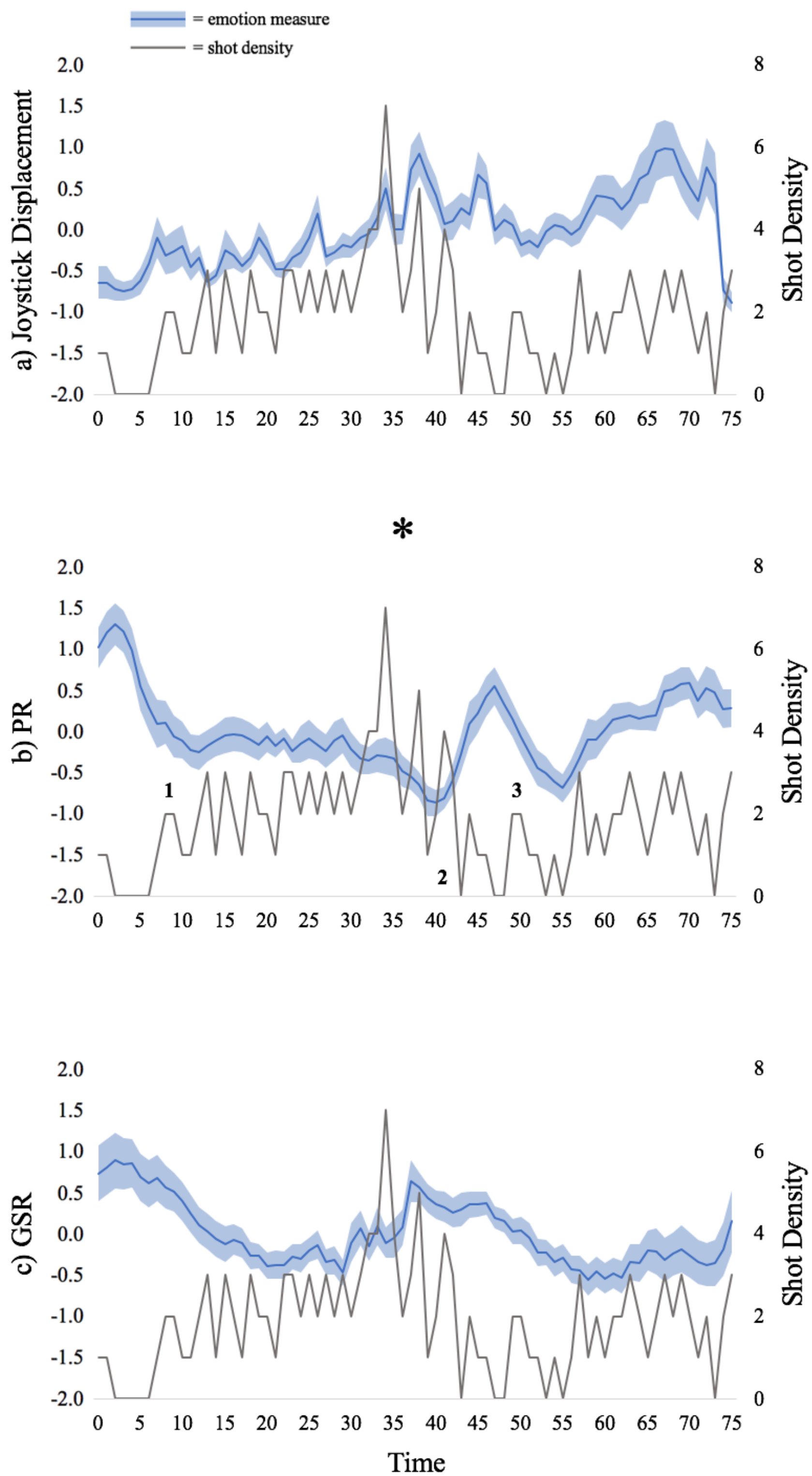

Fig. 23. Mean z-standardized emotion measures as a function of shot density for *Paperman*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

Stutterer. Magnitude of joystick displacement (Fig. 24a) was not predicted by shot density ($t(2905) = 1.89$, $p = .059$), and neither was PR (Fig. 24b) ($t(2931) = .20$, $p = .843$). However, shot density did exert a significant influence on GSR ($t(2931) = 3.42$, $p = .001$), such that as shot density increased, so did GSR, and vice versa (Fig. 24c). In the context of the narrative, this positive relationship is most prominently illustrated in two areas. First, prior to trough 1, both shot density and GSR decrease for 10 seconds. This portion of the film consists of a series of shots depicting Greenwood sitting in his apartment with a bandaged nose, then working distractedly in his studio, followed by a visit to his father's house, where Greenwood watches television while checking his phone to find that he has no messages from Ellie. Second, prior to peak 2, shot density increases for 10 seconds, accompanied by a rise in GSR. During this portion of the film, Greenwood's journey across town ends as he arrives on the street where Ellie stands; they see each other for the first time and tenderly hold each other's gaze.

In sum, subjective emotional tension was unaffected by shot density in *Stutterer*. Of the objective measures of emotion, GSR—but not PR—was significantly predicted by shot density in a positive direction.

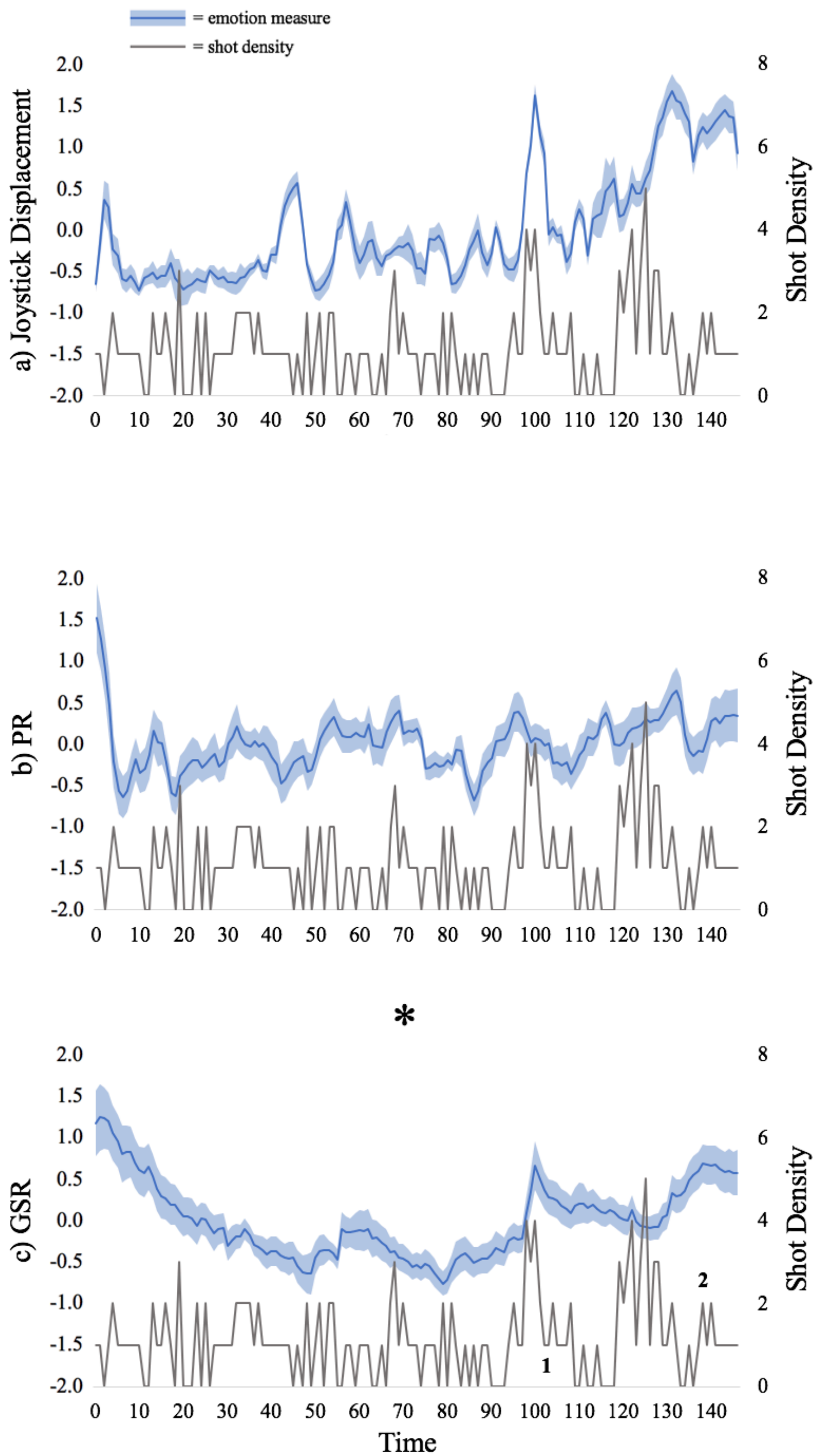

Fig. 24. Mean z-standardized emotion measures as a function of shot density for *Stutterer*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

The Voorman Problem. Shot density was not a significant predictor of joystick displacement ($t(2649) = 1.42$, $p = .156$) (Fig. 25a), PR ($t(2649) = -.63$, $p = .532$) (Fig. 25b), or GSR ($t(2651) = -1.20$, $p = .230$) (Fig. 25c).

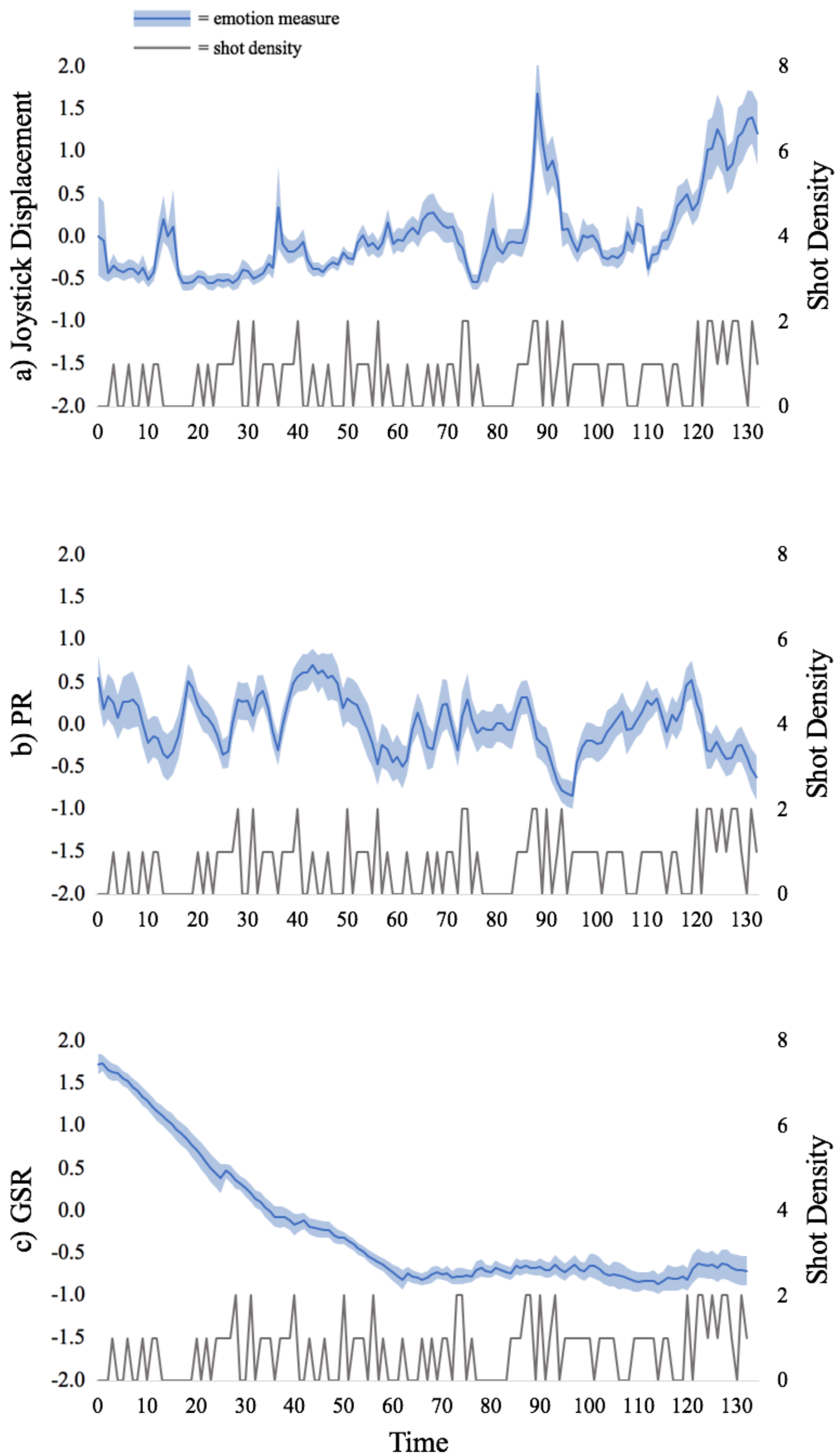

Fig. 25. Mean z-standardized emotion measures as a function of shot density for *The Voorman Problem*. Darker bands indicate 95% confidence intervals.

Interim discussion. As expected, in *Borrowed Time* and *Stutterer*, a positive relationship emerged between shot density and emotional tension. To understand why a greater number of shots over a given time frame would increase emotional response among viewers, a brief foray into film history is worthwhile.

The earliest feature films often relied on the *tableau* style of filmmaking, in which a distant camera would record actors playing out scenes in lengthy shots that were stitched together with little cutting—resulting in the appearance of a carefully composed picture. Early in the twentieth century, however, filmmakers veered away from this style as they embraced the cut, which led to a flourish of cutaways and close-up shots (Bordwell, 2012). One might expect that, in contrast to the ease of tableau style, a greater number of cuts over a given time frame would disrupt the smooth flow of visual information that contributes to film comprehension (thus reducing potential emotional response)—but this is not the case.

While cuts between shots objectively disrupt the depiction of visual space, Smith's (2012) Attentional Theory of Cinematic Continuity describes how filmmakers overcome this obstacle by satisfying the viewer's perceptual expectations. They do this by simulating the way humans attend to their environments in the real world. For instance, the 180-degree rule stipulates that the camera must record all shots in a given scene from one side of an imaginary axis that links two characters (mimicking the typical positioning of an observer to an interaction), thus allowing the viewer to visually connect adjacent shots. Likewise, a cutaway shot to a particular object immediately following a shot of a character's gaze serves to satisfy the viewer's curiosity surrounding the object of the character's scrutiny, and this unfolding of visual information mirrors gaze following in natural settings (Smith et al., 2012; also see Münsterberg, 1916/1970).

Common editing strategies also allow viewers to effortlessly infer a narrative from a series of shots. For instance, *montage* editing (in which dissimilar shots are strung together in a particular combination) prompts viewers to create a story out of the images presented. Two sequential shots of a man, then a child, looking off-screen while waving implies that they are waving to each other; likewise, a shot of an expressionless man followed by a cutaway of an attractive woman imbues the former with a perceived expression of lust (Bordwell, 2012; Kuleshov, 1974). In other words, the stitching together of two shots suggests a symbolic relationship between the images despite a lack of obvious temporal continuity. Fortunately for filmmakers, the human brain is on their side—it makes meaning out of the information it is presented; it cannot help it (see Shermer, 2008).

Over time, filmmakers have further honed their craft to render cuts less detectable to the average viewer (Bordwell 2002, 2006). They achieve this subtlety through a range of strategies, including match-action editing, in which filmmakers jump from one shot of an action to a new shot of the same action from a different viewpoint (e.g., Shimamura, Cohn-Sheehy, & Shimamura, 2014; Smith, 2012; Smith & Henderson, 2008), and intensified continuity, which utilizes a rapid pace of editing that removes extraneous narrative details (Bordwell 2002, 2006).

Collectively, the techniques described above represent an arsenal of tools that filmmakers deploy to achieve an impression of continuity not in spite of the cut, but because of it. Thus, greater shot density, when shrewdly executed, should not disrupt cinematic engagement; rather, it should harness viewer attention to foster absorption in the narrative. Indeed, evidence suggests that looking preference increases with shot pace, such that a quicker pace drives viewer motivation to gather visual information as images are presented on screen. Hochberg and Brooks

(1978) termed this phenomenon *visual momentum* and argued that it contributes to the captivation of film.

But does it follow that increased shot density should also drive emotional response? Not necessarily. Indeed, a positive relationship between shot density and emotional tension emerged for only two of the four short films utilized in this study; shot density exerted a negative influence on *Paperman*, and no influence on *The Voorman Problem.* An explanation for this discrepancy may lie in the conventions of genre, as filmmakers tend to adhere to genre-specific structural formulas. Action films, for example, generally feature shots that are shorter in duration, tighter in scale, and higher in motion (Cutting et al., 2010; Cutting, DeLong, & Brunick, 2011). Given that these patterns are amplified during the climax of a film (Cuttting, Brunick, & DeLong, 2011, 2012; Cutting, 2016b, 2016a), filmmakers appear to implicitly promote narrative absorption by applying a quicker editing pace to more suspenseful segments of the plot.

A striking example of this strategy is observed in *Borrowed Time*, where cross-cutting is utilized to rapidly switch between two scenes: that in which the sheriff approaches the edge of the cliff, and the flashback scene in which the carriage accident takes place, ultimately sending the boy's father over the same cliff approached by the sheriff in the present day. This technique contributes to narrative engagement not only by transporting the viewer from one line of action to the other (and thus, potentially creating anxiety; Pronin & Jacobs, 2008), but also by reflecting how the human mind oscillates between two different events, thus lending a tangible quality to the phenomenon of memory as it is projected onto the cinema screen (Bordwell, 2012; Münsterberg, 1916/1970). Filmmakers appear to utilize such attention-harnessing techniques when they are needed most—during moments of suspense. And this strategy is perhaps most salient across films that unfold at a quicker pace in general (indeed, of the four films utilized in this study, *Borrowed Time* boasts the second highest mean shot density, coming in a very close second to *Paperman*).

On the other hand, shot density also positively influenced emotional tension in *Stutterer*, a drama with a slower pace. However, one must not take an essentialist approach to genre. The category of a film is not always clear-cut, as the conventions of one genre may bleed into another, depending on the goals of the filmmaker. Indeed, *Stutterer* appears to borrow techniques from the action genre throughout its climax, as evidenced by the use of jump cuts to depict Greenwood's anxiety-ridden journey through town on his way to meet Ellie (described in greater detail above; see "Motion"). As a result, shot density is vastly increased throughout this particular sequence, arguably causing the viewer to become more absorbed in the narrative and, consequently, emotionally moved by its contents.

In contrast, shot density negatively influenced emotional tension in *Paperman* which, paradoxically, was marked by the highest mean shot density of all four films. However, as a romantic comedy, *Paperman* may not utilize shot density to the same ends as *Borrowed Time* and *Stutterer.* In fact, moments of greater shot density in *Paperman* generally do not coincide with emotionally evocative plot points—instead, shot density most sharply increases when George is folding and throwing paper airplanes. (Speculatively, this pattern may represent a stylistic choice made by the filmmakers to artfully communicate the paper airplane's thematic prominence in the narrative.)

It is important to note Cutting's (2016b, 2016a) observation that shot duration generally decreases during the beginning of the climax of Hollywood movies, regardless of genre. Thus, Paperman may be an outlier in terms of its pattern of shot pacing. Nonetheless, these findings support the notion that shot density drives emotional tension only when the narrative warrants it. It follows that the lack of relationship between shot density and emotional tension in *The Voorman Problem* may be explained by the remarkably slow pacing that persists throughout the film: shot density never exceeds two cuts per 5-second interval.

In sum, shot density likely drives emotional response based on the ways in which it is linked to the narrative trajectory of a film. Filmmakers seem to apply quicker pacing to more emotionally impactful moments of the narrative in order to harness the viewer's attention to the screen during these pivotal plot points. This is likely influenced, but not determined, by genre.

Shot scale

Fig. 26 displays the pattern of shot scale over the course of each film. *Stutterer* consists of the highest mean shot scale, equivalent to a medium close-up shot (M = 4.98, SD = 1.55), followed by *The Voorman Problem* (M = 4.50, SD = 1.54), then *Borrowed Time*, with a mean shot scale equivalent to a medium shot (M = 3.98, SD = 2.20). Comparatively, *Paperman* is marked by a longer mean shot scale, equivalent to a medium long shot (M = 3.05, SD = 1.48). The influence of shot scale on emotional engagement among viewers of each film is described in greater detail below.

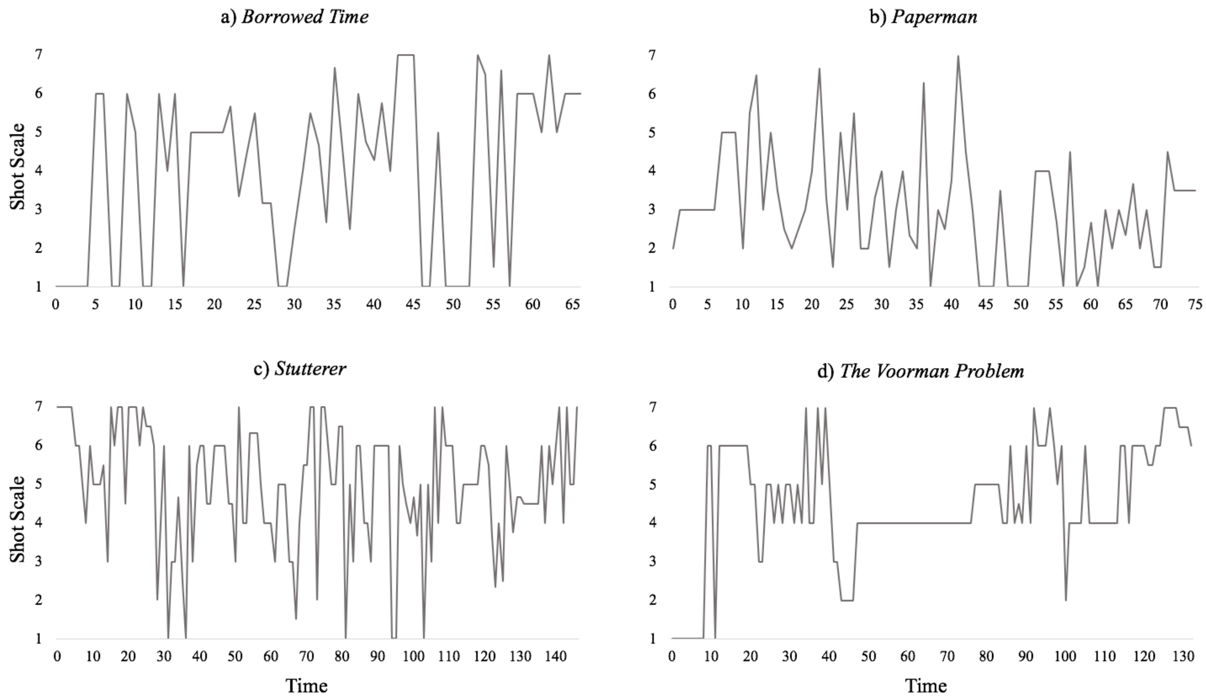

Fig. 26. Shot scale over time in a) *Borrowed Time*, b) *Paperman*, c) *Stutterer*, and d) *The Voorman Problem*. Shot scale units refer to the relative size of a character within the frame of a shot, coded on a scale from 1-7 (see Fig. 1), such that higher values indicate a closer mean shot scale across the preceding 5-second interval, while lower values indicate a more distant mean shot scale.

Borrowed Time. Neither PR ($t(1331) = 1.33$, $p = .183$) (Fig. 27b) nor GSR ($t(1331) = -.66$, $p = .511$) (Fig. 27c) were predicted by shot scale. However, shot scale did exert a significant influence on joystick displacement ($t(1318) = 3.37$, $p = .001$), such that as shot scale increased (toward closer-scaled shots), so did joystick displacement, and vice versa (Fig. 27a). In the context of the narrative, this positive relationship is most prominently illustrated in four areas. First, prior to trough 1, shot scale decreases for 15 seconds, accompanied by a steep decline in joystick displacement. During this portion of the film, the boy recovers from the carriage accident and then runs to the edge of the cliff from which his father has just fallen. Second, prior to peak 2, shot scale sharply increases for 5 seconds, accompanied by a steep rise in joystick displacement. During this segment, the boy accidentally shoots his father with a rifle, then the father's pocket watch flies through the air and lands on the ground nearby, splattered with blood. Third, prior to peak 3, shot scale again sharply increases alongside joystick displacement for 5 seconds, during which the sheriff struggles to pull himself up over the cliff from which he has just fallen. Immediately following this segment, shot scale sharply decreases for 10 seconds, accompanied by another steep decline in joystick displacement. This portion of the films consists of the sheriff recovering from his fall, then crawling toward his father's stopwatch and holding it in his hands.

In sum, objective emotional tension was unaffected by shot scale in *Borrowed Time*, while subjective emotional tension was significantly predicted by shot scale in a positive direction.

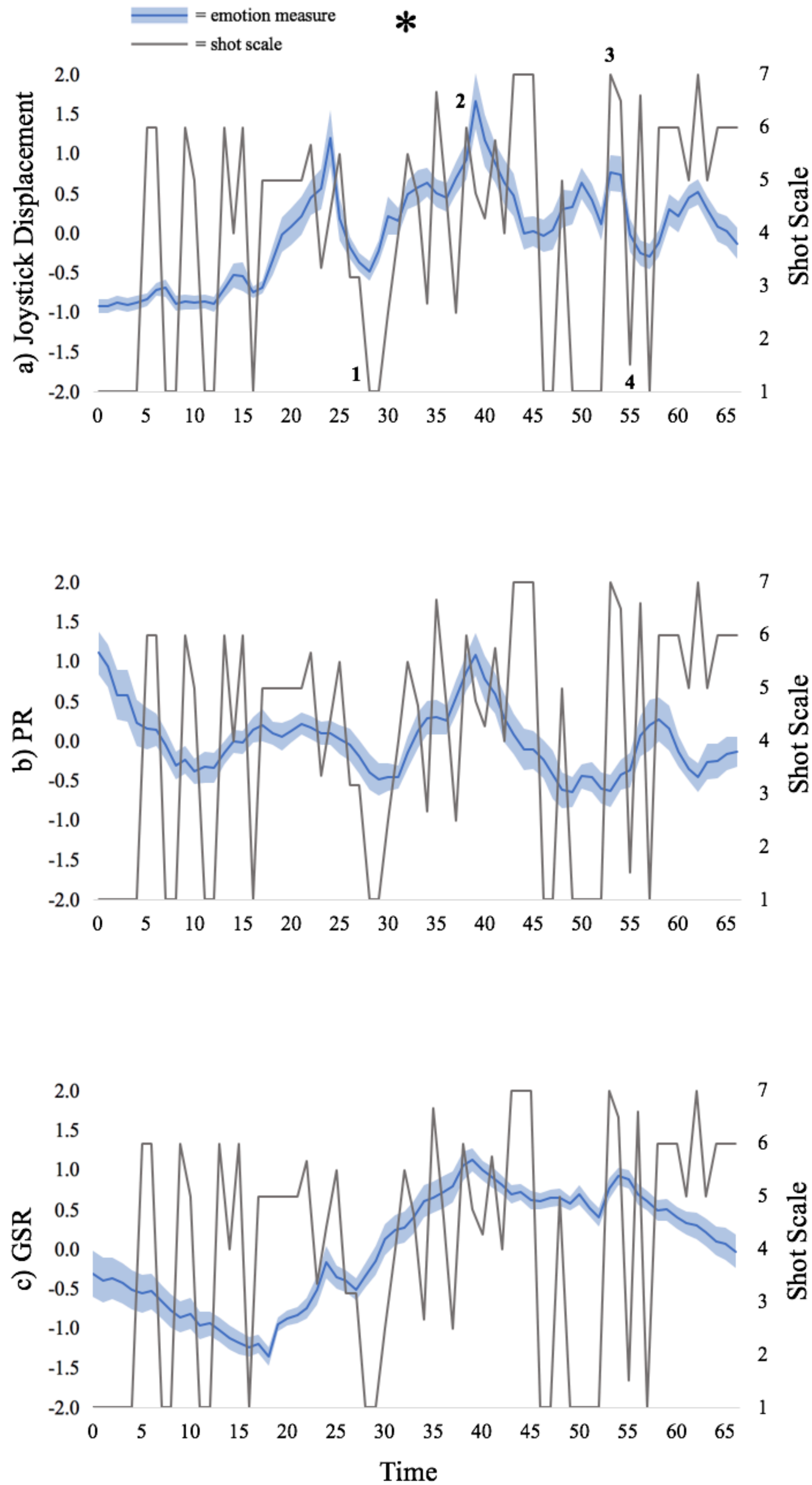

Fig. 27. Mean z-standardized emotion measures as a function of shot scale for *Borrowed Time*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

Paperman. Magnitude of joystick displacement was not predicted by shot scale ($t(1503) = .17$, $p = .864$) (Fig. 28a), and neither was GSR ($t(1511) = -1.65$, $p = .099$) (Fig. 28c). However, shot scale did exert a significant influence on PR in a negative direction ($t(1504) = -4.81$, $p < .001$) (Fig. 28b). In the context of the narrative, this relationship is most prominently illustrated in two areas. First, following peak 1, shot scale sharply declines for 15 seconds, accompanied by a steep rise in PR. This portion of the film depicts George running out of his office building to catch up to Meg. Second, prior to trough 2, shot scale again decreases for 5 seconds, accompanied by a modest rise in PR. During this segment of the film, George is pushed onto a train by a fleet of paper airplanes.

In sum, subjective emotional tension was unaffected by shot scale in *Paperman*. Of the objective measures of emotion, PR—but not GSR—was significantly predicted by shot scale in a negative direction.

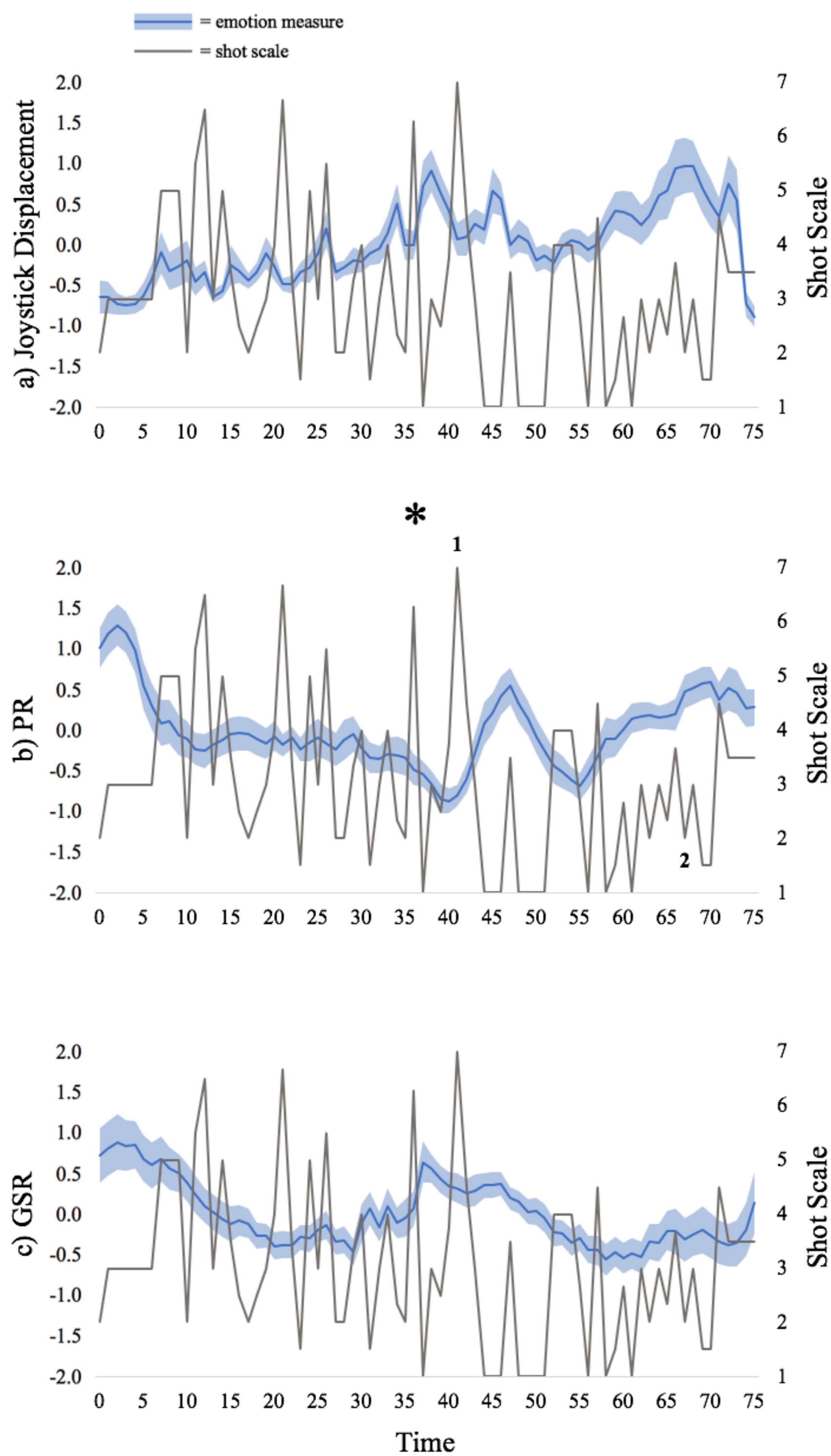

Fig. 28. Mean z-standardized emotion measures as a function of shot scale for *Paperman*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

Stutterer. Shot scale was not a significant predictor of joystick displacement ($t(2905) = 2.62$, $p = .009$) (Fig. 29a), PR ($t(2931) = -1.53$, $p = .125$) (Fig. 29b), or GSR ($t(2931) = -.13$, $p = .898$) (Fig. 29c).

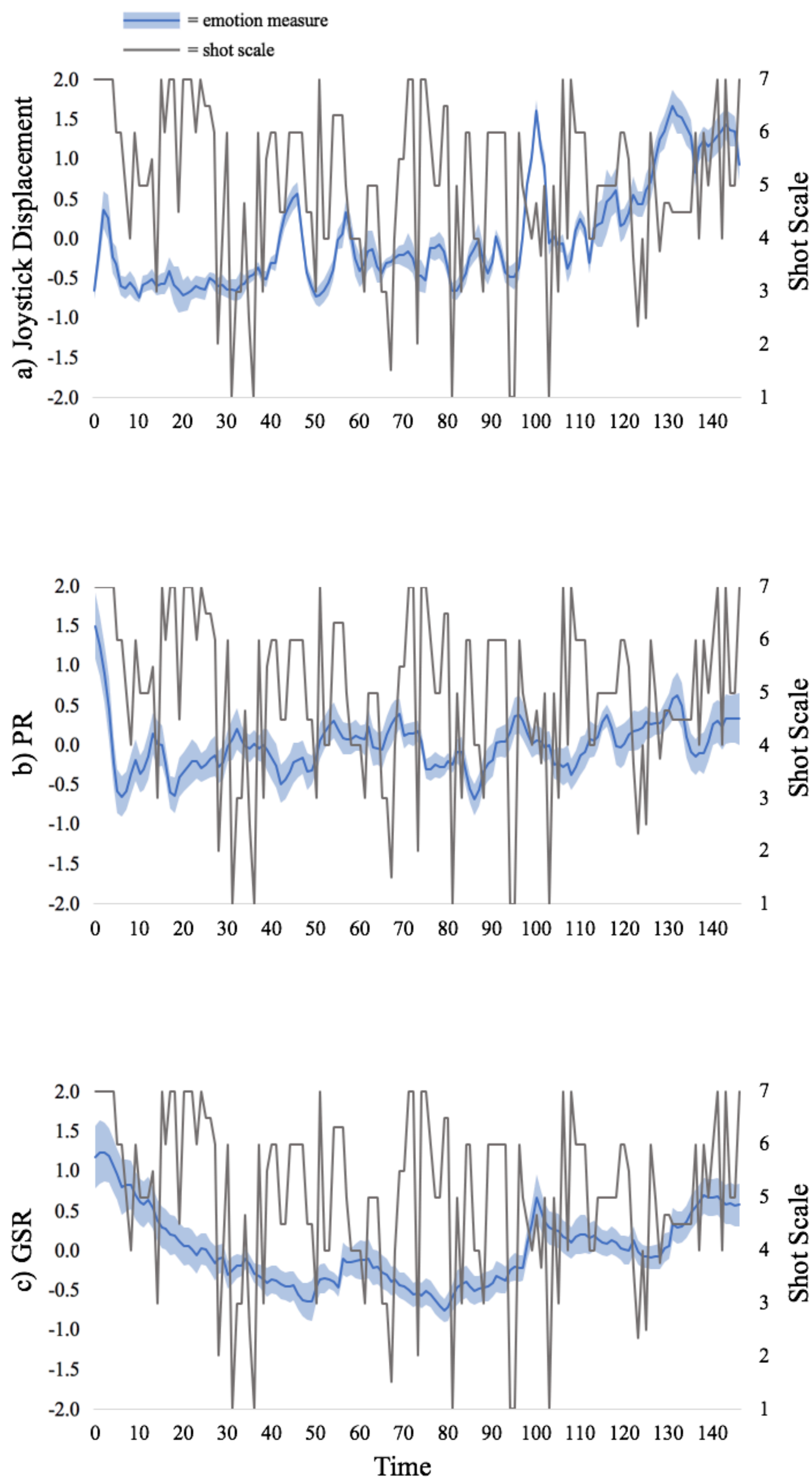

Fig. 29. Mean z-standardized emotion measures as a function of shot scale for *Stutterer*. Darker bands indicate 95% confidence intervals.

The Voorman Problem. PR was predicted by shot scale ($t(2649) = -5.05$, $p < .001$) (Fig. 30b), and so was GSR ($t(2651) = -3.38$, $p = .001$) (Fig. 30c). However, shot scale did not exert a significant influence on joystick displacement ($t(2649) = 1.22$, $p = .224$) (Fig. 30a). A negative relationship was observed between both shot scale and PR and shot scale and GSR, as illustrated most prominently in two areas for each measure.

In terms of PR (Fig. 30b), shot scale sharply increases for 5 seconds prior to peak 1, accompanied by a decrease in PR. During this segment of the film, Dr. Williams asks Governor Bentley how he will know when to release the prisoners (following the governor's admission that the computer systems had crashed). Second, prior to peak 2, shot scale again increases for 5 seconds as PR sharply declines. This portion of the film consists of Dr. Williams returning to the prison to meet with Voorman for the second time.

In terms of GSR (Fig. 30c), shot scale sharply increases for 10 seconds prior to peak 1, accompanied by an ongoing decrease in GSR. This portion of the film consists of Dr. Williams arriving at the prison and listening to faint chanting through the speaker on the front door. Second, prior to trough 2, shot scale decreases for 10 seconds, accompanied by a modest rise in GSR. During this segment of the film, Voorman switches places with Dr. Williams.

In sum, subjective emotional tension was unaffected by shot scale in *The Voorman Problem*, while objective emotional tension was significantly predicted by shot scale in a negative direction.

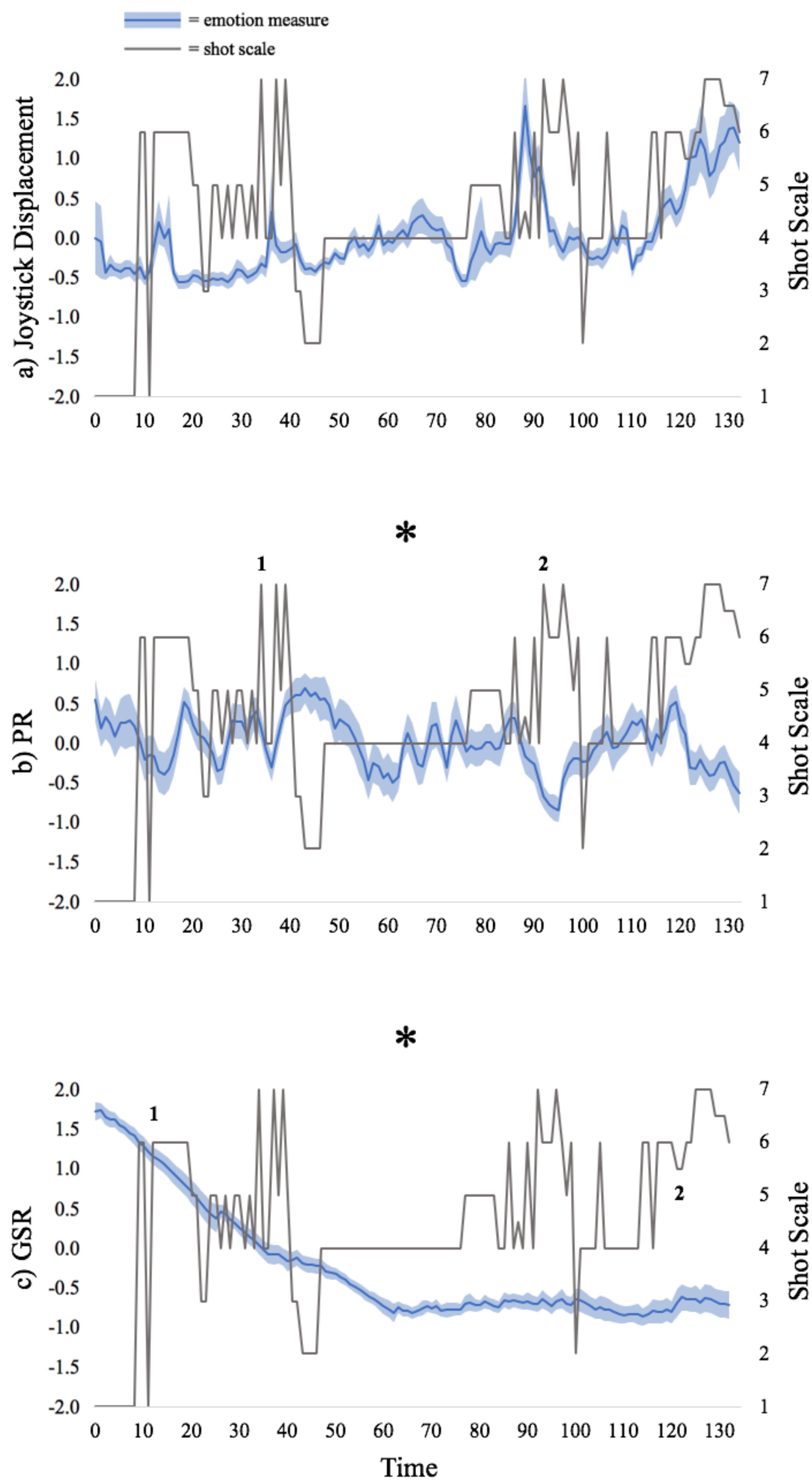

Fig. 30. Mean z-standardized emotion measures as a function of shot scale for *The Voorman Problem*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

Interim discussion. Narrative transportation theory stipulates that, in order for transportation to fully occur, viewers must not only successfully interpret a given story, but they must also experience *empathy* for its characters (Green & Sestir, 2017; van Laer et al., 2014). Scholars have appealed to the embodied nature of emotion to explore empathic engagement in the context of film. Specifically, Coplan (2006) and Plantinga (2009) highlight the role of emotional contagion, which describes the tendency for humans to automatically mimic the facial gestures of another person (via rapid facial reactions; see Dimberg, 1982; Dimberg & Thunberg, 1998; Dimberg, Thunberg, & Elmehed, 2000), resulting in a convergence of affective experience between two individuals (Hatfield, Cacioppo, & Rapson, 1993).

Rooney and Bálint (2018) provide empirical support for this notion. They found that closer-scaled shots of an animated character prompted Theory of Mind among viewers when the facial expression was sad (but not when it was neutral). Theory of Mind describes the ability to understand the mental states of others, including their beliefs, desires, and perspectives, even as these states differ from one's own (for a review, see Carruthers & Smith, 1996). Thus, Rooney and Bálint (2018) concluded that shot scale, as a low-level cinematic feature, can elicit Theory of Mind responses by directing viewer attention to the features of a character's face that communicate their emotional state. These results support previous findings that social reality is not required for emotional contagion to occur, as facial mimicry is reliably elicited by virtual characters in works of animation (Mojzisch et al., 2006).

Given the evidence presented above, one might expect that shot scale would positively influence emotional tension across the four short films utilized in this study. However, the data do not support this hypothesis, as the expected pattern emerged for *Borrowed Time* only. In fact, shot scale negatively predicted emotional tension for both *Paperman* and *The Voorman Problem*, and it had no effect on emotion tension for *Stutterer.*

On the surface, the lack of effect for *Stutterer* is particularly surprising, given that the film boasts the highest mean shot scale of all four films (equivalent to a medium close-up shot). However, a closer look at the time course of shot scale reveals that *Stutterer* is marked by several periods of extreme close-up shots: the data show 13 peaks that indicate 5-second time intervals comprised entirely of extreme close-up shots, which far exceeds any other film (*The Voorman Problem* shows the next highest number of extreme close-up peaks at just 6). And given that medium close-up shots minimize the effort required of the viewer to read emotional information from a character's face (as the viewer is able to simultaneously gather information from the eyes and the mouth of a character in a single glance; Smith, 2012, 2013), extreme close-up shots may be detrimental to emotional engagement by obscuring areas of the character's face or by forcing the viewer to saccade across the screen in order to process the constellation of facial features that constitute an emotional expression. Indeed, Fig. 31 reveals that the subjective emotional tension reported by viewers of *Stutterer* was greatest for medium close-up shots, and then dipped to below average values for both close-up and extreme close-up shots.

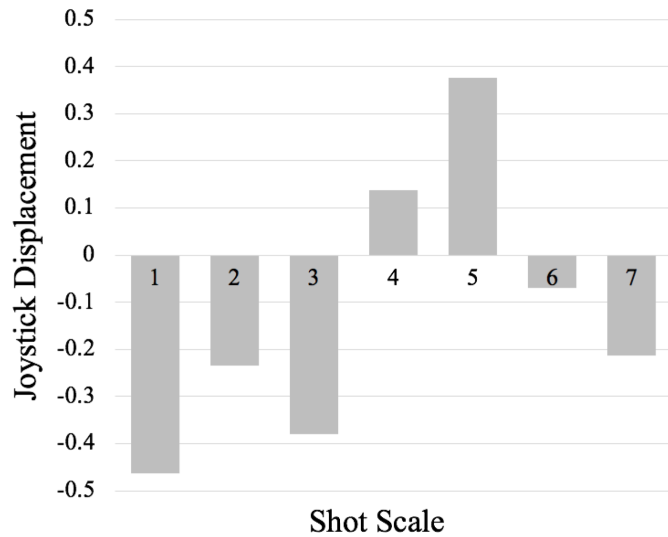

Fig. 31. Mean z-standardized joystick displacement as a function of binned shot scale for *Stutterer*. Bins were determined by classic rounding, such that all 5-second intervals marked by a mean average shot scale between 2.50 and 3.49 were binned in shot scale category 3, and so on.

Critically, the facial expressions displayed by characters in a film *are* relevant to emotional engagement. As previously discussed, Cutting & Armstrong (2018) found that reaction shots (which increasingly end conversations in Hollywood film) tend to be perceived by viewers as slightly negative in valence with modest arousal—that is, slightly perturbed and slightly aroused, as though stifling the urge to speak. By withholding a character's overt emotional response in favor of a clear struggle *not* to speak, filmmakers invite viewers to imagine what might have been said, thus driving viewer engagement with characters and building dramatic impact. Pertinent to this finding is the fact that, over the last century, mean shot scale has shifted from a medium long shot toward a medium close-up, thus facilitating the viewer's ability to read these facial expressions at all (Cutting & Armstrong, 2018). As such, an overabundance of extreme close-up shots might disrupt emotional engagement among viewers by presenting a challenge to facial expression processing.

Notably, the one positive relationship that emerged between shot scale and emotional response was in the case of subjective emotional tension reported by viewers of *Borrowed Time* (which, incidentally, was marked by the second lowest mean shot scale of all four films). The remaining associations between the two variables occurred in a negative direction across objective measures of emotional tension. This discrepancy might be explained by the two categories of emotion variables capturing different facets of emotional experience—namely, subjective measures of emotion may be comparatively more sensitive to the visual structure of film, while objective measures may be more sensitive to the narrative elements of film (as described in greater detail below; see "Correlations Among Emotion Measures").

Even if this is the case, the fact remains that shot scale did not predict subjective emotional tension for three of the four short films utilized in this study. These unexpected findings warrant further research to better isolate the effects of shot scale (ideally without the presence of competing low-level cinematic variables).

Sound amplitude

Fig. 31 displays the pattern of sound amplitude over the course of each film. *Stutterer* consists of the highest degree of sound (M = 7.06, SD = 7.91), followed by *The Voorman Problem* (M = 6.50, SD = 3.99). Comparatively, *Paperman* (M = 4.65, SD = 5.39) and *Borrowed Time* (M = 3.56, SD = 4.51) are somewhat quieter, although the latter shows a prominent peak at the point in the narrative in which a gunshot is fired. The influence of sound amplitude on emotional engagement among viewers of each film is described in greater detail below.

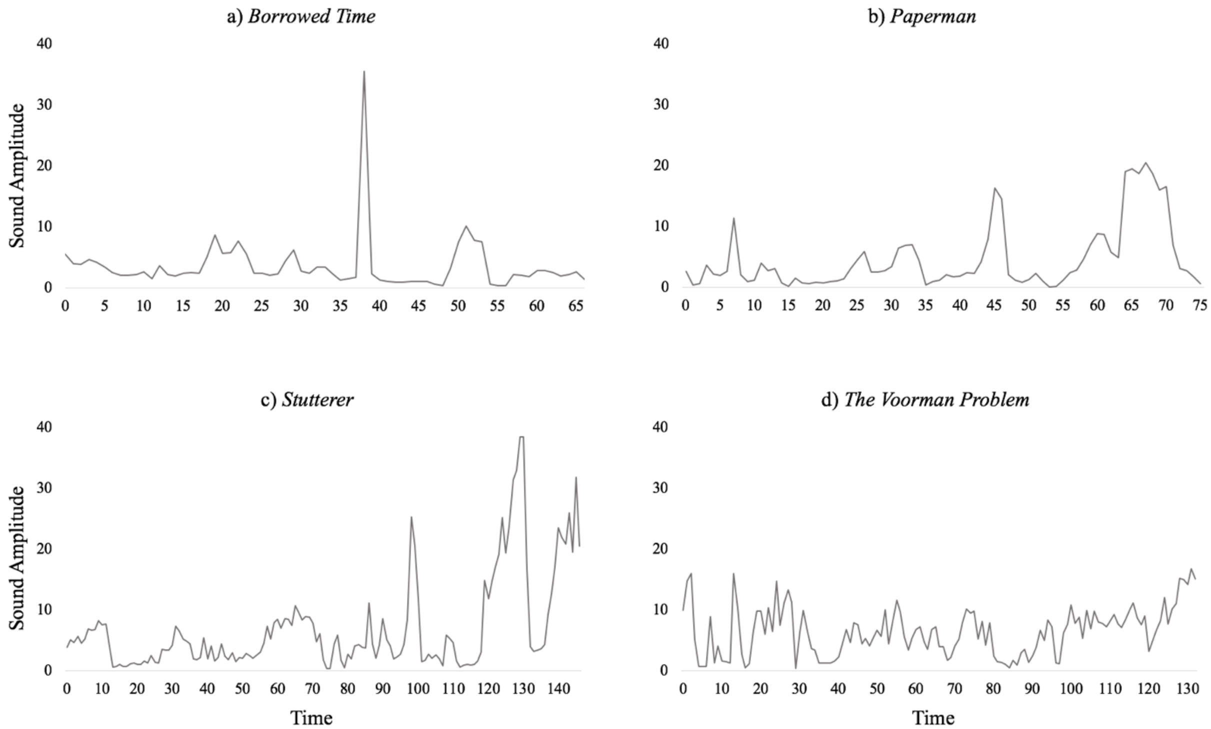

Fig. 32. Sound amplitude over time in a) *Borrowed Time*, b) *Paperman*, c) *Stutterer*, and d) *The Voorman Problem*. Sound amplitude units refer to mean volume on a scale from 0-40, such that higher values indicate greater volume across the preceding 5-second interval, while lower values indicate lesser volume.

Borrowed Time. Sound amplitude was not a significant predictor of joystick displacement ($t(1318) = 2.10$, $p = .036$) (Fig. 33a), PR ($t(1331) = 2.61$, $p = .009$) (Fig. 33b), or GSR ($t(1331) = 2.55$, $p = .011$) (Fig. 33c).

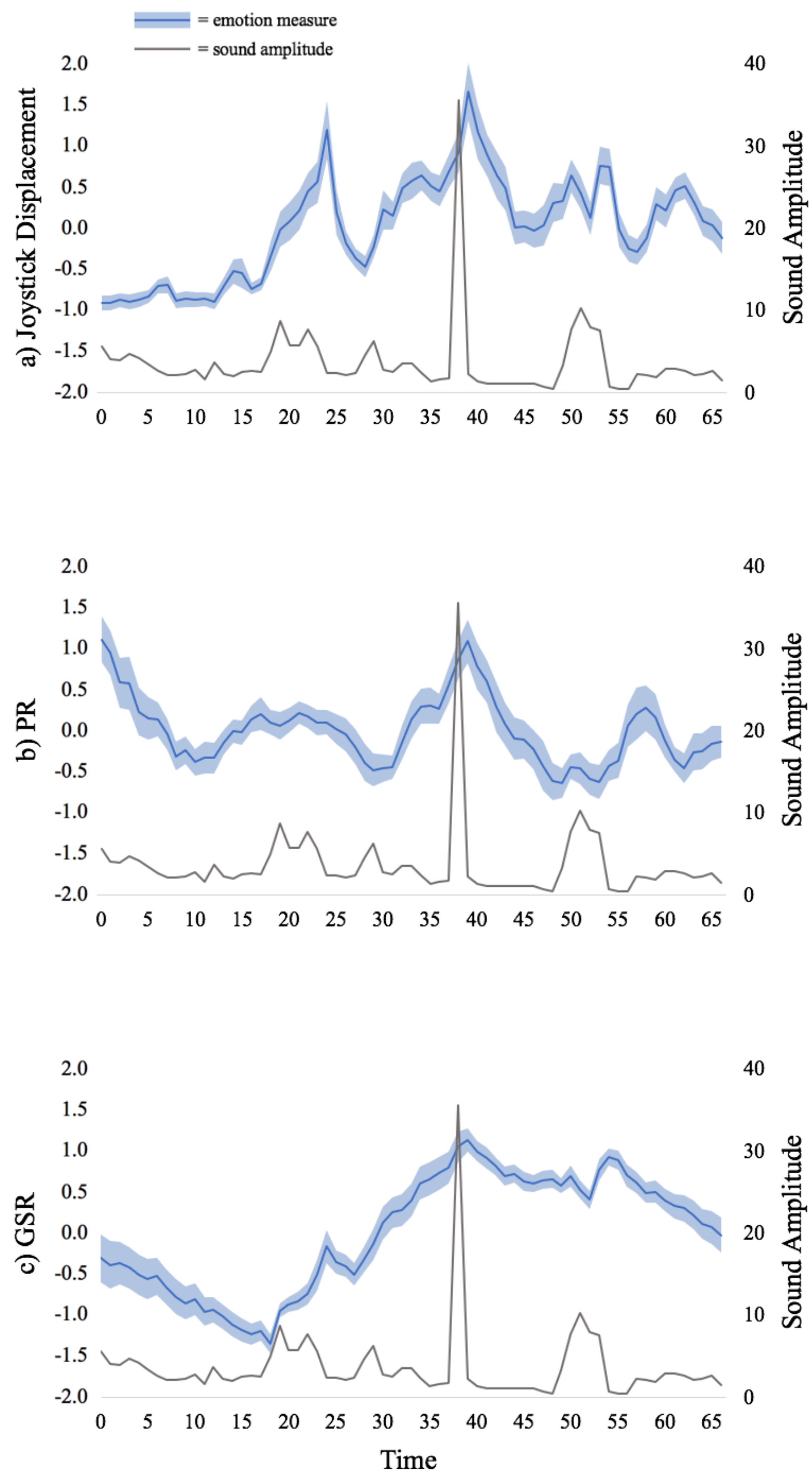

Fig. 33. Mean z-standardized emotion measures as a function of sound amplitude for *Borrowed Time*. Darker bands indicate 95% confidence intervals.

Paperman. Magnitude of joystick displacement was predicted by sound amplitude ($t(1503) = 4.92$, $p = .001$) (Fig. 34a), and so was PR (Fig. 34b) ($t(1504) = 3.07$, $p = .002$). However, sound amplitude did not exert a significant influence on GSR ($t(1511) = 2.09$, $p = .037$) (Fig. 34c). A positive relationship was observed between sound amplitude and joystick displacement and sound amplitude and PR, as illustrated most prominently in four areas for the former and three areas for the latter.

In terms of subjective emotional tension (Fig. 34a), sound amplitude sharply increases for 5 seconds prior to peak 1, accompanied by an increase in joystick displacement. During this segment of the film, George has just peeled a piece of paper off of Meg's face as they stand together on a train platform, and she smiles at him. Second, prior to trough 2, sound amplitude sharply decreases for 10 seconds alongside joystick displacement. This portion of the film consists of George throwing a paper airplane far into the air, sending it toward a dark alley. Third, prior to peak 3, sound amplitude increases for 20 seconds, accompanied by an increase in joystick displacement. During this segment, George is pushed onto a train by a fleet of paper airplanes while Meg chases the paper airplane with her lipstick onto a different train; then, as George tries to leave his seat on the train, the airplanes force him back down. Finally, prior to trough 4, sound amplitude sharply decreases for 20 seconds, alongside a steep decline in joystick displacement, during which George and Meg finally meet on the train platform, followed by a series of still shots depicting the two talking in a café.

In terms of PR (Fig. 34b), sound amplitude sharply decreases for 10 seconds prior to trough 1, accompanied by a decrease in PR. This portion of the film consists of George and Meg enjoying a moment of laughter together before Meg disappears onto a nearby train. Second, prior to peak 2, sound amplitude sharply increases for 15 seconds, accompanied by steep rise in PR.

During this segment of the film, George runs out of his office to catch up to Meg, then dangerously crosses through traffic. Finally, prior to peak 3, sound amplitude increases for 20 seconds, accompanied by an increase in PR, during which both George and Meg are pushed and lured, respectively, onto their trains by paper airplanes.

In sum, subjective emotional tension was significantly predicted by sound amplitude in *Paperman*. Of the objective measures of emotion, PR—but not GSR—was also predicted by sound amplitude. Both effects revealed a positive relationship.

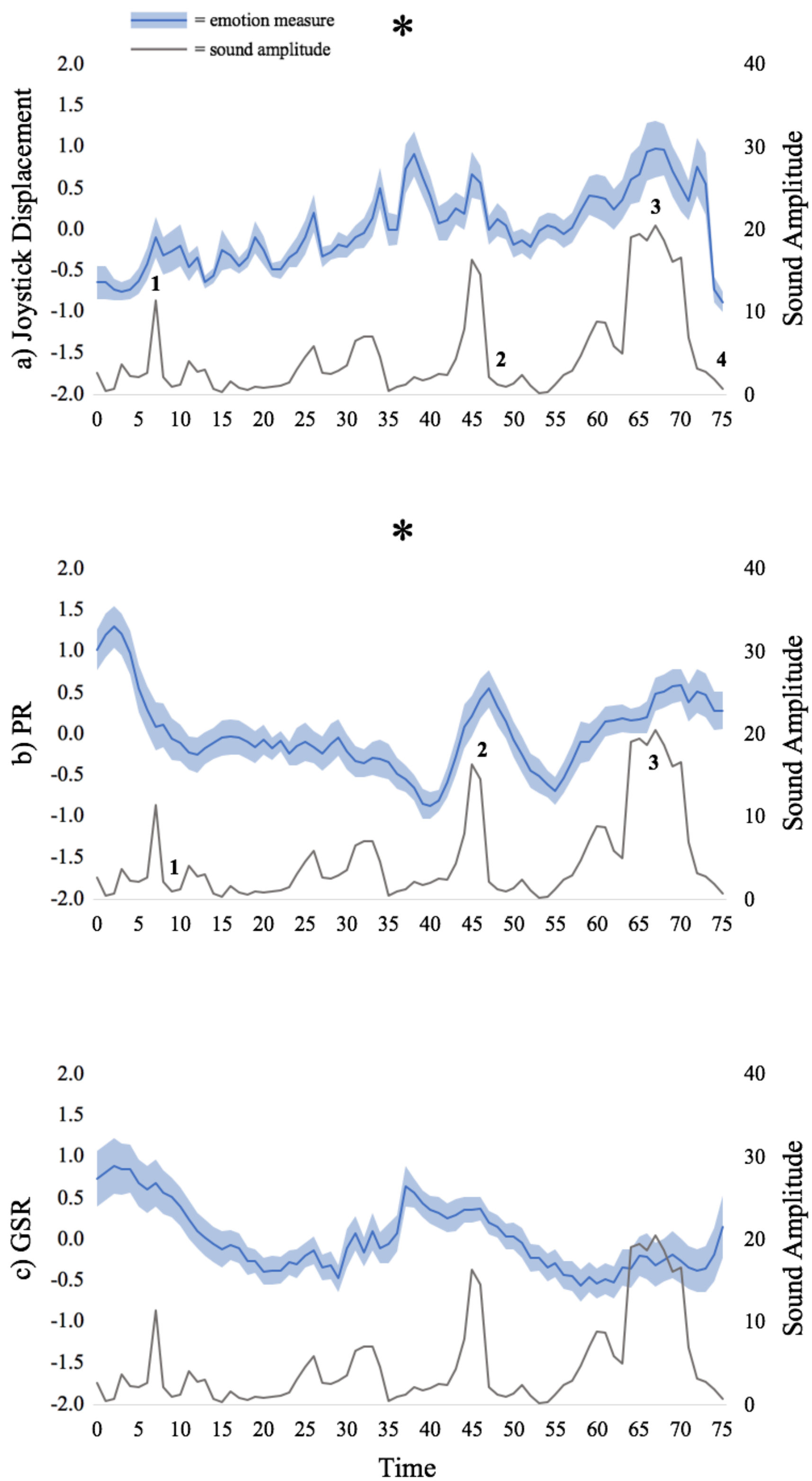

Fig. 34. Mean z-standardized emotion measures as a function of sound amplitude for *Paperman*. Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

Stutterer. Sound amplitude was not a significant predictor of joystick displacement ($t(2905) = 2.37$, $p = .018$) (Fig. 35a), PR ($t(2931) = 1.13$, $p = .260$) (Fig. 35b), or GSR ($t(2931) = -1.96$, $p = .050$) (Fig. 35c).

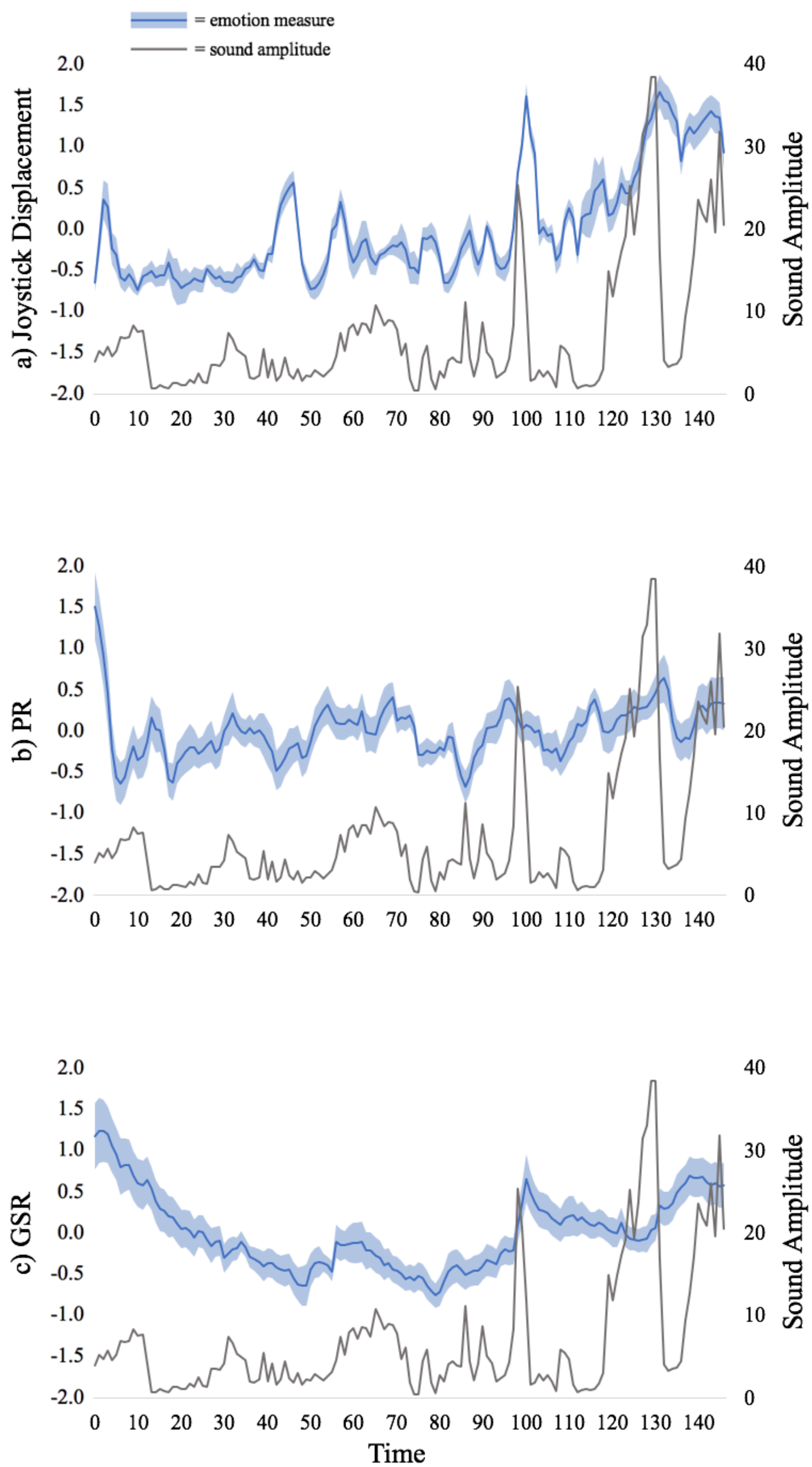

Fig. 35. Mean z-standardized emotion measures as a function of sound amplitude for *Stutterer*. Darker bands indicate 95% confidence intervals.

The Voorman Problem. Magnitude of joystick displacement was not predicted by sound amplitude ($t(2649) = -1.90$, $p = .058$) (Fig. 36a), and neither was GSR ($t(2651) = .50$, $p = .617$) (Fig. 36c). However, sound amplitude did exert a significant influence on PR in a negative direction ($t(2649) = -2.68$, $p = .007$) (Fig. 36b). In the context of the narrative, this relationship is most prominently illustrated in four areas. First, prior to trough 1, sound amplitude decreases for 10 seconds, accompanied by a rise in PR. During this segment of the film, Governor Bentley asks Dr. Williams to evaluate Voorman. Second, prior to peak 2, sound amplitude jaggedly increases for 25 seconds, accompanied by a sharp decline in PR. This portion of the film depicts Dr. Williams returning to the prison to meet with Voorman for the second time (while chanting can be heard in the background), followed by Voorman teasing Dr. Williams about the disappearance of Belgium. Immediately following this segment, sound amplitude sharply decreases for 15 seconds prior to trough 3, accompanied by a steep rise in PR, during which Voorman asks if Dr. Williams believes he is a demon who should be exorcised; Dr. Williams responds by asking Voorman if he thinks he should be, and Voorman shrugs. Finally, prior to peak 4, sound amplitude jaggedly increases for 55 seconds, matched by a similarly jagged decline in PR, during which Dr. Williams struggles against a straitjacket after Voorman has switched places with him. Voorman leaves the room as Dr. Williams grows increasingly agitated, shouting louder and louder against a backdrop of chanting.

In sum, subjective emotional tension was unaffected by shot density in *Paperman*. Of the objective measures of emotion, PR—but not GSR—was significantly predicted by shot density in a negative direction.

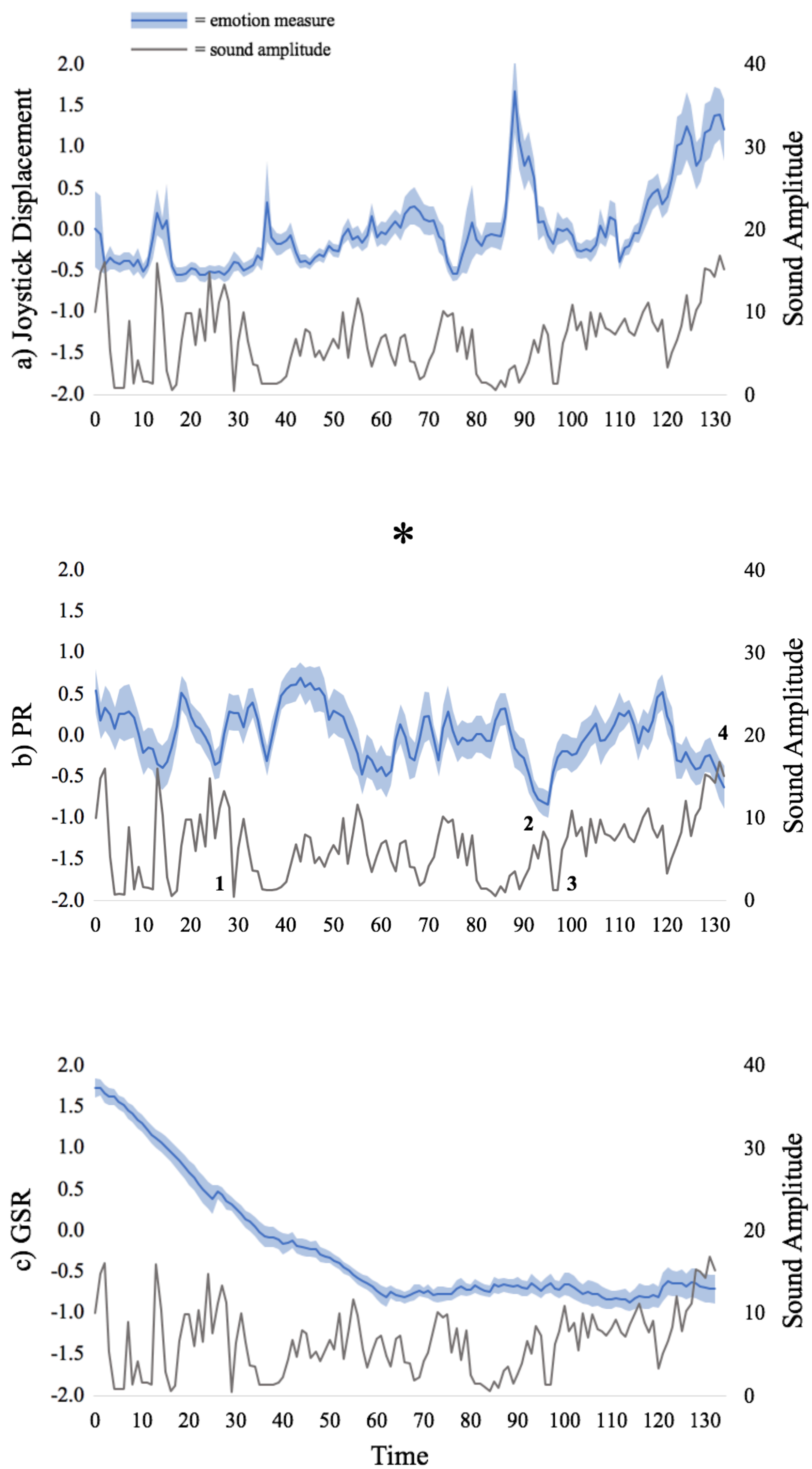

Fig. 36. Mean z-standardized emotion measures as a function of sound amplitude for *The Voorman Problem.* Darker bands indicate 95% confidence intervals. Asterisks indicate significant effects ($p \leq .008$).

Interim discussion. Once again, the effects of sound amplitude were not uniform across films. I had tentatively predicted that sound amplitude would positively influence emotional response measures, such that increased volume would drive heightened emotional tension, and vice versa. This pattern emerged for *Paperman* only. Sound amplitude had no effect on emotional tension among viewers of *Borrowed Time* and *Stutterer*, and it negatively influenced PR among viewers of *The Voorman Problem.*

Speculatively, the lack of effect for *Borrowed Time* and *Stutterer* might be due to different ways in which diegetic versus non-diegetic sources of sound influence emotional response—for instance, perhaps volume caused by dialogue among characters affects emotional tension differently than volume caused by the score of a film. Unfortunately, the sources of sound across *Borrowed Time* and *Stutterer* cannot easily be parsed in the confines of this study. However, given that *Paperman* contains no dialogue, the effects of sound can be narrowed down to sound effects and music, and the latter appears to be the primary contributor to sound amplitude in *Paperman*, as music persists throughout most of the final three quarters of the film (through the complicating action, development, and climax; see Appendix A). Thus, the relationship that emerged between sound amplitude and emotional response among viewers of *Paperman* suggests that the volume of a film's score may positively influence both subjective and objective measures of emotional tension.

On the other hand, sound amplitude negatively influenced PR among viewers of *The Voorman Problem.* While this film boasts the second highest mean sound amplitude, a closer inspection of the plot reveals that the sound in *The Voorman Problem* originates almost entirely with diegetic sound (accompanied by very little music, which plays solely at the beginning and end of the film; see Appendix A). Thus, when considering the influence of a film's score on emotional response among viewers, *The Voorman Problem* is marked by virtual silence.

A number of film scholars have theorized about the role of silence in film, particularly as it pertains to score. For instance, Walker (2010) describes how Austrian director Michael Haneke's sparing use of music imbues his films with an "unnerving power," as the resulting focus on silence punctuated by diegetic sound creates an unsettling tension. This technique is readily apparent in the psychological thriller *Caché* (2005) and the horror film *Funny Games* (2007), both of which exemplify Haneke's artistic focus on social issues, isolation, and estrangement. Likewise, Jim Jarmusch, American director of independent films, often utilizes short musical phrases interspersed with prolonged periods of silence in order to build dramatic impact (Kulezic-Wilson, 2009), as exemplified in *Paterson* (2016)—a drama marked by quiet pauses that seemingly reflect the slow unravelling of poetry created in the protagonist's mind.

Strategies such as these may raise emotional tension by spoiling the entertainment value of cinema with a sense of reality, as music typically cues the viewer how to feel during moments that would evoke uncomfortable silence if they were to occur in the real world (Kuperberg & Kuperberg, 2005). Similarly, Balazs (1923/2012) argues that the vivid effects of silence occur precisely because soundless spaces are nonexistent in the wild, and therefore, unfamiliar to the viewer's senses. Balazs (1923/2012) also indirectly credits silence for the emotional impact of reaction shots, claiming that "the physiognomy of men is more intense when they are silent," thus communicating a tension that rouses the events to follow. Such theories are reflected in practical resources for contemporary filmmakers: Aldredge (2017) recommends using silence to provide a narrative punch, to suggest isolation, or to imbue characters with emotional depth (e.g., by conveying the emptiness of grief), and Harriott (2017) describes how cinematic silence drives deeper narrative engagement by prompting the viewer to fill in the voids where sound might otherwise live.

The Voorman Problem provides tentative psychological evidence for these guidelines. As a film that makes sparing use of music, the negative association between sound amplitude and PR suggests that periods of relative silence drive emotional tension. Of course, further research should corroborate these findings with a larger corpus of films.

Additionally, it is worth noting that, while no relationship emerged between sound amplitude and GSR in *The Voorman Problem*, the film begins with a relatively loud and high-pitched segment of music, which may explain the elevated GSR among viewers at the start of the film. This burst of GSR activity suggests that sound amplitude may have interfered with the potential emotional influence exerted by the other low-level features of *The Voorman Problem*.

In sum, these findings suggest that non-diegetic sound (e.g., music) drives an increase in both subjective and objective measures of emotional tension, while diegetic sound (e.g., sound effects, dialogue) drives a negative relationship between sound amplitude and emotional tension, and that this effect likely manifests in the silences that linger in between sound.

Before moving on, it is worth noting that the soundtracks to all four films lacked music with vocalized lyrics (i.e., songs). Recent evidence suggests that songs in film might have an additive effect on emotional tension. Albouy et al. (2020) found that humans decode music in much the same way that songbirds do, such that the left hemisphere of the brain processes the words of a song while the right hemisphere processes the melody. These findings indicate that songs engage the brain in a way that is fundamentally different than speech or music alone, which may explain why songs are especially meaningful. Thus, future research should explore the differential effects of speech, instrumental music, and songs on emotional tension among film viewers.

Omnibus Analyses

To assess the overall influence of each low-level cinematic variable, I performed omnibus analyses by removing timepoint as a fixed effect from each regression model, collapsing the data across films (and adding film as a fixed effect), and then performing a stepwise regression to determine the percent variance explained by each cinematic variable (Table 1). While these analyses allow a broader look at the influence of film structure on emotional synchrony, the results come with a caveat: removing timepoint as a fixed effect ignores the time-varying nature of the variables, rendering the analyses a rough approximation of the overall influence of cinematic structure on both subjective and objective measures of emotional response. As such, these findings should be taken with a grain of salt.

However, I chose to proceed with these analyses for the following reason: in the context of cinema, time as a metric is arguably not as informative as it is in more objective settings. Simply stated, time is imbued with subjective influence while viewing film. To illustrate, Pollatos, Laubrock, and Wittmann (2014) reported that fear-inducing film clips resulted in greater subjective time dilation among viewers, while amusing film clips resulted in greater subjective time contraction. This distortion is potentially linked to physiological signals. To explain their findings, Pollatos et al. (2014) appealed to Craig's (2009a) proposal of an internal clock that keeps time based on the accumulation of bodily signals processed in the insular cortex (a region of the brain involved in perception and consciousness; see Craig, 2009b). Specifically, physiological arousal results in an increased accumulation of perceived temporal units during a given time frame. Thus, emotions known to heighten physiological arousal increase the pacemaker rate, resulting in an overestimation of time duration (also see Gil & Droit-Volet, 2012; Wittmann, 2009; Wittmann & Wassenhove, 2009). Additionally, Lositsky et al. (2016) found that estimates of time duration between two clips in a radio story were correlated with the distance between neural patterns at encoding of each of the clips. Specifically, fMRI pattern changes in the right insula predicted subsequent duration estimates.

In sum, narrative processing appears to result in subjective temporal distortion. This phenomenon may be amplified by filmmakers' increasing usage of intensified continuity, which describes a technique of rapid editing marked by the removal of extraneous narrative details (Bordwell 2002, 2006). Thus, given the constructed nature of time in film, it is perhaps preferred but not entirely necessary to map timepoint as an objective, fixed effect onto emotional response variables.

	t	*p*	% variance explained
Joystick			
Clutter	-1.26	.207	11.51
Luminance	-12.21	<.001*	9.99
Motion	9.63	<.001*	10.67
Shot density	12.28	<.001*	8.69
Shot scale	8.64	<.001*	11.49
Sound amplitude	17.10	<.001*	7.10
PR			
Clutter	0.66	.512	0.83
Luminance	-1.94	.052	0.80
Motion	1.19	.234	0.82
Shot density	-0.95	.340	0.83
Shot scale	-4.40	<.001*	0.77
Sound amplitude	5.31	<.001*	0.49
GSR			
Clutter	5.89	<.001*	0.36
Luminance	-2.68	.007*	0.56
Motion	-4.63	<.001*	0.83
Shot density	0.03	.973	0.93
Shot scale	-3.81	<.001*	0.69
Sound amplitude	3.06	.002*	0.92

Table 1. Results of omnibus regression analyses (collapsed across film). Asterisks denote significant predictors, as determined by a Bonferroni corrected α of 0.008.

The results of the omnibus analysis reveal that the visual structure of film exerts a heavier influence on subjective emotional tension than objective measures of emotion. Specifically, five out of six cinematic variables predicted joystick displacement, and the percent variance explained by each significant measure ranged from 7.10% (sound amplitude) to 11.49% (shot scale). As expected, subjective emotional tension increased alongside diminished luminance (i.e., a darker screen), greater motion, greater shot density (i.e., a faster shot pace), closer-scaled images, and greater sound amplitude—suggesting that these patterns in film structure at least partially explain the remarkable synchrony in emotional response observed across viewers.

In comparison, while longer-scaled shots and greater sound amplitude predicted an increase in PR, and while greater clutter, diminished luminance, decreased motion, longer-scaled shots, and greater sound amplitude predicted an increase in GSR, the effects were markedly smaller, with percent variance explained by each variable lingering below 1%. Although small values might be expected (given the numerous cinematic factors—both structural and narrative—that not only drive emotional response but potentially compete with each other), the difference in effect size across emotion variables suggests a dissociation between subjective and objective measures of emotional tension. This dissociation is amplified by the opposing directions of effect across variables. For instance, motion positively predicted joystick displacement but negatively predicted GSR. A similar pattern emerged for shot scale, such that closer shots predicted an increase in subjective emotional tension but a decrease in PR and GSR.

Furthermore, splitting the omnibus analysis between live action films (*Stutterer* and *The Voorman Problem*) and animated films (*Borrowed Time* and *Paperman*) revealed that opposing directions of effect also occurred within emotion variables across film style (Table 2). Specifically, clutter negatively predicted PR in live action films but positively predicted PR in animated films. Likewise, closer shots negatively predicted GSR in live action films but positively predicted GSR in animated films. However, given the relatively mild effects across objective measures of emotional tension, as well as the small sample size within each film style category, these discrepancies are perhaps unnoteworthy. Indeed, the only cinematic variable that differed consistently between live action and animated films across all three measures of emotion was luminance, which negatively predicted joystick displacement, PR, and GSR for live action films, but exerted no influence on emotional tension (subjective or objective) for animated films.

	Live action films			Animated films		
	t	*p*	% variance explained	*t*	*p*	% variance explained
Joystick						
Clutter	-0.49	.625	14.05	-4.49	<.001*	9.68
Luminance	-13.49	<.001*	11.96	-0.42	.673	10.68
Motion	4.34	<.001*	13.81	3.62	<.001*	10.32
Shot density	3.88	<.001*	14.04	11.11	<.001*	5.02
Shot scale	8.84	<.001*	13.29	3.43	.001*	10.68
Sound amplitude	12.54	<.001*	9.01	7.89	<.001*	8.84
PR						
Clutter	-3.51	<.001*	1.22	5.25	<.001*	1.47
Luminance	-2.81	.005*	0.95	0.82	.412	1.62
Motion	3.15	.002*	1.07	1.30	.195	1.63
Shot density	-1.37	.170	1.24	-0.69	.490	1.65
Shot scale	-5.71	<.001*	0.47	-0.54	.588	1.66
Sound amplitude	1.47	.143	0.81	4.51	<.001*	1.12
GSR						
Clutter	4.91	<.001*	1.43	0.54	.593	0.93
Luminance	-3.19	.001*	1.77	-1.09	.276	1.09
Motion	-4.04	<.001*	2.59	-3.79	<.001*	0.50
Shot density	-4.00	<.001*	2.03	2.13	.033	1.08
Shot scale	-6.22	<.001*	1.03	2.78	.006*	0.75
Sound amplitude	5.52	<.001*	2.33	0.87	.386	1.09

Table 2. Results of omnibus regression analyses (by film style). Asterisks denote significant predictors, as determined by a Bonferroni corrected α of 0.008.

The robust effects of luminance may be explained by an overall difference in brightness between live action and animated films: live action films have steadily grown darker over time, while animated films have remained comparatively bright (Brunick & Cutting, 2014). Aligning with these findings, *Stutterer* and *The Voorman Problem* are significantly darker than *Borrowed Time* and *Paperman* ($t(8458) = 31.8$, $p < .001$). Thus, given the omnibus findings that darker images drive heightened emotional tension, it follows that luminance would exert a greater influence on emotional response for darker (live action) films compared to brighter (animated) films.

Table 3 summarizes significant effects across each of the four films and the omnibus analysis. Discrepancies in the presence, size, and direction of these effects across emotion variables motivated the following analysis, which explores the degree of alignment between the three measures of emotion within each film.

		Clutter	Luminance	Motion	Shot density	Shot scale	Sound amplitude
Borrowed Time (animated adventure)	**GSR**		↓	↓			
	PR		↓		↑		
	Joystick	↓			↑	↑	
Paperman (animated romance)	**GSR**						
	PR				↓	↓	↑
	Joystick		↑				↑
Stutterer (live action drama)	**GSR**	↓	↓		↑		
	PR	↓		↑			
	Joystick		↓	↑			
The Voorman Problem (live action science fiction)	**GSR**	↑	↑			↓	
	PR					↓	↓
	Joystick						
Omnibus (aggregate analysis of all four films)	**GSR**	↑	↓	↓		↓	↑
	PR					↓	↑
	Joystick		↓	↑	↑	↑	↑

Table 3. Significant effects of cinematic variables on emotional response variables across all four films. White arrows indicate a positive relationship and black arrows indicate a negative relationship.

Correlation Among Emotion Measures

Given the observed dissociation between the three measures of emotion, an analysis of the degree of correlation among the variables is warranted. For visualization purposes, Fig. 37 displays joystick displacement, PR, and GSR plotted against each other within each of the four short films.

Fig. 37. Mean z-standardized emotion measures over time for a) *Borrowed Time*, b) *Paperman*, c) *Stutterer*, and d) *The Voorman Problem*.

To assess the degree of correlation between the three emotion variables, a cross-correlation analysis was performed (accounting for a lag of ± 3). However, the resulting correlations diminished or remained relatively stable in comparison to standard correlation measures (i.e., at lag 0). The lack of any robust cross-correlation may be due to the nature of data collection in this study, as the dependent variables were recorded every five seconds, and physiological measures are not known to lag on the order of 5-second time intervals (see Benedek & Kaernbach, 2010). Thus, having eliminated the possibility that stronger correlations would emerge when accounting for physiological lag, I turned to standard correlation procedures to assess the similarity between the three measures of emotion (Table 4).

	Pearson's *r*	*p*
Joystick vs. PR		
Borrowed Time	-0.03	.299
Paperman	-0.03	.283
Stutterer	0.12	<.001
The Voorman Problem	-0.09	<.001
Joystick vs. GSR		
Borrowed Time	0.28*	<.001
Paperman	-0.03	.211
Stutterer	0.11	<.001
The Voorman Problem	-0.21*	<.001
PR vs. GSR		
Borrowed Time	-0.04	.198
Paperman	0.23*	<.001
Stutterer	0.12	<.001
The Voorman Problem	0.08	<.001

Table 4. Pearson's correlation coefficients among pairs of emotion variables within each film. Asterisks denote significant weak correlations (as defined by a threshold of $r \geq 0.20$).

Table 4 reveals very little correlation among the three emotion variables. According to standard thresholds (Campbell & Swinscow, 2009), nine of the twelve variable pairings resulted in very weak correlations (only five of which were significant). The remaining three pairings resulted in weak, but significant, correlations: joystick versus GSR in *Borrowed Time* and *The*

Voorman Problem (the latter of which is a negative relationship), and PR versus GSR in *Paperman.* This general lack of correlation is corroborated by a principle components analysis (collapsed across film; Table 5), indicating that the three indices of emotion are reasonably independent measures of psychological state.

	Joystick	PR	GSR
Joystick	1.00	-0.14	0.01
PR	--	1.00	0.18
GSR	--	--	1.00

Table 5. Correlation coefficients among emotion variable pairings (collapsed across film), as determined by a principle components analysis.

There are a number of plausible explanations for the limited alignment between measures. For instance, the lack of correlation between PR and GSR may be due to the ways in which heart rate and electrodermal activity are differentially affected by the autonomic nervous system (ANS). While both measures are tightly linked to sympathetic activity, heart rate alone is antagonistically controlled by both the sympathetic and parasympathetic branches of the ANS, producing increased or decreased PR, respectively (see Mendelowitz, 1999; Wang et al., 2018). As such, compared to GSR, PR may be more reflective of emotional states linked to dampened physiological arousal (e.g., empathy or tenderness; Gross, Frederickson, & Levenson, 1994; Marsh, Beauchaine, & Williams, 2008; Sternbach, 1962). This differentiation may partially explain the lack of correlation among PR and joystick measures as well, given that participants were instructed to move the joystick in a single direction regardless of the nature of emotional tension they felt. In other words, while the joystick measure collapsed all types of emotional response into a single construct of emotional tension, PR may have better captured the multifaceted nature of emotional reactivity. In contrast, the joystick and GSR measures may have captured emotional arousal in a more comparable manner, which aligns with the fact that two of the three most robust correlations emerged among joystick / GSR pairings. However, even these correlations were weak.

From a psychological standpoint, the general lack of correlation between the three variables is unsurprising, given evidence surrounding the differential effects of emotion regulation strategies on subjective vs. objective measures of emotion. Gross and Levenson (1993) found that emotional suppression, which involves the conscious inhibition of outward signs of emotion, increases physiological arousal among film viewers but has no impact on the subjective experience of negative emotions. On the other hand, reappraisal refers to changing one's construal of a situation in order to reduce its emotional impact, and unlike suppression, reappraisal does not appear to influence physiological arousal, but it does decrease the subjective experience of negative emotions (for a review, see Gross, 2002).

Although participants in the current study were given no indication that they should regulate their emotional response to each of the films, the mere presence of a researcher in the laboratory may have influenced them to do so (e.g., Buck et al., 1992; Lee & Wagner, 2002). As such, the discrepancies observed across emotion variables might reflect a dissociation between measures of emotion that are comparatively more sensitive to the visual structure of film (i.e., subjective measures), and those that are comparatively more sensitive to the narrative elements of film (i.e., physiological measures).

As the subjective experience of emotion is known to be dampened by cognitive reappraisal mechanisms (Gross, 2002), such mechanisms may leave more room for the influence of low-level cinematic structure on subjective emotional response (as such features are arguably less subject to reappraisal). On the other hand, given that physiological measures are unabated by reappraisal and can actually increase in magnitude during suppression, it follows that objective measures of emotion may be more sensitive to the narrative elements of film. (Indeed, the physiological correlates of emotion during film viewing reflect the granularity and strength of emotional response in the real world; Davydov, Zech, & Luminet, 2011; Kreibig, Wilhelm, Roth, & Gross, 2007). As such, narrative processing may interfere with the potential influence of low-level visual stimulation on objective emotional response measures.

In sum, cognitive understanding of the artificiality of film may render the joystick measure more sensitive to the visual structure of film as an emotional elicitor, while automatic, unconscious measures of emotional tension (i.e., PR and GSR) are likely unaffected by such reappraisal mechanisms (and therefore marked by competition between the narrative and structural elements of film, which could dilute the relative influence of cinematic visual structure on such measures). If this is the case, then the synchrony observed within all three emotion variables suggests that both visual and narrative elements play a substantial role in syncing emotional response among viewers, with the visual structure of film exerting a stronger effect on subjective emotional tension.

To conclude, these findings align with established research surrounding the differential effects of emotion regulation strategies on objective vs. subjective measures of emotion. However, further research is needed to better understand the role of low-level film structure in the dissociation observed among various measures of cinematically-induced emotion. For now, though, the peculiarities of the emotional impact of film can be further investigated through a closer look at *Paperman*, due to a unique quirk of its audience.

Naïve Versus Familiar Viewers: The Case of Paperman

Paperman was the only film that a portion of participants had seen prior to participating in the study. Twelve participants had previously seen *Paperman*; eight had not. Given the fairly even split, an analysis of the similarities and differences between groups could provide insight into the effects of familiarity on emotional response during film viewing. While one might expect that emotional response would be diminished among familiar viewers, a remarkably strong correlation emerged between familiar and naïve viewers across all three measures of emotion (Fig. 38).

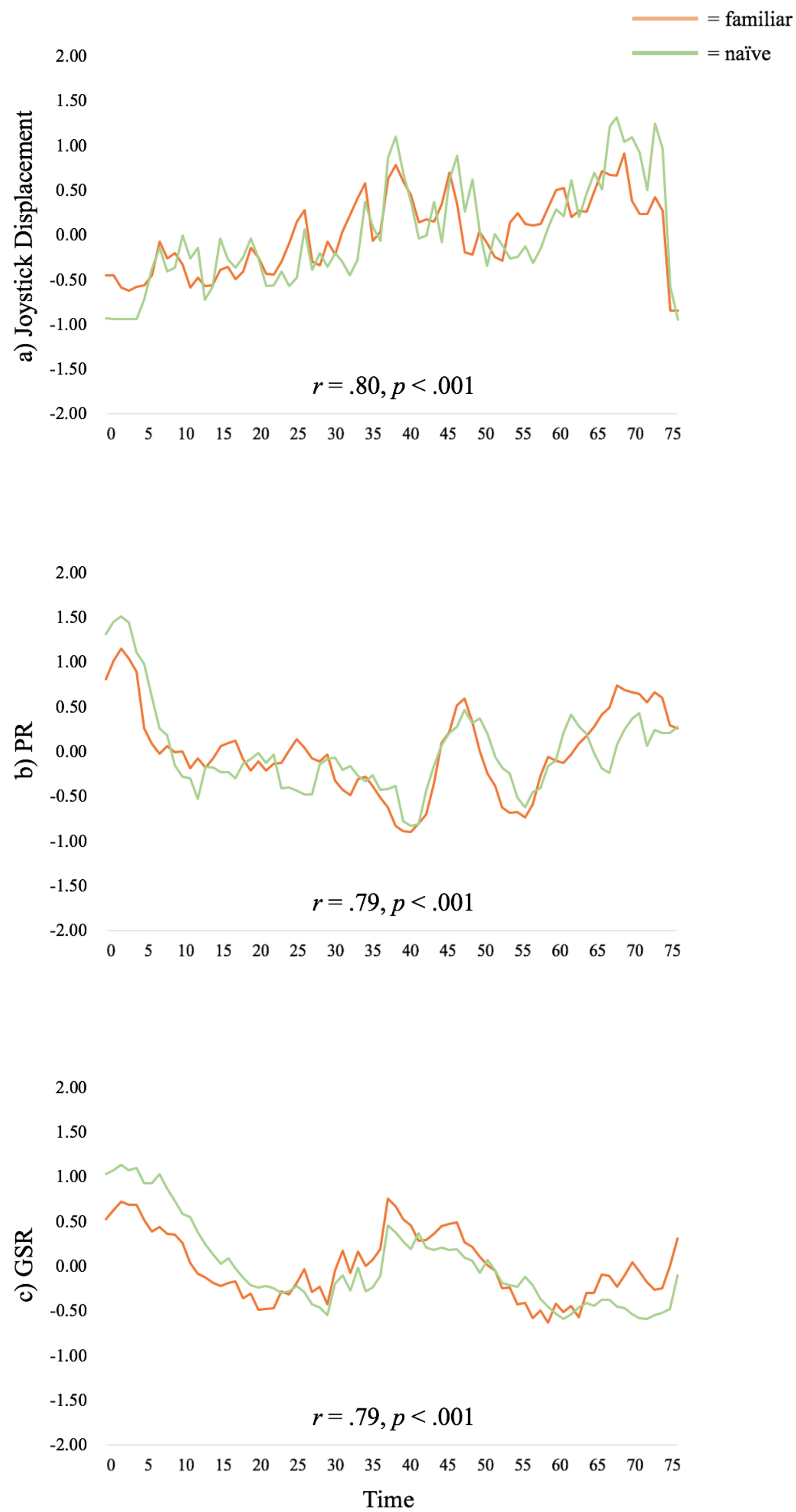

Fig. 38. Mean z-standardized emotion measures and Pearson's correlation coefficients for a) joystick displacement, b) PR, and c) GSR among viewers of *Paperman*.

Furthermore, in assessing the influence of the low-level features of *Paperman* on emotional response, very little difference was detected between familiar and naïve viewers (Table 6). Specifically, luminance predicted joystick displacement among naïve viewers only, and while sound amplitude emerged as a significant predictor of joystick displacement among both groups, it predicted PR among familiar viewers only. In contrast, shot density emerged as a significant predictor of PR among both groups, but it predicted GSR among naïve viewers only. Once again, given the relatively mild effects that were observed, these discrepancies are not especially striking, particularly as they are eclipsed by the overarching similarities between the two groups.

In sum, familiar and naïve viewers displayed remarkably similar patterns in terms of the trajectories observed for each emotion measure, as well as the influence of the visual structure of *Paperman* on emotional response. These findings suggest that film has the power to captivate regardless of familiarity, further attesting to the capacity of film to evoke synchronized responses among viewers that are at least partially grounded in the visual structure of film.

	Familiar viewers		Naïve viewers	
	t	*p*	*t*	*p*
Joystick				
Clutter	-0.13	.899	-1.46	.145
Luminance	1.35	.178	3.85	<.001*
Motion	1.33	.185	1.10	.271
Shot density	1.96	.050	0.80	.422
Shot scale	0.70	.486	-0.62	.537
Sound amplitude	3.66	<.001*	3.29	.001*
PR				
Clutter	0.75	.452	1.16	.247
Luminance	1.59	.113	0.62	.538
Motion	0.61	.539	0.70	.485
Shot density	-3.84	<.001*	-4.37	<.001*
Shot scale	-3.15	.002*	-3.73	<.001*
Sound amplitude	3.55	<.001*	0.55	.582
GSR				
Clutter	-0.02	.985	-0.94	.346
Luminance	0.57	.566	-0.81	.416
Motion	-1.61	.109	-0.80	.424
Shot density	-1.08	.278	-2.84	.005*
Shot scale	-1.47	.141	-0.80	.422
Sound amplitude	1.88	.060	0.99	.322

Table 6. Results of regression analyses across familiar versus naïve viewers of *Paperman*. Asterisks denote significant predictors, as determined by a Bonferroni corrected α of 0.008.

General Discussion

While watching a film, a curious thing happens. Viewers become immersed in a world evoked by a fictional story while the rest of the world drops away. This state, known as narrative transportation, requires that individuals are able to process the story being told as well as empathize with the story's characters (Gerrig, 1993; Green & Sestir, 2017; van Laer et al., 2014).

And filmmakers appear to structure their movies to facilitate this process; previous research suggests that low-level visual changes in Hollywood film over the last 100 years likely trigger attentional mechanisms and emotional responses that align with narrative form (see Cutting, 2016a). Luminance presents a prime example of this phenomenon. The screen is darkest about three quarters of the way through a film, reflecting the protagonist's deepest moment of despair as they struggle to achieve their goal, and brightest during the epilogue, once they have overcome their challenges (Cutting, 2016b, 2016a). However, the influence of these cinematic strategies on audience emotion has largely been unknown.

Thus, this study sought to corroborate the theorized link between the low-level visual structure of film and emotional tension among viewers. A number of findings emerged:

1. Different measures of emotional tension likely capture different aspects of emotional response. Specifically, subjective emotional tension appears to be more sensitive to low-level cinematic variables than objective measures of emotional tension.
2. Despite the discrepancy among emotion measures, both subjective and objective indices of emotional tension are remarkably synchronized across viewers. (And this synchrony persists whether or not viewers are familiar with a given film, as illustrated by the robust correlation within all three emotion measures between viewers who had previously seen *Paperman* and those who had not.)
3. The low-level structure of film partially explains this synchrony (particularly in terms of subjective emotional tension). However, the influence of cinematic variables on emotional response varies across films. This variability is likely due to stylistic differences across genre, as filmmakers utilize visual tools only to the extent that makes sense for the stories they are telling.

The third point is best illustrated through a brief summary of the findings that emerged for each of the six cinematic variables of interest:

- Clutter heightened emotional tension for *The Voorman Problem* only, suggesting that it might be most effectively utilized to suggest psychological turmoil in science fiction films, fantasy films, or psychological thrillers (all of which are defined by features present in *The Voorman Problem*). However, in order for clutter to exert this effect, shots marked by a high degree of clutter must be prolonged in order for viewers to process the content of each shot.
- Luminance affected emotional tension across the board. However, given that emotion measures in this study were designed to capture the entire spectrum of emotional experience (from positively- to negatively-valenced emotions), the direction of this effect varied across genre. Films that seemingly evoke more negative emotional responses (i.e., *Borrowed Time* and *Stutterer*) resulted in a negative relationship between luminance and emotional tension, while films that seemingly evoke more positive emotional responses (i.e., *Paperman* and *The Voorman Problem*) resulted in a positive relationship between luminance and emotional tension. This pattern of results reflects previous research suggesting that brightness elevates mood while darkness lowers it (e.g., McEnany & Lee, 2005; Pinchasov, Shurgaja, Grischin, & Putilov, 2005; Tarvainen et al., 2015; Valdez & Mehrabian, 1994).
- Motion increased emotional tension for *Stutterer* only, and this effect appears to be the result of cinematography and editing techniques. Specifically, segments of heightened motion coincide with scenes marked by handheld camera and jump cuts, known to increase emotional impact by conveying the impression of viewer presence in the action

and by amplifying the density of emotional gestures, respectively. One might expect that motion would also exert an influence on viewers of action films such as *Borrowed Time*, but a closer inspection of the film's storyline suggests that the dramatic impact of the narrative itself may have outweighed the potential influence of motion.

- Shot density appears to heighten emotional tension for films that utilize an established cinematic formula—that is, decreasing shot duration (and, consequently, increasing shot density) during the beginning of a film's climax (Cutting, 2016b, 2016a). Indeed, this structural pattern was limited to *Borrowed Time* and *Stutterer*—the only two films for which a positive relationship emerged between shot density and emotional tension.
- Shot scale appears to affect subjective emotional response such that closer-scaled shots heighten emotional tension—to a point. For *Stutterer*, medium close-up shots were indeed associated with greater subjective emotional tension, but this tension dropped off with close-up and extreme close-up shots, suggesting that shots on the extreme end of the scale impede processing of characters' facial expressions (thus corroborating previous research; Smith, 2012, 2013). Ultimately, a positive relationship between shot scale and emotional tension was limited to *Borrowed Time*, warranting further research on the role of shot scale in emotional engagement.
- In terms of sound amplitude, a closer look at the use (or nonuse) of music across the four films suggests that the amplitude of a film's score may drive heightened emotional tension for films that make use of music; in contrast, films marked primarily by diegetic sound (e.g., sound effects and dialogue) appear to drive heightened emotional tension during periods of relative silence.

What can be made of the inconsistency observed in the emotional effects of cinematic visual structure across films?

First, the fact that 24 (out of a possible 72) significant effects emerged is telling, given that each of the low-level variables studied was not investigated in isolation. All six cinematic variables were not only competing with each other, but they were also competing with elements of the narrative throughout each of the four short films. Indeed, Tan (1996, 2018) outlines a number of emotional effects that a film's narrative might exert on the viewer. For instance, by weaving uncertainty of narrative outcomes throughout a film's storyline, filmmakers address the natural propensity for humans to feel curious, and filmmakers gradually and systematically unfold the protagonist's fate in order to appeal to the viewer's ongoing interest. Additionally, the fundamental human need for bonding and belonging prompts the viewer to sympathize with characters as they navigate their circumstances on screen. On the other hand, there exist cinematically-induced emotions that rely less on narrative content and more on visual spectacle: imagery composed of explosions, cruelty, harmony, serenity, and the absurd can evoke emotional reactions of fear, disgust, awe, and aesthetic appreciation—even out of narrative context (Tan, 2018). In other words, these visceral emotional responses operate independently of both the storyline and the psychological mechanisms of empathy that narrative engagement so heavily relies on.

No matter the impetus for cinematically-induced emotion, Hollywood filmmakers seem to align the visual art of film with its narrative foundation (Cutting, 2016a). Through strategies of cinematography and editing, they demonstrate an implicit understanding of how to focus the viewer's attention and amplify their emotional responses to the most suspenseful or poignant moments of a film. So perhaps, in general, the narrative elements of a film are not competing with its low-level structure at all—when adhering to common conventions of filmmaking, filmmakers manipulate the visual tools they have at their disposal in order to maximize the emotional responses they expect to be evoked through the film's storyline.

But not all films embody this alignment between visual structure and narrative form. Indeed, the results of this study reveal that the narrative elements of a film might outweigh the emotional influence of low-level features when the two do not align (as suggested by the lack of effect of motion on emotional response among viewers of *Borrowed Time*), or vice versa (as suggested by the influence of motion on emotional response in some of the milder scenes of *Stutterer*). Perhaps, then, this mapping does not apply as precisely to short films as it does to Hollywood feature films. Furthermore, in some cases, certain low-level features may better serve attentional focus than emotional engagement per se.

These caveats render the number of effects observed in this study all the more striking. As such, the findings provide considerable support for the heretofore speculated relationship between the visual structure of film and its emotional impact on audiences. By demonstrating that six distinct cinematic variables, in isolation, account for a substantial portion of variance in the subjective emotional tension among viewers, this work illuminates the prominent role that low-level cinematic structure plays in audience engagement. Additionally, these findings provide insight into the perceptual mechanisms underlying emotional synchrony among viewers (a synchrony that nicely rounds out previous findings pointing to attentional and neural synchrony among film viewers; e.g., Carmi & Itti, 2006; Goldstein et al., 2007; Hasson et al., 2004, 2008, 2010; Itti, 2005; Mital et al., 2011; Smith & Henderson, 2008).

Second, the findings of this study build upon the empirical bridge between psychology and film theory. For decades, film theorists have crafted compelling, yet speculative, accounts of the emotional impact of cinema. For example, they have reasoned that luminance and darkness, through the art of chiaroscuro, create suspense through contrast (e.g., Hall, 1987; "What Is Chiaroscuro in Film?" 2019), and they have described how the motion created by handheld camera instills a sense of presence among viewers (see Pandža, 2018). They have also described the unusual cinematic techniques that set apart auteurs such as Michael Haneke (Walker, 2010) and Jim Jarmusch (Kulezic-Wilson, 2009), particularly in terms of how they use silence to drive an unnerving sense of tension. Implicit in these accounts are the stylistic differences that distinguish genres. And now, the differing emotional effects of low-level film structure that emerged across the films utilized in this study support not only these theoretical accounts, but also the notion that filmmakers employ these structural features differently, depending on their goals. And the findings of this study suggest their success: viewers' emotional responses are not only robust, but they are also synchronized. So, filmmakers are perhaps not only filmmakers but also amateur psychologists, demonstrating an implicit and empirically-supported understanding of how to focus viewers' attention, facilitate narrative comprehension, and now, promote emotional engagement with their films.

Critically, audience expectation also appears to govern the stylistic decisions made by filmmakers. For example, while Hollywood films intended for adult audiences have grown darker over time (Cutting, Brunick, DeLong, et al., 2011), children's animated films have maintained a steady level of brightness, and children's live action films have increased in brightness, suggesting that filmmakers manipulate luminance to appeal to audience preferences (Brunick & Cutting, 2014). Indeed, *Paperman* is arguably the most child-friendly of the films utilized in this study, and it is also the brightest. Whether luminance heightens subjective emotional tension among children is a question for future research, as evidence of such an effect would provide support for the notion that filmmakers implicitly manipulate the visual tools of cinema in order to maximize emotional engagement among particular audiences. Furthermore, if children do exhibit patterns of emotional response to luminance that are similar to those of adults, then subsequent research should investigate whether baseline luminance exerts a differential effect on viewers of all ages, such that for relatively darker films, increased darkness drives emotional tension, and for relatively brighter films, increased brightness has the same effect.

Additionally, while the results of this study offer supporting evidence for several theoretical accounts linking cinematic structure and audience emotion, a larger corpus of films would go a long way in corroborating these findings. In the meantime, this first pass offers an exciting glimpse into the alignment of film theory and psychological research.

Finally, this study offers empirical support for Busselle & Bilandzic's (2009) model of narrative engagement from a perceptual perspective. While previous research supports the first and second dimensions of Busselle & Bilandzic's (2009) model (attentional focus and narrative understanding; e.g., Carmi & Itti, 2006; Cutting, Brunick, & Candan, 2012; Goldstein et al., 2007; Hasson et al., 2008; Itti, 2005; Mital et al., 2011; Smith & Henderson, 2008), perceptual evidence surrounding the third dimension (emotional engagement) has been sparse. These findings help to fill that void by demonstrating that viewers' emotional responses to film are not only synchronized, but that this synchrony is driven in part by the visual structure of film. Thus, in the context of film, emotional engagement is not the sole province of the narrative—cinematically-induced emotion is verifiably steeped in sensory stimulation as well.

Future Directions and Social Implications

Future research should seek to replicate these findings on a larger corpus of films, such that multiple exemplars of each genre can be utilized to corroborate these effects on a broader scale.

This work could also be extended in new directions, as it carries relevance in a variety of domains. While one might assume that the findings could be used to construct a practical guide to filmmaking, this would not be necessary. After all, these findings emerged from the work of filmmakers themselves (which is perhaps unsurprising, given that filmmakers are humans with experience perceiving the world, and as a result, they apply principles of human perception to the cinema screen).

On the other hand, guidelines derived from these findings may prove fruitful in the realm of intercultural communication, where meaning is often lost in translation. Indeed, what is lost frequently relates to emotion, even if tangentially (e.g., tone, nuance, and humor; see Cronin, 2008; Sinkeviciute, 2017). Perhaps, then, these losses can be partially recovered through the careful consideration of low-level features in the construction of cross-cultural visual media. Of course, an important intermediate step would be to determine how the emotional influence of the visual structure of film varies across cultures, particularly given long-standing models that distinguish low-context and high-context cultures (Hall, 1976). As previously discussed, varying degrees of emphasis on contextual information in social interaction suggest that individuals across cultures may be differentially sensitive to the manipulation of shot scale in film, which can expand or constrict the viewer's access to a character's emotional gestures. It follows that a trend toward closer-scaled shots in Hollywood film would be detrimental to emotional engagement among viewers from high-context cultures (but not low-context cultures). If this were the case, then the low-level structure of visual media would need to be shaped to appeal to different cultural influences on perceptual processing (in order for these tools to be beneficial to intercultural communication at all)—especially since filmmakers appear to implicitly manipulate low-level variables from their own frame of reference.

This work also speaks to the dynamic field of museum science, which reflects a growing focus on interactive installations (see Wyman, Smith, Meyers, & Godfrey, 2011). With technological advancements bringing museum visitors closer to the content that modern exhibits showcase, such exhibits might benefit from evidence surrounding the emotional impact of the low-level features of screen technologies. Indeed, a market exists for immersive exhibits that borrow from the language and tools of cinema in order to evoke strong emotional responses to scientific ideas and stories (see Snibbe & Raffle, 2009). However, it is important to note that the findings of the present study are derived from traditional narrative film. To best apply these results to museum settings, it may be worthwhile to investigate whether alternative cinematic formats behave in the same way. For instance, how does the visual structure of documentaries, avant-garde films, and surrealist cinema influence emotional response among viewers? These formats may better reflect some of the immersive exhibits featured in modern museums.

In order to facilitate expansion of this work across various cultures and contexts, the following study addresses the lack of diversity in Hollywood film that constrains the current findings—especially in terms of their applicability to media formats that may be vulnerable to group-based perceptual biases.

CHAPTER 2: CREATION OF A RACIALLY REPRESETATIVE DATABASE OF FACIAL EXPRESSION STIMULI

The results of Study 1, while robust, are limited to films that feature predominantly White male characters. Given the growing diversity of Hollywood film (Hunt & Ramón, 2020), it is worthwhile to assess whether the findings that emerged in Study 1 replicate across storylines that center women and, especially, people of color. Such storylines are currently not commonplace; however, one way forward is to investigate the influence of low-level cinematic variables on viewer engagement with a diverse corpus of facial stimuli (to serve as a proxy for film characters). Expanding the current work in this way will provide insight into whether the effects of film structure on viewer engagement are affected by group-based biases in face perception (i.e., own- and other-race effects; see Meissner & Brigham, 2001).

While diverse facial stimuli exist for open use among face perception researchers, few collections have been validated for racial representativeness. As such, the next step toward investigating the influence of film structure on emotional engagement with diverse characters is to create and validate a racially representative collection of facial stimuli. Furthermore, given that film is largely constructed around the faces of characters reacting to various situations (see Cutting & Armstrong, 2018; Plantinga, 2009; Tan, 1996), and that facial expressions of emotion are not always clear (Cutting & Armstrong, 2018), future research should target faces that vary across not only gender and race, but also emotional expression and intensity.

While the creation of these stimuli is motivated by the findings of Study 1, I have approached the stimulus set as a general purpose, open-use tool for all face perception researchers. Thus, these stimuli not only serve the research purposes outlined above, but they also overcome common obstacles encountered in the broader realm of face perception research.

Study Overview

For face perception researchers (particularly those who study facial expressions of emotion), a methodological problem persists in the selection of stimuli: one must choose between photographs of real people or computer-generated likenesses of human faces, each presenting a unique set of obstacles. While computer-generated images raise concerns surrounding ecological validity, the use of photographs to capture real facial expressions (typically produced either by actors or by research participants induced to feel particular emotions) introduces a problem of representation: do the captured expressions display the precise musculature known to map onto basic facial expressions of emotion (Ekman & Friesen, 1975), or are the expressions ambiguous in some way (see Barrett, Adolphs, Marsella, Martinez & Pollak, 2019)? Furthermore, do features of the actor's face (e.g, birthmarks, freckles, hairstyle) distract from the task of identifying, categorizing, or discriminating her facial expressions (see MacLin & Malpass, 2001)? While these challenges can be mitigated through the use of computer-generated images, on the other end of the spectrum, the absence of such features may result in an unrealistic likeness of the human face that appears cartoonish or uncanny (Ferrey, Burleigh, & Fenske, 2015; Mori, 1970).

Additionally, a growing focus on ingroup vs. outgroup influences on face perception (for a review, see Hugenberg, Young, Sacco, & Bernstein, 2011) necessitates a body of stimuli that depict, among other factors, reasonable representations of race. However, given that perceptual boundaries between racial groups are vague (Tishkoff & Kidd, 2004) and dynamic (depending on economic conditions, for example; Krosch & Amodio, 2014), accurate racial representation is not straightforward. Furthermore, the relative contributions of racial cues vs. skin tone per se (i.e., luminance) toward perceptual biases are difficult to parse, resulting in potential confounding factors when investigating the influence of race on face perception.

Finally, a problem of image quality arises for researchers who require stimuli depicting varying degrees of emotional intensity or stimuli that are morphed from one expression to another: basic morphing programs often blur or otherwise distort the facial features depicted in both photographic and digital images. For example, a continuum of expressions generated between an image of a closed-mouth neutral expression and one depicting an open-mouthed smile might result in pink-tinged teeth in the intermediate images due to the lip color in the anchor photograph of neutral expression. Consequently, the resulting images must be polished, calling for image editing skills that require ample time to learn and apply consistently across stimuli.

In sum, the obstacles described above are neither trivial nor tangential to the questions that motivate face perception research, as these obstacles permeate the very target of perceptual investigation—the face as visual stimulus. Thus, in response to the need for a convincing, racially diverse set of facial stimuli covering a comprehensive range of basic emotional expressions, emotional intensities, and morphed expressions, I have created a large database of images (called the International Faces) that combat the problems described above—which I refer to throughout this paper as problems related to, respectively, human likeness, racial representativeness, luminance control, and clarity of expression.

Methods

I approached the creation of this database in three distinct phases: initial rendering of images, validation of racial representativeness, and expansion of the stimulus set—each of which tackles one or more of the core problems described above. First, I digitally rendered a series of basic facial expression images across four racial categories (focusing on the problems of human likeness, racial representativeness, and luminance control). Then, I conducted a validation study to verify accurate racial representativeness of each image, which then gave me license to expand on the range of emotional expressions captured by the stimuli (while addressing the problem of clarity of expression).

Rendering of Stimuli

I created the stimulus set using FaceGen Modeller Core 3.14 (Singular Inversions, 2016), which is based on the empirical study of face perception and has a number of functions that provide users with precise control over the appearance of facial stimuli.

FaceGen stimuli are generated from a digital "face space" based on principal components analysis of 273 high-resolution three-dimensional face scans, resulting in 80 shape dimensions and 50 color dimensions—thus simulating a multi-dimensional cognitive face space, not unlike that proposed by Valentine (1991). From this face space, FaceGen users are able to do the following: generate random facial stimuli within specified gender and racial categories; further manipulate age, gender, and racial attributes of the images; make fine-grained adjustments to physiognomy and skin tone; and apply a range of facial expressions of varying intensity to the facial images.

I began by creating eight distinct faces of neutral expression, including one female and one male for each of four racial categories: Black, White, South Asian, and East Asian. While these racial categories by no means capture the richness of racial diversity across the globe, I selected these groups to allow for the control of a perceptual phenomenon relevant to face perception: that luminance affects evaluative judgments of stimuli, such that—in a stroop-like manner—brighter stimuli are often associated with positive valence and darker stimuli with negative valence (Meier, Robinson, & Clore, 2004). Critically, this effect extends to person perception; for instance, individuals expect that perpetrators of immoral acts will have a darker skin tone, regardless of race (Alter, Stern, Granot, & Balcetis, 2016).

However, given that the average luminance of Black faces is comparable to that of South Asian faces, and the average luminance of White faces is comparable to that of East Asian faces (Alaluf et al., 2002; Brunsting & Sheard, 1929), if carefully controlled (as described below), one can reasonably discount luminance as a confounding factor when observing perceptual biases in facial processing across each of these racial pairings. In other words, controlling for luminance allows for a more convincing analysis of the contribution of racial cues over skin tone per se when probing ingroup vs. outgroup effects on face perception.

Thus, to create faces representative of each of these four racial categories, I utilized the Generate function of FaceGen, which randomly produces a face from the stipulated gender and racial group. Given that I privileged averageness over randomness in an attempt to achieve convincing likenesses within each racial category, I browsed numerous images generated by FaceGen, and, after inspecting each image (against comparison images gathered from a Google search), I used my best judgment to select the faces that appeared to depict the most credible prototypes for each racial category.

By default, FaceGen stimuli are generated with statistically reconstructed skin textures that tend to be very smooth, resulting in a cartoonish quality. I addressed the artificial appearance of the selected faces by utilizing FaceGen's Texture function, which adds small-scale detail to the skin, including minor wrinkles, ruddiness, freckles, and—for men—very short facial hair (i.e., the “five o’clock shadow”), thus addressing the previously discussed problem of preserving human likeness in digitally-rendered facial stimuli. Given the range of textures that FaceGen offers, I avoided those with markings that might be distracting during perceptual tasks, including moles, blemishes, and a deep flushing of the cheeks that could signal emotional arousal (for a review, see Crozier & Jong, 2012).

After adding a different texture to each of the eight faces, I adjusted skin color in order to equate luminance between Black and South Asian faces and between White and East Asian faces. To do so, I computed the mean luminance within each racial pairing of male faces by importing the raw stimuli into Adobe Photoshop and generating a luminosity histogram for each image (on the 8-bit scale of 0-255). The mean luminosity value was 68.18 for the Black and South Asian male faces, and 87.85 for the White and East Asian male faces. I then used the Skin Shade – Dark/Light slider in the Color tab of FaceGen to either lighten or darken the pigmentation of each of the male faces until the luminosity matched the respective mean value stated above (+/- 0.03). As a result, luminance for the Black and South Asian male faces was equated at 68.18, while luminance for the White and East Asian male faces was roughly equated at values of 87.84 and 87.87, respectively (Figure 39).

I took a slightly different approach to adjusting skin color among the four female faces. Rather than focusing on absolute luminance, I measured the mean difference in luminance between the adjusted male face and the original female face within each racial category (again, by generating luminosity values in Adobe Photoshop). Female faces were—on average—10.02 units brighter than male faces in the same racial category, aligning with extensive findings that female skin tone is generally lighter than male skin tone across most racial groups (for a review, see Fitzmaurice & Maibach, 2010). Next, I repeated the skin color adjustment procedure that I applied to the male faces, using the Skin Shade – Dark/Light slider in FaceGen to either lighten or darken the pigmentation of each female face until its luminosity was 10.02 units greater than the male face of the corresponding race (+/- 0.03). As a result, luminance for the Black and South Asian female faces was equated at 78.20, while luminance for the White and East Asian female faces was roughly equated at values of 97.85 and 97.87, respectively (Figure 39). In sum, while absolute luminance varied across racial pairings (i.e., Black/South Asian vs. White/East Asian), the luminance distance between the male face and the female face within each racial category was equated, ensuring yet another level of luminance control.

While manipulating skin pigmentation was an important step toward addressing the psychological effects of luminance, issues regarding the racial representativeness of the stimulus set remained. Indeed, by virtue of being White and living in a relatively homogenous area of the United States, I am limited by my own perceptual experiences. Robust empirical findings suggest that people are better able to distinguish individuals of the racial group with whom they have the most contact (often their own race) over individuals of other races—a phenomenon known as the other-race effect (for a review, see Meissner & Brigham, 2001). As such, my judgment of racial prototypicality is imperfect.

I addressed this limitation by consulting with individuals who identify with the racial categories represented in my stimulus set. Specifically, given the nature of other-race effects, I sought the guidance of those who were either raised or currently embedded in environments with heavy representation of Black, South Asian, and/or East Asian groups. The individuals who agreed to assist me—including one individual who identifies as East Asian (Chinese) and one individual who identifies as half Indian and half African-American and resides in a predominantly Black community—each participated in a two-hour session to modify the stimuli corresponding to their own racial backgrounds. After providing a brief tutorial on the functionality of FaceGen, I worked collaboratively with these individuals to adjust the facial morphology in each image to best reflect a facial structure representative of the respective racial group. These individuals also offered guidance in adjusting the temperature of the skin tone (warm vs. cool) depicted in each image (all adjustments were made while holding luminance constant to preserve the criteria described above). The resulting eight images, comprising what I refer to as the "core" stimuli, are displayed in Figure 39.

Figure 39. Initial core stimuli with associated luminance values. All images were produced using FaceGen Modeller Core 3.14 (Singular Inversions, 2016).

Once the final adjustments to skin texture, skin pigmentation, and facial structure were complete, I transitioned to my goal of creating a stimulus set that covers a broad range of emotional expressions. I used the Expression sliders in FaceGen to generate seven versions of each face—each version depicting a different facial expression of emotion at 100% intensity, including the six basic emotions of anger, disgust, fear, happiness, sadness, and surprise (Ekman & Friesen, 1975), along with the original neutral expression (resulting in a total of 56 images). Although recent evidence has called into question the discrete and universal nature of basic emotions (see Barrett et al., 2019), I chose these facial expressions because they represent a wide range of commonly experienced emotional states, even if not perfectly unambiguous. Additionally, their implementation in FaceGen is based on extensive scientific research; FaceGen generates these expressions according to the Facial Action Coding System developed by Ekman & Friesen (1978), which meticulously describes the anatomical basis of facial movement related to the physical expression of emotion (again, with the caveat that this system may not capture the richness of facial expressions of emotion as they occur in the wild). To illustrate the range of emotion represented in the stimulus set, the six basic facial expressions generated for the East Asian male face are displayed in Figure 40.

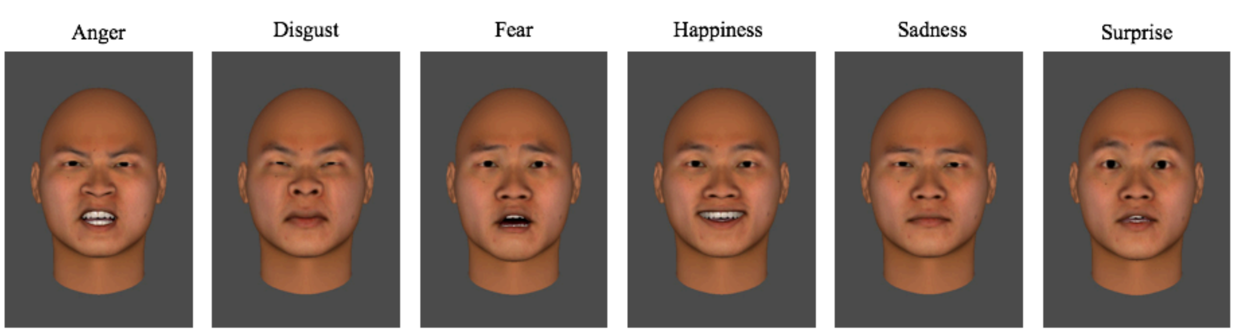

Figure 40. Six basic emotional expressions displayed by the East Asian male face. All images were produced using FaceGen Modeller Core 3.14 (Singular Inversions, 2016).

As a final step toward achieving human likeness across the stimulus set, I utilized Photoshop to add hair to each of the 56 stimuli. The hair images were located through a Google search for open-source Photoshop (PSB) files and edited to ensure that the layering of hair over FaceGen images did not obscure relevant facial features of the stimuli. After adding the hair files, I changed the default gray background of each image to white to ensure that the stimuli more seamlessly integrate into testing displays. I also cropped the bottom of each image to remove the "floating head" appearance of the original stimuli. The finalized core stimuli are displayed in Figure 41.

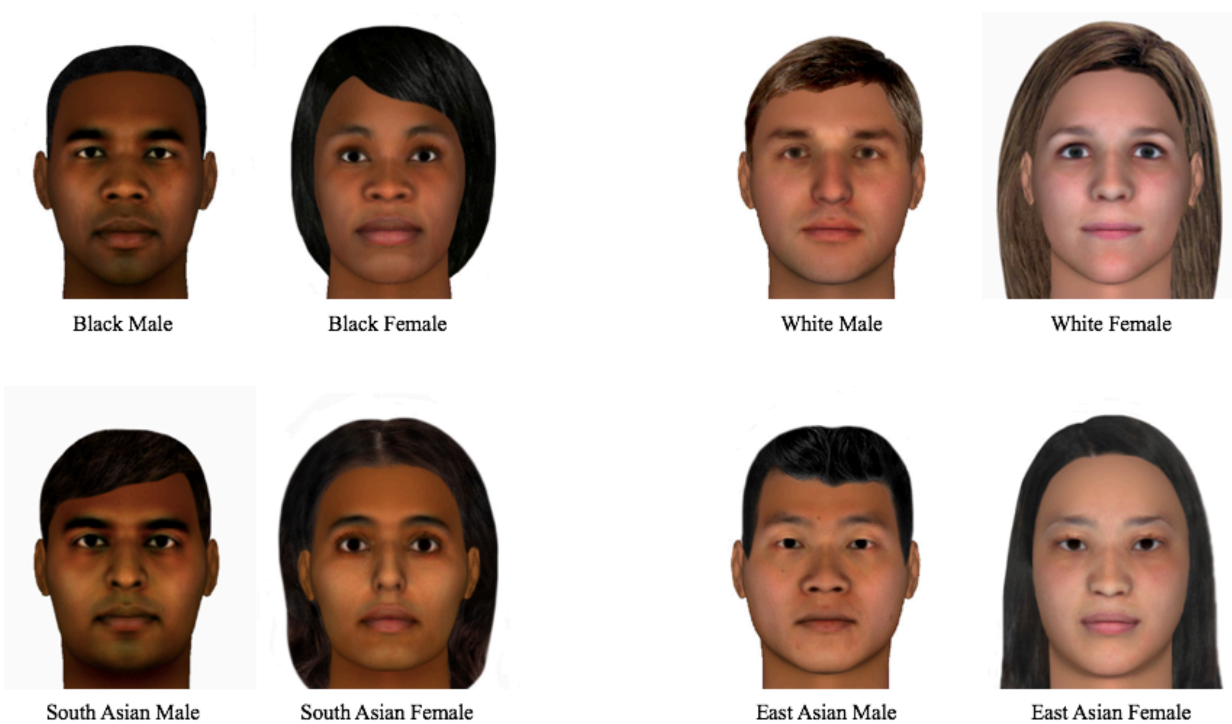

Figure 41. Finalized core stimuli. All images were produced using FaceGen Modeller Core 3.14 (Singular Inversions, 2016) and Adobe Photoshop.

Validation of Racial Representativeness

To ensure that this stimulus set consisted of convincing portrayals of Black, White, South Asian, and East Asian exemplars, I asked a geographically diverse group of participants to judge the degree to which each image represented the intended racial category. Given that facial expressions of emotion are known to influence perception of race and vice versa (e.g., Hugenberg & Bodenhausen, 2004; Li & Tse, 2016), I sought to validate racial representativeness for each face across all seven emotional expressions.

Methods

To locate individuals with perceptual expertise for Black, White, South Asian, and East Asian faces, I used Amazon Mechanical Turk (MTurk) to recruit a target of 200 English-speaking participants living in countries where the majority race matched these particular racial categories—including, respectively, 50 participants from Sub-Saharan Africa, 50 participants from the United States and Canada, 50 participants from South Asian countries, and 50 participants from East Asian countries (see Appendix B). The location of participants was restricted via MTurk settings, such that only users registered with MTurk in the specified regions could view and participate in the study (although users were not aware of these location restrictions).

After consenting to the study, participants were directed to a Qualtrics survey that presented each of the 56 images, one at a time and in randomized order, below the following text: On a scale of 1 to 7, how convinced are you that this person is of (European/African/East Asian/South Asian) descent? (1 = not at all convinced, 7 = very convinced). Participants used a digital slider to indicate their response to each image. At the end of the task, participants were asked to provide demographic information, including age, gender, race, country of residence (to verify that participants completed the survey from the regions in which they were registered with MTurk), and their MTurk ID numbers (to enable detection of duplicates). On average, the survey took 7.42 minutes to complete. As per MTurk recommendations, participants were paid $1.50 USD for their time.

In advance of data collection, I reasoned that the cut-off value for race validation should be above the neutral midpoint of the scale while also allowing for conservative response tendencies. Thus, I set a value of 5.00 as the validation standard.

Analysis

Given my intention to infer a broad, cross-cultural perception of race among this stimulus set, I reasoned that mean ratings with 95% confidence intervals would be the most informative measure for these purposes. However, the ratings were non-normally distributed with a strong negative skew (-0.936, SE = 0.022), as confirmed by the Kolmogorov-Smirnov test (.213, $p < .001$). Natural log transformation of the data did not successfully normalize the distribution (.229, $p < .001$). Given that confidence intervals are typically constructed under assumptions of normality, I performed a bootstrap analysis in R to achieve a normally-distributed set of resamples (running 10,000 bootstrap replicates for each data subset of interest), thus allowing population parameters to be inferred. Indeed, bootstrapping is frequently recommended as an approach to address issues of non-normality (see Desharnais, Camirand-Lemyre, Mireault, & Skinner, 2015; Pek, Wong, & Wong, 2017). I calculated mean values on the resampled data, along with bias-corrected and accelerated (BCa) confidence intervals to correct for any remaining skew in the distribution of bootstrap estimates.

Results

While I intended to recruit 200 participants, the Qualtrics survey was completed a total of 231 times (likely due to simultaneous sign-ups on MTurk). However, five participants completed the survey twice (as determined by their MTurk ID numbers), so I excluded the second instance of each duplicate. An additional five participants were excluded for failing to provide their MTurk ID numbers (as I were unable to detect duplicates among these participants). Finally, seven participants were excluded for reporting countries of residence outside of the specified regions. These participants reported living in North African countries (four in Algeria, two in Morocco, and one in Egypt) rather than Sub-Saharan African countries. Given that the population of North Africa reflects a heterogenous racial profile (including West Asian and Caucasian influence; Maca-Meyer et al., 2003), I were not convinced that these participants would have the desired perceptual expertise for Black faces. Therefore, I analyzed the data for 214 participants: 57 from Sub-Saharan Africa, 56 from the United States and Canada, 51 from South Asian countries, and 50 from East Asian countries. Sixty-eight participants were female, and the mean age was 32 years, with a range of 18 to 64. A complete breakdown of demographic information (including the countries of residence and racial self-identification of all participants) is provided in Appendix B.

Figure 42 captures the bootstrapped mean ratings and confidence intervals for each of the core stimuli (collapsed across facial expression and geographical location of raters). Table 7 further breaks down these values by facial expression (again collapsed across geographical location of raters). Seven of the eight core stimuli, along with their associated facial expressions, resulted in mean ratings greater than 5.00. Only the South Asian female face resulted in an aggregate mean rating below 5.00 (M = 4.71, 95% CI = 4.63, 4.79), as did all seven of her facial expressions, ranging from 4.32 [4.11, 4.53] for disgust to 4.88 [4.67, 5.08] for happiness.

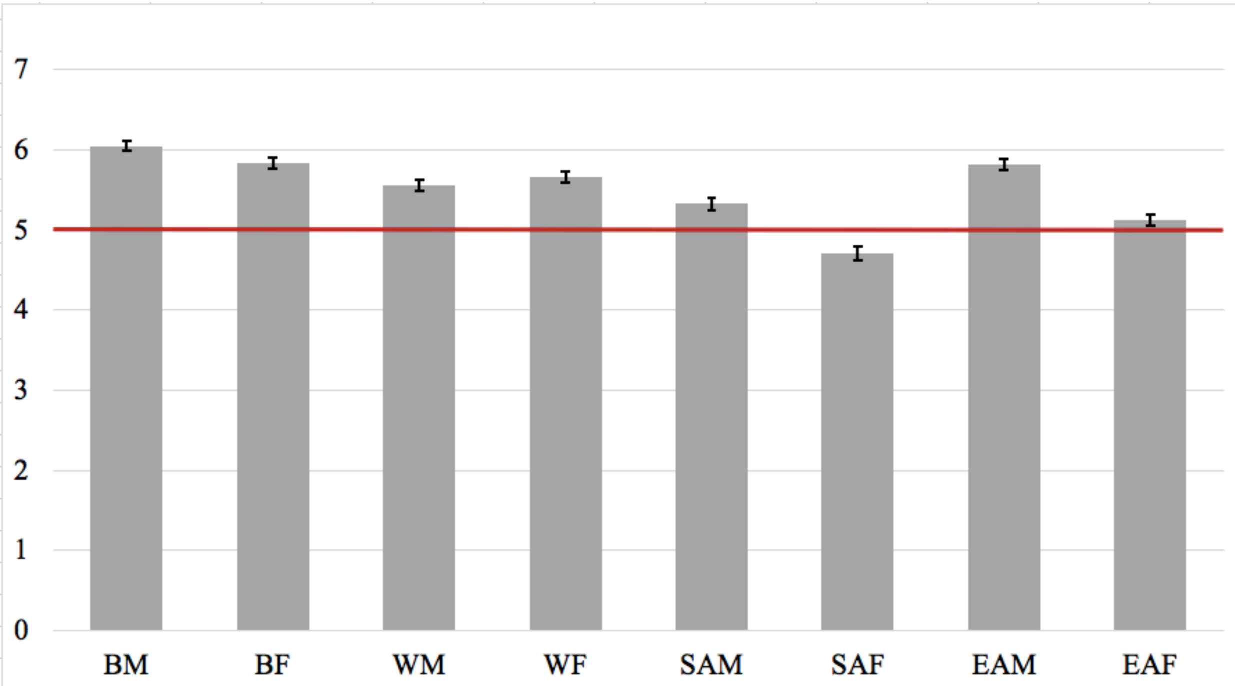

Figure 42. Bootstrapped mean ratings for core stimuli, collapsed across facial expression and geographical location of raters (B = Black; W = White; SA = South Asian; EA = East Asian; M = Male; F = Female). Error bars represent bootstrapped 95% confidence intervals. All values are derived from 10,000 bootstrapped replicates. The red line marks the validation criterion of 5.00.

	Core Stimuli							
	Black Male	**Black Female**	**White Male**	**White Female**	**S. Asian Male**	**S. Asian Female**	**E. Asian Male**	**E. Asian Female**
Neutral	6.09 [5.93, 6.25]	5.95 [5.79, 6.11]	5.77 [5.60, 5.94]	5.88 [5.71, 6.04]	5.43 [5.22, 5.63]	4.88* [4.66, 5.09]	5.89 [5.72, 6.05]	5.27 [5.08, 5.46]
Anger	5.94 [5.76, 6.11]	5.61 [5.42, 5.79]	5.36 [5.17, 5.55]	5.54 [5.35, 5.72]	5.16 [4.95, 5.36]	4.59* [4.38, 4.79]	5.74 [5.58, 5.90]	5.06 [4.85, 5.26]
Disgust	5.89 [5.71, 6.07]	5.87 [5.69, 6.04]	5.12 [4.91, 5.32]	5.27 [5.07, 5.46]	5.02 [4.81, 5.22]	4.32* [4.11, 4.53]	5.62 [5.43, 5.80]	5.06 [4.86, 5.25]
Fear	6.08 [5.92, 6.23]	5.68 [5.50, 5.86]	5.63 [5.45, 5.81]	5.74 [5.56, 5.91]	5.36 [5.16, 5.55]	4.68* [4.47, 4.89]	5.80 [5.63, 5.96]	5.14 [4.95, 5.32]
Happiness	6.13 [5.97, 6.28]	5.87 [5.70, 6.03]	5.60 [5.43, 5.77]	5.75 [5.58, 5.91]	5.28 [5.14, 5.42]	4.88* [4.67, 5.08]	5.89 [5.74, 6.04]	5.00 [4.80, 5.20]
Sadness	6.05 [5.88, 6.21]	5.92 [5.75, 6.09]	5.60 [5.43, 5.77]	5.75 [5.58, 5.91]	5.48 [5.29, 5.67]	4.78* [4.57, 4.99]	5.92 [5.77, 6.06]	5.26 [5.07, 5.45]
Surprise	6.11 [5.95, 6.26]	5.82 [5.64, 5.99]	5.65 [5.48, 5.82]	5.63 [5.45, 5.81]	5.44 [5.24, 5.63]	4.80* [4.58, 5.01]	5.78 [5.60, 5.95]	5.05 [4.85, 5.24]
Aggregate	**6.05 [5.99, 6.11]**	**5.83 [5.76, 5.89]**	**5.56 [5.49, 5.62]**	**5.66 [5.59, 5.72]**	**5.33 [5.25, 5.40]**	**4.71* [4.63, 4.79]**	**5.82 [5.75, 5.88]**	**5.13 [5.06, 5.20]**

Table 7. Bootstrapped mean ratings [95% CI] by facial expression (collapsed across geographical location of raters). All values are derived from 10,000 bootstrapped replicates. Asterisks denote mean ratings that fall below the validation criterion of 5.00.

However, when isolated to raters living in South Asian countries, bootstrapped mean ratings for all of the South Asian female facial expressions—except for that of disgust—grew to greater than 5.00 (Figure 43). Table 8 further illustrates that, of the four geographical regions represented, only participants living in South Asia rated the South Asian female face as a reasonable depiction of an individual of South Asian descent. This pattern is not due to an overall trend of South Asian raters responding liberally to the stimuli, as the mean South Asian rating for all 56 stimuli ($M = 5.32$, $SD = 1.26$) was significantly lower than the mean rating for all 56 stimuli across the other three geographical regions combined (M = 5.57, SD = 1.48) [t(11982) = -7.99, p < 0.001, *d* = 0.18].

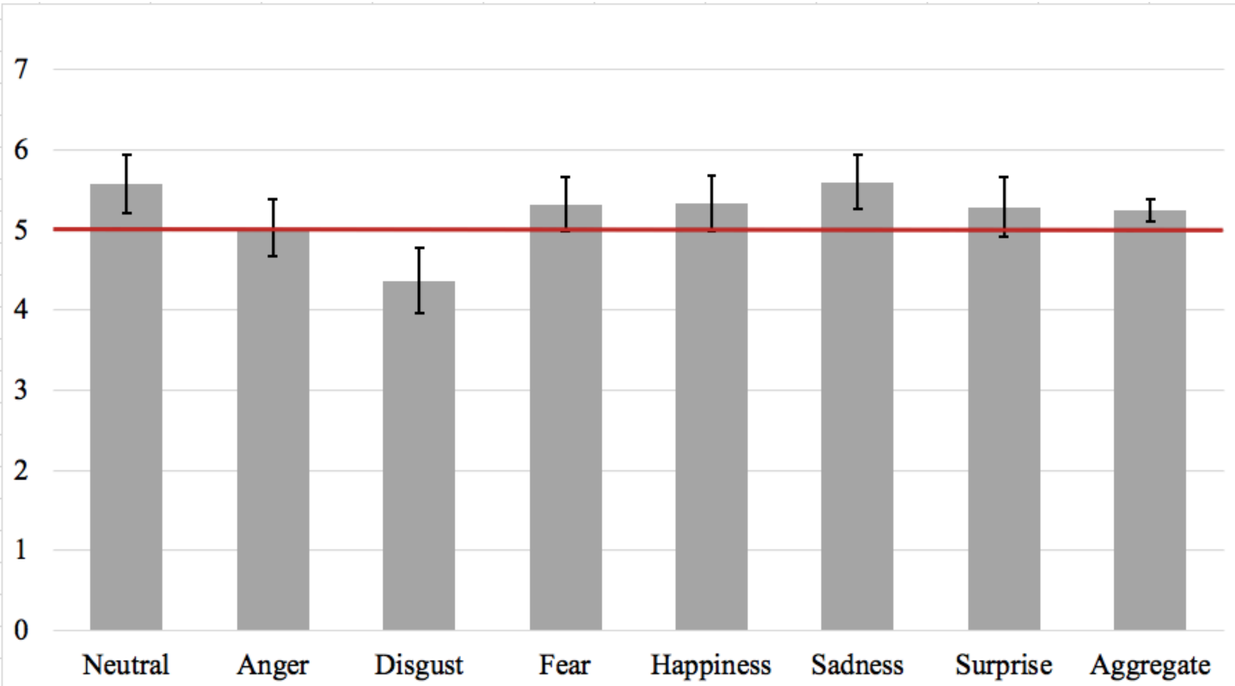

Figure 43. Bootstrapped mean ratings for South Asian female facial expressions, isolated to raters living in South Asia. Error bars represent bootstrapped 95% confidence intervals, derived from 10,000 bootstrapped replicates. The red line marks the validation criterion of 5.00.

	Geographical Location of Raters			
	Africa	**North America**	**South Asia**	**East Asia**
Neutral	4.44 [3.97, 4.90]	4.48 [4.07, 4.88]	**5.58 [5.22, 5.94]**	5.00 [4.56, 5.44]
Anger	4.44 [3.98, 4.90]	4.24 [3.84, 4.63]	**5.03 [4.67, 5.39]**	4.57 [4.20, 4.94]
Disgust	4.25 [3.81, 4.68]	4.31 [3.89, 4.73]	**4.37* [3.96, 4.77]**	4.30 [3.90, 4.70]
Fear	4.66 [4.21, 5.11]	4.04 [3.63, 4.45]	**5.32 [4.98, 5.65]**	4.69 [4.30, 5.08]
Happiness	4.73 [4.26, 5.19]	4.18 [3.77, 4.59]	**5.33 [4.98, 5.67]**	5.26 [4.94, 5.58]
Sadness	4.70 [4.28, 5.12]	3.99 [3.57, 4.41]	**5.60 [5.26, 5.94]**	4.77 [4.38, 5.16]
Surprise	4.69 [4.21, 5.16]	4.18 [3.73, 4.63]	**5.29 [4.92, 5.65]**	5.01 [4.70, 5.32]
Aggregate	4.59 [4.41, 4.76]	4.22 [4.06, 4.38]	**5.25 [5.11, 5.39]**	4.84 [4.69, 4.98]

Table 8. Bootstrapped mean ratings [95% CI] for South Asian female facial expressions, split by geographical location of raters. All values are derived from 10,000 bootstrapped replicates. Asterisks denote South Asian ratings that fall below the validation criterion of 5.00.

Figure 44 captures the bootstrapped mean ratings and confidence intervals for each of the seven facial expressions (collapsed across stimuli and geographical location of raters). Two facial expressions resulted in significantly lower racial validation ratings than the other facial expressions combined: anger [$t(213) = -5.36$, $p < 0.001$, $d = 0.37$] and disgust [$t(213) = -7.44$, $p < 0.001$, $d = 0.51$].

A complete geographical breakdown of mean ratings and confidence intervals across core stimuli and facial expression is provided in Appendix C.

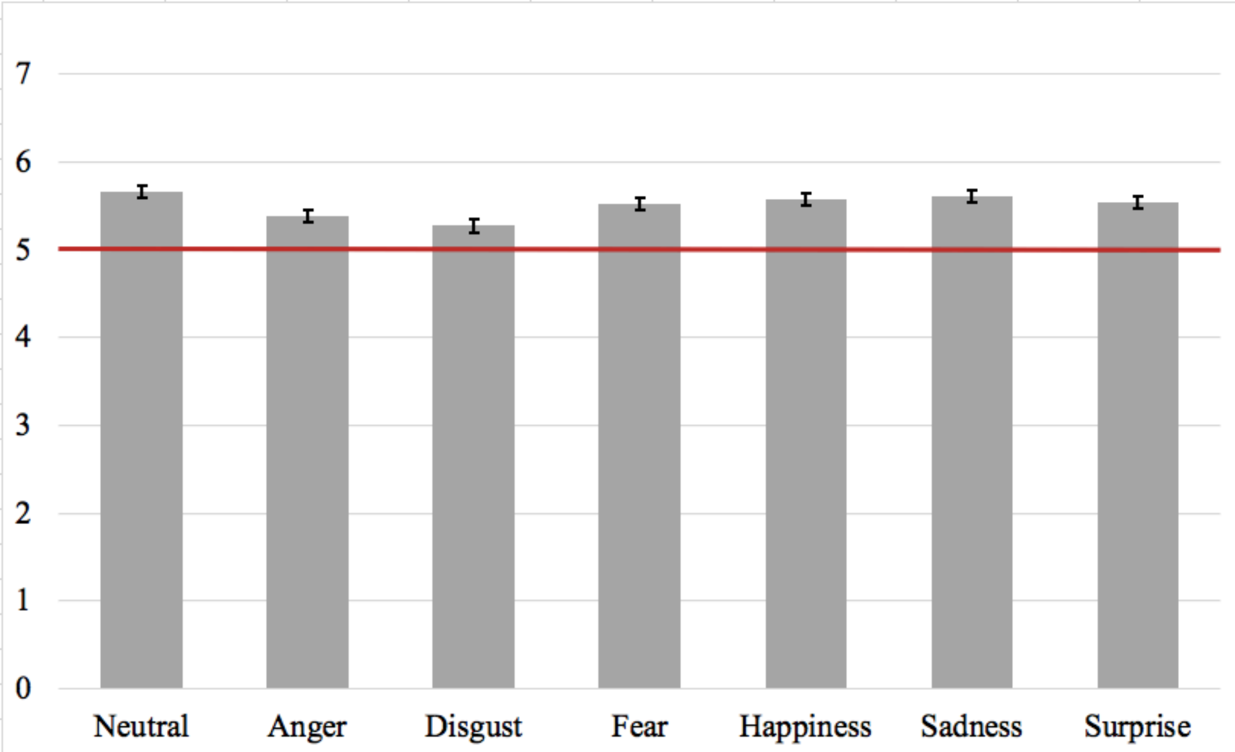

Figure 44. Bootstrapped mean ratings by facial expression (collapsed across stimuli and geographical location of raters). Error bars represent bootstrapped 95% confidence intervals, derived from 10,000 bootstrapped replicates. The red line marks the validation criterion of 5.00.

Interim discussion

Given the presumed perceptual expertise that individuals living in South Asia exhibit for South Asian faces, their responses arguably carry more weight for the South Asian female stimuli than responses provided by participants living outside of this region. Thus, I consider 55 of the 56 stimulus images to be sufficiently validated for racial representativeness.

Only the South Asian female facial expression of disgust does not appear to be a convincing representation of the intended racial category, even among raters living in South Asia. This is unlikely related to cultural nuance in the expression of disgust. Ninety-eight percent of participants in South Asia reported living in India. Even though Hindu Indians (who account for the largest religious group in India; “India’s religions by numbers,” 2015) associate disgust primarily with moral violations (in comparison to Westerners, who more often associate disgust with food; Rozin, 1999), Chapman et al. (2009) found that exposure to moral violations triggers a pattern of facial motor activity similar to that evoked by oral disgust. Thus, low ratings for the South Asian female facial expression of disgust are not likely to be explained by differential facial expressions associated with moral vs. gustatory disgust across cultures (as further evidenced by the fact that the South Asian male facial expression of disgust was validated for racial representativeness among raters living in South Asia).

Furthermore, while facial expressions of disgust were validated for racial representativeness across all other core stimuli, they were consistently rated lower than the other facial expressions combined, as were facial expressions of anger. I turn to an evolutionary premise to explain this divergence: humans are adept at detecting threatening stimuli due to the survival advantage that such a skill confers (e.g., Mathews & Mackintosh, 1998; Öhman & Mineka, 2001). As such, threatening facial expressions of emotion—particularly expressions of anger—are widely known to capture involuntary attention (e.g., Fox, Russo, & Dutton, 2002; Sawada & Sato, 2015; Vuilleumier & Schwartz, 2001; for a review, see Yiend, 2010). Similarly, stimuli characterized by disgust interfere with Stroop color-naming tasks (Charash & McKay, 2002) and digit categorization tasks (Carretié, Ruiz-Padial, López-Martín, & Albert, 2011), and facial expressions of disgust prompt automatic encoding followed by top-down attentional modulation (Zhang, Liu, Zhou, Chen, & Luo, 2014). Collectively, these findings suggest that stimuli evoking anger and disgust capture exogenous attention due to the salience of threat. Perhaps, then, in the context of this study, the attention-capturing properties of angry and disgusted facial features distracted from the task of racial categorization (which relies on both skin color and physiognomy; Dunham, Dotsch, Clark, & Stepanova, 2016; Dunham, Stepanova, Dotsch, & Todorov, 2015), rendering raters comparatively less confident in their evaluations of race.

Finally, it is worth noting that prior research suggests a cognitive association between anger and Black faces, such that racially ambiguous individuals displaying facial expressions of anger are more likely to be categorized as Black compared to the same individuals expressing happiness (Dunham, 2011). However, these data do not reflect a similar bias. Black facial expressions of anger were rated lower on the racial validation scale ($M = 5.78$, $SD = 1.17$) than Black facial expressions of happiness ($M = 6.01$, $SD = 1.09$) [$t(213) = -4.06$, $p < 0.001$, $d = 0.29$]. Furthermore, when isolated to the predominantly White sample of participants in North America—98% of whom reported living in the United States and who therefore likely exhibit pronounced racial biases toward Black individuals (whether implicit or explicit) due to a history of racial conflict (for a review, see Amodio & Mendoza, 2010; Banks, Eberhardt, & Ross, 2006)—no difference in ratings was observed for Black facial expressions of anger ($M = 5.75$, $SD = 1.21$) vs. Black facial expressions of happiness ($M = 5.88$, $SD = 1.21$) [$t(55) = -1.25$, $p = 0.22$].

In sum, I speculate that the comparatively lower racial validation ratings for angry and disgusted facial expressions is likely due to attentional capture associated with threatening stimuli, which may have interfered with the validation process by reducing cognitive resources allocated to the task—regardless of the racial category of the target stimulus. Nonetheless, all of the stimulus images, except for the South Asian female facial expression of disgust, were ultimately validated for racial representativeness. Further research is needed to determine how facial expressions of disgust interact with race and gender cues for South Asian faces. In the meantime, I advise that researchers who choose to utilize the South Asian facial female expression of disgust do so with caution.

Expansion of Stimuli

The largely successful validation of racial representativeness across the stimulus set gave me license to expand the stimuli to include a more comprehensive range of facial expression images. However, an important caveat follows: I neither expect nor encourage the use of these images as prototypes of the emotion categories they represent. Barrett et al. (2019) performed a recent review of evidence surrounding the production and perception of emotional expressions, and the findings reveal that commonly accepted portrayals of emotion (i.e., those based on facial musculature stereotypically associated with the six basic emotions) do not sufficiently capture the richness of ways in which humans communicate emotion. For instance, stereotypical facial expression stimuli do not sufficiently reflect cultural variation in the production and interpretation of emotional expressions, nor do they account for the robust effects of contextual cues on facial expression processing. Additionally, emotional inference following exposure to such images is subject to many-to-many mappings between facial musculature and specific categories of emotion, such that—for example—anger may be expressed in myriad ways that do not involve the typically portrayed scowl, while the scowl might imply a number of different emotional states (Barrett et al., 2019).

The conclusions of Barrett et al. (2019) have prompted me not to attempt to generate "accurate" depictions of facial expression categories (if such a thing exists), but to instead create a large body of stimuli that captures the ambiguity inherent to the facial expression of emotion (without imposing emotional labels onto the images themselves). As such, I have created numerous continua of facial expression stimuli that include both fine-grained emotional intensity and images morphed incrementally between pairs of basic anchor expressions.

First, I generated a continuum of emotional expression intensity at 5% intervals for each of the six basic emotions, resulting in 20 images of each facial expression for each of the eight core stimuli (plus the original neutral expression). Thus, a total of 968 images were created to cover a comprehensive range of emotional intensity across gender and race. To generate these continua, I utilized the Expression sliders in FaceGen to specify the desired percentage of emotional intensity. Critically, in contrast to many traditional morphing programs, FaceGen's application of the Facial Action Coding System (Ekman & Friesen, 1978) across the full range of emotional intensity ensures clarity of expression, such that the visual integrity of all facial features is retained. To illustrate, an abbreviated continuum of the White male facial expression of fear is displayed in Figure 45.

Figure 45. Abbreviated continuum of the White male facial expression of fear. All images were produced using FaceGen Modeller Core 3.14 (Singular Inversions, 2016) and Adobe Photoshop.

Next, given that facial expressions of emotion in the real world are often ambiguous—a phenomenon that has generated considerable scientific attention (e.g., Barrett et al., 2019; Hassin, Aviezer, & Bentin, 2013; Motley & Camden, 1988), I further expanded the stimulus set to include images that are morphed between two distinct facial expressions. Specifically, I generated a continuum of morphed expressions at 10% intervals for each of the fifteen possible basic emotion pairings, thus lending varying degrees of ambiguity to the stimuli. I applied all fifteen continua to each of the eight core stimuli (again, by utilizing the Expression sliders in FaceGen), resulting in 1080 images. An abbreviated continuum of the Black female morph between facial expressions of happiness and surprise is displayed in Figure 46.

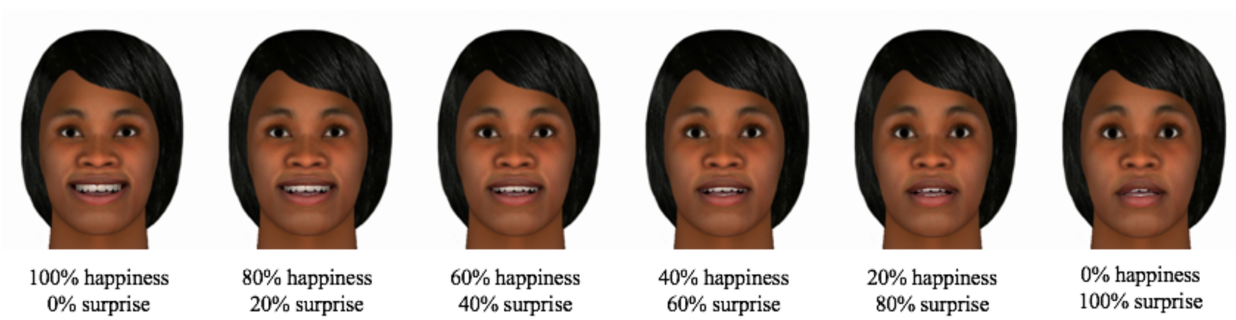

Figure 46. Abbreviated continuum of the Black female face morphed between happiness and surprise. All images were produced using FaceGen Modeller Core 3.14 (Singular Inversions, 2016) and Adobe Photoshop.

General Discussion

Through three distinct phases, creation of the International Faces database addressed obstacles that face perception researchers frequently encounter when faced with stimuli selection. First, by generating the stimuli through FaceGen, which consists of a digital "face space" accompanied by facial textures, I were able to achieve a reasonably high degree of human likeness across the stimuli—a quality that is often impoverished in computer-generated images of faces compared to photographs of real humans. Although, the issue of ecological validity surrounding computer-generated images may become less relevant with the growing sophistication of facial imaging software programs. Indeed, Schindler, Zell, Botsch, & Kissler (2017) found that neural activation in brain regions associated with higher-order perceptual evaluation and memory encoding increases linearly with face realism. Additionally, recent evidence surrounding amygdala activation in response to human vs. avatar emotional expressions suggests that the use of animated faces may be nearly as effective as human faces in the investigation of emotion perception, particularly for male participants (Cheetham, 2017).

Furthermore, creating computer-generated facial stimuli allowed for greater control over the potential confounding factor of luminance. I adjusted skin tone pigmentation to ensure equal luminosity within gender between Black and South Asian faces and between White and East Asian faces. Thus, perceptual biases observed for these racial pairings can more reasonably be attributed to true racial cues rather than skin tone per se.

Next, I validated the racial representativeness of the stimulus set by recruiting a geographically diverse sample of participants to rate each image. While most of the core stimuli were successfully validated, the South Asian female faces resulted in relatively low ratings. However, these ratings vastly improved when isolated to raters living in South Asia, who presumably exhibit greater perceptual expertise for South Asian faces than do raters living outside of South Asia. Ratings for the South Asian female facial expression of disgust remained low, however, warranting further investigation.

Finally, I expanded the stimulus set to include fine-grained continua of facial expression intensity, as well as continua of images morphed between basic facial expression pairings, resulting in 2048 unique stimuli. The use of FaceGen to produce these continua (as based on the Facial Action Coding System; Ekman & Friesen, 1978) ensured clarity of expression, thus combating the problem of visual distortion that can occur with basic morphing programs.

Lastly, it is important to note that, while racial representativeness of the stimulus set has largely been validated, the precision of the facial expressions displayed in the images has not. In the future, I intend to assess the *perceived* emotion across a subset of the stimuli by utilizing free-labeling tasks (in which participants generate their own words to describe the emotional state conveyed by a particular facial expression), given that the more commonly utilized choice-from-array tasks (in which participants must choose the correct answer from a small collection of options) may falsely amplify the reliability of basic facial expression images (Barrett et al., 2019). In the meantime, I encourage researchers to heed the advice put forth by Barrett et al. (2019), including instructions to manipulate the contexts in which facial expression images are embedded and to build variation into stimulus sets—ideally by sampling a large collection of stimuli for a given emotion category and treating stimuli as a random variable.

CONCLUSION

Across two studies, this dissertation first looks through the lens of the present to investigate how contemporary film is structured in order to optimize emotional engagement among viewers, and then looks forward in terms of working toward an application of these findings to Hollywood film as it diversifies. Thus, these studies begin to establish the emotional effects of visual film structure and to build a foundation for translating these findings to a broader context.

My foray into the influence of low-level film structure on cinematic engagement is not new. In fact, it was over a century ago that psychologist Hugo Münsterberg (1916/1970) declared that the visual structure of film mimics how the mind is engineered to attend to stimuli in the external environment. Today, a body of empirical evidence exists in retrospective support of Münsterberg's position, providing insight into the ways in which film capitalizes on human psychological processes to foster several aspects of narrative engagement, including attentional focus, narrative understanding, and now, emotional engagement.

Critically, this work provides a unique first glimpse into the dynamic nature of emotional response to whole films against a backdrop of cinematic visual structure. The findings demonstrate that emotional response among viewers is highly synchronized, and that this synchrony is partially driven by the low-level structure of film, suggesting that filmmakers hold implicit knowledge of the mechanisms underlying perceptual influences on emotional arousal. These practices do not seem to be applied consistently across films—but rather than suggesting a fickleness in the art of cinema, this inconsistency illustrates that filmmakers utilize the visual tools they have to cultivate a specific experience for viewers in the context of the stories they wish to tell. Ultimately, while viewers' emotional response to film is undoubtedly influenced by the narrative content of film itself, these results demonstrate that low-level film structure is also a crucial ingredient in the emotional experience that unites viewers as they witness cinematic narratives unfold on screen.

However, the applicability of these findings is limited—largely due to the fact that the stimuli utilized in this study feature White characters in Westernized storylines. Given robust evidence that face perception is strongly affected by ingroup vs. outgroup influences (such that own-race faces are individuated faster and more accurately than other-race faces, while other-race faces are more quickly categorized into a single group; Meissner & Brigham, 2001; Levin, 1996, 2000), it follows that such biases might affect emotional engagement with more diverse characters on screen. As such, the purpose of the second study was to create and validate a collection of racially diverse facial expression stimuli (dubbed the International Faces) that could serve as a canvas on which to investigate how the manipulation of low-level cinematic variables affects emotional engagement with more diverse characters. Thus, these stimuli overcome the lack of diverse representation in Hollywood film by creating a proxy for still images of characters who vary in terms of gender and race. While the use of still images precludes direct replication of Study 1 (given that this study relied on whole films to explore dynamic emotional tension over time), these stimuli nonetheless provide a strong starting point to expand on this research.

In addition, the International Faces database was created with broader research needs in mind, as the stimuli are intended for open use among face perception researchers. As such, the creation and validation of these stimuli addressed four issues central to the use of facial stimuli in psychological experimentation: human likeness, luminance control, racial representativeness across cultures, and clarity of expression.

In sum, extending the present work as described above will not only further illuminate how low-level film structure affects emotional engagement with diverse characters on screen, but it will also serve as a vehicle for hypothesis generation. By investigating the mechanisms underlying emotional engagement with a medium as ubiquitous and absorbing as film—particularly as film continues to better reflect the rich diversity of an increasingly globalized world—researchers are likely to formulate predictions that might not have otherwise been imagined. For instance, the findings of Study 1 led me to think differently about the effects of luminance on mood and to shift my focus from sound to silence when considering the impact of volume on emotional tension. Perhaps, then, follow-up studies on diversified stimuli will reveal new insight into the ways in which face perception across social groups interacts with the visual structure of film.

References

Adelmann & Zajonc. (1989). Facial efference and the experience of emotion. *Annual Review of Psychology*, *40*(1), 249–280. doi: 10.1146/annurev.ps.40.020189.001341

Albouy, P., Benjamin, L., Morillon, B., & Zatorre, R. J. (2020). Distinct sensitivity to spectrotemporal modulation supports brain asymmetry for speech and melody. *Science, 367*(6481), 1043-1047. doi: 10.1126/science.aaz3468

Aldredge, J. (2017). Understanding the importance of silence in filmmaking [Blog post]. *The Beat.* Retrieved from https://www.premiumbeat.com/blog/importance-silence-filmmaking-projects/

Alter, A. L., Stern, C., Granot, Y., & Balcetis, E. (2016). The "Bad Is Black" effect: Why people believe evildoers have darker skin than do-gooders. *Personality and Social Psychology Bulletin*, *42*(12), 1653–1665. doi: 10.1177/0146167216669123

Amodio, D. M., & Mendoza, S. A. (2010). Implicit intergroup bias: Cognitive, affective, and motivational underpinnings. In B. Gawronski, & B. K. Payne (Eds.), *Handbook of implicit social cognition measurement, theory, and applications* (pp. 353-374). New York: Guilford Press.

Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of General Psychology, 10*(3), 229–240. doi: 10.1037/1089-2680.10.3.229

Aristotle (1991). *On rhetoric: A theory of civic discourse* (G. A. Kennedy, Trans.) New York: Oxford University Press. (Original work published 350 BCE)

Aristotle (2013). *Poetics* (A. Kenny, Trans.) Oxford, UK: Oxford University Press. (Original work published 335 BCE)

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science, 338*(6111), 1225-1229. doi: 10.1126/science.1224313

Balazs, B. (2012). *Theory of the film.* Bristol, UK: Bristol Typesetting Company. (Original work published 1923)

Balters, S., & Steinert, M. J. (2017). Capturing emotion reactivity through physiology measurement as a foundation for affective engineering in engineering design science and engineering practices. *Journal of Intelligent Manufacturing, 28*(7), 1585-1607. doi: 10.1007/s10845-015-1145-2

Banks, R., Eberhardt, J., & Ross, L. (2006). Discrimination and implicit bias in a racially unequal society. *California Law Review, 94*(4), 1169-1190. doi:10.2307/20439061

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and nurture in own-race face processing. *Psychological Science*, *17*(2), 159–163. doi: 10.1111/j.1467-9280.2006.01679.x

Bard, P. (1934). On emotional experience after decortication with some remarks on theoretical views. *Psychological Review*, 41, 309-329. doi: 10.1037/h0070765

Barrett, L. F. (2017). *How emotions are made: The secret life of the brain.* Boston, MA: Houghton Mifflin Harcourt.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest, 20*(1), 1-68. doi: 10.1177/1529100619832930

Barsam, R. M., & Monahan, D. (2010). Looking at movies: An introduction to film. New York:

W. W. Norton & Co.

Beaudry, O., Roy-Charland, A., Perron, M., Cormier, I., & Tapp, R. (2014). Featural processing in recognition of emotional facial expressions. *Cognition and Emotion, 28*(3), 416-432. doi: 10.1080/02699931.2013.833500

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods, 190*(1), 80–91. doi: 10.1016/j.jneumeth.2010.04.028

Bergstrom, J. R., Duda, S., Hawkins, D., & McGill, M. (2014). Physiological response measurements. In J. R. Bergstrom & A. J. Schall (Eds.), *Eye tracking in user experience design* (pp. 81-108). New York: Elsevier.

Blau, J. J. C., Petrusz, S. C., & Carello, C. (2013). Fractal structure of event segmentation: Lessons from reel and real events. *Ecological Psychology*, *25*(1), 81–101. doi: 10.1080/10407413.2013.753811

Boltz, M., Schulkind, M., & Kantra, S. (1991). Effects of background music on the remembering of filmed events. *Memory & Cognition, 19*(6), 593-606. doi: 10.3758/BF03197154

Bordwell, D. (2002). Intensified continuity: Visual style in contemporary American film. *Film Quarterly, 55*(3), 16-28. doi: 10.1525/fq.2002.55.3.16

Bordwell, D. (2006). *The way Hollywood tells it.* Berkeley, CA: University of California Press.

Bordwell, D. (2012). The viewer's share: Models of mind in explaining film [Blog post]. *David Bordwell's Website on Cinema.* Retrieved from: http://www.davidbordwell.net/essays/viewersshare.php

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Publication recommendations for electrodermal measurement. *Psychophysiology, 49*, 1017-1034. doi: 10.1111/j.1469-8986.2012.01384.x

Brouwer, A-M., van Dam, E., van Erp, J. B. F., Spangler, D. P., & Brooks, J. R. (2018). Improving real-life estimates of emotion based on heart rate: A perspective on taking metabolic heart rate into account. *Frontiers in Human Neuroscience, 12*(284). doi: 10.3389/fnhum.2018.00284

Bruce, V., & Young, A. (2012). *Face perception*. New York: Psychology Press.

Brunick, K. L. & Cutting, J. E. (2014). Coloring the animated world: Exploring human color perception and preference through the animated film. In P. Taberham & T. Nanicelli (Eds.), *Cognitive media theory* (pp. 124-138). New York: Routledge.

Brunick, K. L., Cutting, J.E., & DeLong, J.E. (2013). Low-level features of film: What they are and why we would be lost without them. In A. P. Shimamura (Ed.), *Psychocinematics: Exploring cognition at the movies* (133-148). New York: Oxford University Press.

Brunsting, L. A., & Sheard, C. (1929). The color of the skin as analyzed by spectrophotometric methods. *Journal of Clinical Investigation*, *7*(4), 593–613. doi: 10.1172/JCI100245

Buck, R., Losow, J. I., Murphy, M. M., & Costanzo, P. (1992). Social facilitation and inhibition of emotional expression and communication. *Journal of Personality and Social Psychology, 63*(6), 962–968. doi: 10.1037/0022-3514.63.6.962

Busselle, R., & Bilandzic, H. (2009). Measuring narrative engagement. *Media Psychology*, *12*(4), 321-347. doi: 10.1080/15213260903287259

Byatt, G., & Rhodes, G. (2004). Identification of own-race and other-race faces: Implications for the representation of race in face space. *Psychonomic Bulletin & Review*, *11*(4), 735–741. doi: 10.3758/BF03196628

Calvo, M.G., Fernández-Martín, A. & Nummenmaa, L. (2014). Facial expression recognition in peripheral versus central vision: Role of the eyes and the mouth. *Psychological*

Research, 78, 180–195. doi: 10.1007/s00426-013-0492-x

Campbell, M. J., & Swinscow, T. D. V. (2009). *Statistics at square one* (11th ed.). Hoboken, NJ: Wiley-Blackwell.

Cannon, W. B. (1929). *Bodily changes in pain, hunger, fear, and rage: An account of recent research into the function of emotional excitement* (2nd ed.). New York: Appleton-Century-Crofts.

Carmi, R., & Itti, L. (2006). Visual causes versus correlates of attentional selection in dynamic scenes. *Vision Research*, *46*(26), 4333–4345. doi: 10.1016/j.visres.2006.08.019

Carretié, L., Ruiz-Padial, E., López-Martín, S., & Albert, J. (2011). Decomposing unpleasantness: Differential exogenous attention to disgusting and fearful stimuli. *Biological Psychology, 86,* 247–53. doi: 10.1016/j.biopsycho.2010.12.005

Carruthers, P., & Smith, P. K. (1996). *Theories of theories of mind.* Cambridge, UK: Cambridge University Press.

Chapman, H. A., Kim, D. A., Susskind, J. M., & Anderson, A. K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science*, *323*(5918), 1222–1226. doi: 10.1126/science.1165565

Charash, M., & McKay, D. (2002). Attention bias for disgust. *Journal of Anxiety Disorders, 16,* 529–41. doi: 10.1016/s0887-6185(02)00171-8

Cheetham, M. (2017). Editorial: The uncanny valley hypothesis and beyond. *Frontiers in Psychology, 8*, 1738. doi: 10.3389/fpsyg.2017.01738

Choi, K.-H., Kim, J., Kwon, O. S., Kim, M. J., Ryu, Y. H., & Park, J.-E. (2017). Is heart rate variability (HRV) an adequate tool for evaluating human emotions?—A focus on the use of the International Affective Picture System (IAPS). *Psychiatry Research, 251,* 192–

196. doi: 10.1016/j.psychres.2017.02.025

Coplan, A. (2006). Catching characters' emotions: Emotional contagion responses to narrative fiction film. *Film Studies*, *8*(1), 26–38. doi: 10.7227/fs.8.5

Craig, A. D. (2009a). Emotional moments across time: A possible neural basis for time perception in the anterior insula. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, *364*(1525), 1933–1942. doi:10.1098/rstb.2009.0008

Craig, A. D. (2009b). How do you feel—now? The anterior insula and human awareness. *Nature Reviews Neuroscience, 10*(1), 59-70. doi: 10.1038/nrn2555

Crivelli, C., Russell, J. A., Jarillo, S., & Fernández-Dols, J. M. (2016). The fear gasping face as a threat display in Melanesian society. *PNAS, 113*(44), 12403-12407. doi: 10.1073/pnas.1611622113

Cronin, M. (2008). *Translation goes to the movies*. New York: Routledge.

Crozier, W. R., & Jong, P. J. (2012). *The psychological significance of the blush*. New York: Cambridge University Press.

Cutting, J. E. (2014). How light and motion bathe the silver screen. *Psychology of Aesthetics, Creativity, and the Arts, 8*(3), 340-353. doi: 10.1037/a0036174

Cutting, J. E. (2016a). Narrative theory and the dynamics of popular movies. *Psychonomic Bulletin & Review*, *23*(6), 1713–1743. doi: 10.3758/s13423-016-1051-4

Cutting, J. E. (2016b). The evolution of pace in popular movies. *Cognitive Research: Principles and Implications, 1*(1), Article 30: 1-21. doi: 10.1186/S41235-016-0029-0.

Cutting, J. E., & Armstrong, K. L. (2016). Facial expression, size, and clutter: Inferences from movie structure to emotion judgments and back. *Attention, Perception, &*

Psychophysics, *78*(3), 891-90. doi: 10.3758/s13414-015-1003-5

Cutting, J. E., & Armstrong, K. L. (2018). Cryptic emotions and the emergence of a metatheory of mind in popular filmmaking. *Cognitive Science, 42,* 1317-1344. doi: 10.1111/cogs.12586

Cutting, J. E., Brunick, K. L., & Candan, A. (2012). Perceiving event dynamics and parsing Hollywood films. *Journal of Experimental Psychology. Human Perception and Performance*, *38*(6), 1476–1490. doi: 10.1037/a0027737

Cutting, J. E., Brunick, K. L., & DeLong, J. E. (2011). How act structure sculpts shot lengths and shot transitions in Hollywood film. *Projections: The Journal for Movies and Mind, 5,* 1-16. doi: 10.3167/proj.2011.050102

Cutting, J. E., Brunick, K. L., & DeLong, J. E. (2012). On shot lengths and film acts: A revised view. *Projections: The Journal for Movies and Mind, 6*, 142-145. doi: 10.3167/proj.2012.06010

Cutting, J. E., Brunick, K. L., DeLong , J. E., Iricinschi, C., & Candan, A. (2011). Quicker, faster, darker: Changes in Hollywood film over 75 years. *i-Perception*, 2, 569–576. doi: 10.1068/i0441aap

Cutting, J. E., DeLong, J. E., & Brunick, K. L. (2011). Visual activity in Hollywood film: 1935 to 2005 and beyond. *Psychology of Aesthetics, Creativity and the Arts*, 5, 115–125. doi: 10.1037/a0020995

Cutting, J. E., DeLong, J. E., & Nothelfer, C. E. (2010). Attention and the evolution of Hollywood film. *Psychological Science*, 21(3), 432–439. doi: 10.1177/0956797610361679

Cutting, J., & Iricinschi, C. (2015). Re-presentations of space in Hollywood movies: An eventindexing analysis. *Cognitive Science*, *39*(2), 434–456. doi: 10.1111/cogs.12151

Darwin, C. (2009). *The expression of the emotions in man and animals*. London: Penguin Classics. (Original work published 1872)

Davydov, D. M., Zech, E., & Luminet, O. (2011). Affective context of sadness and physiological response patterns. *Journal of Psychophysiology*, 25(2), 67–80. doi: 10.1027/0269-8803/a000031

Dean, R. T., Bailes, F., & Schubert, E. (2011). Acoustic intensity causes perceived changes in arousal levels in music: An experimental investigation. *PLOS ONE*, *6*(4), e18591. doi: 10.1371/journal.pone.0018591

DeLong, J.E., Brunick, K. L., & Cutting, J.E. (2013). Film through the human visual system: Finding patterns and limits. In J. C. Kaufman & D. K. Simonton (Eds.), *The social science of the cinema* (pp. 123-137). New York: Oxford University Press.

Desharnais, B., Camirand-Lemyre, F., Mireault, P., & Skinner, C. D. (2015). Determination of confidence intervals in non-normal data: Application of the bootstrap to cocaine concentration in femoral blood. *Journal of Analytical Toxicology*, *39*(2), 113–117. doi: 10.1093/jat/bku127

Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology*, *19*(6), 643–647. doi: 10.1111/j.1469-8986.1982.tb02516.x

Dimberg, U., & Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scandinavian Journal of Psychology*, *39*(1), 39–45. doi: 10.1111/1467-9450.00054

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, *11*(1), 86–89. doi: 10.1111/1467-9280.00221

Dunham, Y. (2011). An angry = outgroup effect. *Journal of Experimental Social Psychology, 47*, 668-671. doi: 10.1016/j.jesp.2011.01.003

Dunham, Y., Dotsch, R., Clark, A. R., & Stepanova, E. V. (2016). The development of White-Asian categorization: Contributions from skin color and other physiognomic cues. *PLOS ONE*, *11*(6), e0158211. doi: 10.1371/journal.pone.0158211

Dunham, Y., Stepanova, E. V., Dotsch, R., & Todorov, A. (2015). The development of race-based perceptual categorization: Skin color dominates early category judgments. *Developmental Science*, *18*(3), 469–483. doi: 10.1111/desc.12228

Ekman, P. (1971). Universals and cultural differences in facial expressions of emotion. *Nebraska Symposium on Motivation, 19,* 207–283.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face*. Englewood Cliffs, NJ: Prentice Hall.

Ellis, R. J., & Simons, R. F. (2005). The impact of music on subjective and physiological indices of emotion while viewing films. *Psychomusicology, 19*, 15-40. doi: 10.1037/h0094042

Favelle, S., Tobin, A., Piepers, D., Burke, D., & Robbins, R. A. (2015). Dynamic composite faces are processed holistically. *Vision Research, 111*, 26-32. doi: 10.1016/j.visres.2015.05.002

Ferrey, A. E., Burleigh, T. J., & Fenske, M. J. (2015). Stimulus-category competition, inhibition, and affective devaluation: A novel account of the uncanny valley. *Frontiers in Psychology*, *6*, 249. doi: 10.3389/fpsyg.2015.00249

Fields, H. L., & Margolis, E. B. (2015). Understanding opioid reward. *Trends in Neurosciences, 38*(4), 217-225. doi: 10.1016/j.tins.2015.01.002

Findlay, J. M. (2004). Eye scanning and visual search. In: Henderson J. M., & Ferreira F. (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 134-159). New York: Psychology Press.

Fitzmaurice, S., & Maibach, H. I. (2010). Gender differences in skin. In M. A. Farage, K. W. Miller, & H. I. Maibach (Eds.), *Textbook of aging skin* (pp. 999–1017). Heidelberg, Germany: Springer-Verlag Berlin Heidelberg.

Fox, E., Russo, R., & Dutton, K. (2002). Attentional bias for threat: Evidence for delayed disengagement from emotional faces. *Cognition & Emotion*, *16*(3), 355–379. doi:10.1080/02699930143000527

Frijda, N. H. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.

Frijda, N. H. (2007). *The laws of emotion.* New York: Routledge.

Gelman, S. A., & Hirschfeld, L. A. (1999). How biological is essentialism? In D. Medin & S. Atran (Eds.), *Folkbiology* (pp. 403–446). Cambridge, MA: MIT Press.

George, D., & Mallery, M. (2010). *SPSS for Windows step by step: A simple guide and reference*, 17.0 update (10th ed.). Boston: Pearson.

Gerrig, R. J. (1993). *Experiencing narrative worlds: On the psychological activities of reading*. New Haven, CT: Yale University Press.

Gil, S., & Droit-Volet, S. (2012). Emotional time distortions: The fundamental role of arousal. *Cognition & Emotion, 26*(5), 847–862. doi: 10.1080/02699931.2011.625401

Gilden, D. L. (2001). Cognitive emissions of 1/*f* noise. *Psychological Review*, *108*(1), 33–56. doi: 10.1037/0033-295x.108.1.33

Glass, L., Beuter, A., & Larocque, D. (1998). Time delays, oscillations, and chaos in physiological control systems. *Mathematical Biosciences, 90*(1-2), 111-125. doi: 10.1016/0025-5564(88)90060-0

Goldstein, E. B. (2013). *Sensation and perception* (9th ed.). Belmont, CA: Wadsworth Publishing.

Goldstein, R. B., Woods, R. L., & Peli, E. (2007). Where people look when watching movies: Do all viewers look at the same place? *Computers in Biology and Medicine*, *37*(7), 957–964. doi: 10.1016/j.compbiomed.2006.08.018

Gottschall, J. (2012). *The storytelling animal: How stories make us human.* New York: Houghton, Mifflin, Harcourt.

Green, M. C., & Brock, T. C. (2002). In the mind's eye: Transportation-imagery model of narrative persuasion. In M. C. Green, J. J. Strange, & T. C. Brock (Eds.), *Narrative impact: Social and cognitive foundations* (pp. 315–341). Mahwah, NJ: Erlbaum.

Green, M. C., & Sestir, M. (2017). Transportation theory. In *The International Encyclopedia of Media Effects*. John Wiley & Sons, Inc. doi: 10.1002/9781118783764.wbieme0083

Grodal, T. (2009). *Embodied visions: Evolution, emotion, culture, and film.* New York: Oxford University Press.

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology*, *39*(3), 281–291. doi: 10.1017/s0048577201393198

Gross, J. J., Fredrickson, B. L., & Levenson, R. W. (1994). The psychophysiology of crying. *Psychophysiology, 31*, 460–468. doi: 10.1111/j.1469-8986.1994.tb01049.x

Gross, J. J., & Levenson, R. W. (1993). Emotional suppression: Physiology, self-report, and expressive behavior. *Journal of Personality and Social Psychology, 64*(6), 970-986. doi: 10.1037//0022-3514.64.6.970

Hall, E. T. (1976). *Beyond culture*. New York: Doubleday.

Hall, M. B. (1987). *Color and technique in renaissance painting: Italy and the north*. New York: J. J. Augustin.

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories.

British Journal of Social Psychology, *39*(1), 113–127. doi: 10.1348/014466600164363

Hassin, R. R., Aviezer, H., & Bentin, S. (2013). Inherently ambiguous: Facial expressions of emotions, in context. *Emotion Review*, *5*(1), 60–65. doi: 10.1177/1754073912451331

Hasson, U., Furman, O., Clark, D., Dudai, Y., & Davachi, L. (2008). Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. *Neuron, 57*(3), 452-462. doi: 10.1016/j.neuron.2007.12.009

Hasson, U., Malach, R., & Heeger, D. J. (2010). Reliability of cortical activity during natural stimulation. *Trends in Cognitive Sciences, 14*(1), 40-48. doi: 10.1016/j.tics.2009.10.011

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science, 303*(5664), 1634-1640. doi: 10.1126/science.1089506

Harriott, M. (2017). Silence and why it's important in film [Blog post]. *The ECG Blog.* Retrieved from https://www.ecgprod.com/silence-and-why-its-important-in-film/

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). *Emotional contagion*. Cambridge, UK: Cambridge University Press.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223-233. doi: 10.1016/s1364-6613(00)01482-0

Hecht, S., & Shlaer, S. (1936). Intermittent stimulation by light: The relation between intensity and critical frequency for different parts of the spectrum. *The Journal of General Physiology, 19*(6), 965-977. doi: 10.1085/jgp.19.6.965

Henderson, J. M., Chanceaux, M., & Smith, T. J. (2009). The influence of clutter on real-world scene search: Evidence from search efficiency and eye movements. *Journal of Vision, 9*(1), Article 32, 1–8. doi:10. 1167/9.1.32

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, *21*(7), 1229–1243. doi: 10.1162/jocn.2009.21189

Hill, H., Bruce, V., & Akamatsu, S. (1995). Perceiving the sex and race of faces: The role of shape and colour. *Proceedings of the Royal Society B: Biological Sciences*, *261*(1362), 367–373. doi: 10.1098/rspb.1995.0161

Hochberg, J., & Brooks, V. (1978). Film cutting and visual momentum. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and the higher psychological functions* (pp. 293-317). Hillsdale, NJ: Lawrence Erlbaum.

Hogan, P.C. (2007). Sensorimotor projection, violations of continuity, and emotion in the experience of film. *Projections: The Journal for Movies and Mind*, 1(1), 41-58. doi: 10.3167/proj.2007.010104

Holland, N. (2003). The willing suspension of disbelief: A neuro-psychoanalytic view. *PSYART: A Hyperlink Journal for the Psychological Study of the Arts.* http://psyartjournal.com/article/show/n_holland-the_willing_suspension_of_disbelief_a_ne. Accessed February 1, 2017.

Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in social categorization: The role of prejudice and facial affect in race categorization. *Psychological Science*, *15*(5), 342–345. doi: 10.1111/j.0956-7976.2004.00680.x

Hunt, D., & Ramón, A-C. (2020). Hollywood Diversity Report 2020 [PDF file]. Retrieved from https://socialsciences.ucla.edu/wp-content/uploads/2020/02/UCLA-

Hollywood-Diversity-Report-2020-Film-2-6-2020.pdf

Hunt, D., Ramón, A-C., & Tran, M. (2019). Hollywood Diversity Report 2019 [PDF file]. Retrieved from https://socialsciences.ucla.edu/wp-content/uploads/2019/02/UCLA-Hollywood-Diversity-Report-2019-2-21-2019.pdf

Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual Review of Psychology*, *60*(1), 653–670. doi: 10.1146/annurev.psych.60.110707.163604

India's religions by numbers. (2015, August 26). *The Hindu*. Retrieved from https://www.thehindu.com/news/national/religious-communities-census-2011-what-the-numbers-say/article7582284.ece.

Itti, L. (2005). Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. *Visual Cognition*, *12*(6), 1093–1123. doi: 10.1080/13506280444000661

Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations. *Psychological Review, 99*(3), 561-565. doi: 10.1037/0033-295X.99.3.561

Izard, C. E., Huebner, R. R., Risser, D., & Dougherty, L. M. (1980). The young infant's ability to produce discrete emotion expressions. *Developmental Psychology, 16*(2), 132-140. doi: 10.1037/0012-1649.16.2.132

James, W. (1884). What is an emotion? *Mind*, 9, 188-205. doi: 10.1093/mind/os-IX.34.188

Jenkins, R., Dowsett, A. J., & Burton, A. M. (2018). How many faces do people know? *Proceedings of the Royal Society B: Biological Sciences, 285*(1888), 20181319. doi: 10.1098/rspb.2018.1319

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302-4311. doi: 10.1523/jneurosci.17-11-04302.1997

Kaplan, J. T., & Iacoboni, M. (2006). Getting a grip on other minds: Mirror neurons, intention understanding, and cognitive empathy. *Social Neuroscience*, *1*(3–4), 175–183. doi: 10.1080/17470910600985605

Kelley, E., Andrick, G., Benzenbower, F., & Devia, M. (2014). Physiological arousal response to differing musical genres. *Modern Psychological Studies, 20*(1), 25-33.

Keltner, D, Oatley, K., & Jenkins, J. M. (2013). *Understanding emotions* (3rd ed.). Hoboken, NJ: John Wiley & Sons.

Keysers, C., & Gazzola, V. (2010). Social neuroscience: Mirror neurons recorded in humans. *Current Biology*, *20*(8), R354. doi: 10.1016/j.cub.2010.03.013

Kilner, J. M., & Lemon, R. N. (2013). What we know currently about mirror neurons. *Current Biology*, *23*(23), R1057–R1062. doi: 10.1016/j.cub.2013.10.051

Kim H. Y. (2013). Statistical notes for clinical researchers: Assessing normal distributions using skewness and kurtosis. *Restorative Dentistry & Endodontics*, *38*(1), 52–54. doi: 10.5395/rde.2013.38.1.52

Kragel, P. A., & LaBar, K. S. (2016). Somatosensory representations link the perception of emotional expressions and sensory experience. *eNeuro*, *3*(2). doi: 10.1523/eneuro.0090-15.2016

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology, 84*(3), 394–421. doi: 10.1016/j.biopsycho.2010.03.010

Kreibig, S. D., Wilhelm, F. H., Roth, W. T., & Gross, J. J. (2007). Cardiovascular, electrodermal, and respiratory response patterns to fear- and sadness-inducing films. *Psychophysiology, 44*(5), 787–806. doi: 10.1111/j.1469-8986.2007.00550.x

Krosch, A. R., & Amodio, D. M. (2014). Economic scarcity alters the perception of race.

Proceedings of the National Academy of Sciences, *111*(25), 9079–9084. doi: 10.1073/pnas.1404448111

Kuleshov, L. (1974). *Kuleshov on film*. R. Levaco (Ed.). Berkeley: University of California Press.

Kulezic-Wilson, D. (2009). The music of film silence. *Music and the Moving Image, 2*(3), 1-10. doi: 10.5406/musimoviimag.2.3.0001

Lamm, C., & Majdandžić, J. (2015). The role of shared neural activations, mirror neurons, and morality in empathy – A critical comment. *Neuroscience Research*, *90*, 15–24. doi: 0.1016/j.neures.2014.10.008

Lange, C. (1887). *Ueber gemüthsbewegungen*. Leipzig, Germany: Theodor Thomas.

Lecker, M., Dotsch, R., Bijlstra, G., & Aviezer, H. (2019). Bidirectional contextual influence between faces and bodies in emotion perception. *Emotion.* Advance online publication. doi: 10.1037/emo0000619

Lee, V., Wagner, H. (2002). The effect of social presence on the facial and verbal expression of emotion and the interrelationships among emotion components. *Journal of Nonverbal Behavior, 26*, 3–25. doi: 10.1023/A:1014479919684

Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1364–1382. doi: 10.1037/0278-7393.22.6.1364

Levin, D. T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology*, *129*(4), 559–574. doi: 10.1037/0096-3445.129.4.559

Li, Y., & Tse, C.-S. (2016). Interference among the processing of facial emotion, face race, and face gender. *Frontiers in Psychology*, *7*, 1700. doi: 10.3389/fpsyg.2016.01700

Likowski, K. U., Mühlberger, A., Gerdes, A. B. M., Wieser, M. J., Pauli, P., & Weyers, P. (2012). Facial mimicry and the mirror neuron system: Simultaneous acquisition of facial electromyography and functional magnetic resonance imaging. *Frontiers in Human Neuroscience*, *6*, 214. doi: 10.3389/fnhum.2012.00214

Lisetti, C. L., & Nasoz, F. (2004). Using noninvasive wearable computers to recognize human emotions from physiological signals. *Journal on Applied Signal Processing, 11,* 1672-1687. doi: 10.1155/S1110865704406192

Loschky, L. C., Larson, A. M., Magliano, J. P., & Smith, T. J. (2015). What would Jaws do? The tyranny of film and the relationship between gaze and higher-level narrative film comprehension. *PLOS ONE*, *10*(11), e0142474. doi: 10.1371/journal.pone.0142474

Lositsky, O., Chen, J., Toker, D., Honey, C. J., Shvartsman, M., Poppenk, J. L., … Norman, K. A. (2016). Neural pattern change during encoding of a narrative predicts retrospective duration estimates. *eLife*, *5*, e16070. doi: 10.7554/eLife.16070

Magliano, J. P., Dijkstra, K., & Zwaan, R. A. (1996). Generating predictive inferences while viewing a movie. *Discourse Processes*, 22(3), 199–224. doi: 10.1080/01638539609544973

Magliano, J. P., & Zacks, J. M. (2011). The impact of continuity editing in narrative film on event segmentation. *Cognitive Science*, *35*(8), 1489–1517. doi: 10.1111/j.1551-6709.2011.01202.x

Mannan, S. K., Ruddock, K. H., & Wooding, D. S. (1997). Fixation sequences made during visual examination of briefly presented 2D images. *Spatial Vision*, *11*(2), 157–178. doi: 10.1163/156856897x00177

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Marsh, P., Beauchaine, T. P., & Williams, B. (2008). Dissociation of sad facial expressions and autonomic nervous system responding in boys with disruptive behavior disorders. *Psychophysiology, 45*(1), 100–110. doi: 10.1111/j.1469-8986.2007.00603.x

Mathews, A., & Mackintosh, B. (1998). A cognitive model of selective processing in anxiety. *Cognitive Therapy and Research, 22*(6), 539–560. doi: 10.1023/A:1018738019346

McEnany, G. W., & Lee, K. A. (2005). Effects of light therapy on sleep, mood, and temperature in women with nonseasonal major depression. *Issues in Mental Health Nursing, 26*(7), 781-794. doi: 10.1080/01612840591008410

McGrail, L. (n. d.). 5 essential elements of successful mise en scène in film [Blog post]. *Lights Film School.* Retrieved from https://www.lightsfilmschool.com/blog/mise-en-scene-in-film-afk.

McIntosh, D. (1996). Facial feedback hypothesis: Evidence, implications, and directions. *Motivation and Emotion*, *20*(2), 121–147. doi: 10.1007/bf02253868

McIntosh, D. (2006). Spontaneous facial mimicry, liking and emotional contagion. *Polish Psychological Bulletin*, *37*(1), 31–42.

Mehrabian, A. (1995). Framework for a comprehensive description and measurement of emotional states. *Genetic, Social, and General Psychology Monographs, 121*(3), 339-361.

Meier, B. P., Robinson, M. D., & Clore, G. L. (2004). Why good guys wear white: Automatic inferences about stimulus valence based on brightness. *Psychological Science*, *15*(2), 82–87. doi: 10.1111/j.0963-7214.2004.01502002.x

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, & Law, 7*(1), 3-35. doi: 10.1037/1076-8971.7.1.3

Mendelowitz, D. (1999). Advances in parasympathetic control of heart rate and cardiac function. *News in Physiological Sciences, 14*, 155-161. doi: 10.1152/physiologyonline.1999.14.4.155

Mital, P. K., Smith, T. J., Hill, R. L., & Henderson, J. M. (2011). Clustering of gaze during dynamic scene viewing is predicted by motion. *Cognitive Computation*, 3(1), 5-24. doi: 10.1007/s12559-010-9074-z

Mojzisch, A., Schilbach, L., Helmert, J. R., Pannasch, S., Velichkovsky, B. M., & Vogeley, K. (2006). The effects of self-involvement on attention, arousal, and facial expression during social interaction with virtual others: A psychophysiological study. *Social Neuroscience*, *1*(3–4), 184–195. doi: 10.1080/17470910600985621

Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2009). Is the mirror neuron system involved in imitation? A short review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *33*(7), 975–980. doi: 10.1016/j.neubiorev.2009.03.010

Moody, E. J., McIntosh, D. N., Mann, L. J., & Weisser, K. R. (2007). More than mere mimicry? The influence of emotion on rapid facial reactions to faces. *Emotion, 7*(2), 447–457. doi: 10.1037/1528-3542.7.2.447

Motion Picture Association of America (2019). 2018 THEME report [PDF file]. Retrieved from https://www.mpaa.org/wp-content/uploads/2019/03/MPAA-THEME-Report-2018.pdf.

Motley, M. T., & Camden, C. T. (1988). Facial expression of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. *Western Journal of Speech Communication*, *52*(1), 1–22. doi:

10.1080/10570318809389622

Münsterberg, H. (1970). *The film: A psychological study: The silent photoplay in 1916*. New York, NY: D. Appleton and Company. (Original work published 1916)

Neumann, D. L., & Westbury, H. R. (2011). The psychophysiological measurement of empathy. In D. J. Scapaletti (Ed.), *Psychology of empathy* (pp. 119-142). New York: Nova Science Publishers.

Newman, M. E. J. (2013). Power laws, Pareto distributions and Zipf's law. *Cities*, *30*, 59–67. doi: 10.1080/00107510500052444

Niedenthal, P. M., Halberstadt, J. B., Margolin, J., & Innes-Ker, Å. H. (2000). Emotional state and the detection of change in facial expression of emotion. *European Journal of Social Psychology, 30(2)*, 211-222. doi: 10.1002/(SICI)1099-0992(200003/04)30:2<211::AID-EJSP988>3.0.CO;2-3

Niedenthal, P. M., Mermillod, M., Maringer, M., & Hess, U. (2010). The Simulation of Smiles (SIMS) model: Embodied simulation and the meaning of facial expression. *Behavioral and Brain Sciences, 33*(6), 417-433. doi: 10.1017/s0140525x10000865

Oatley, K. (1999). Why fiction may be twice as true as fact: Fiction as cognitive and emotional simulation. *Review of General Psychology, 3*(2), 101–117. doi: 10.1037/1089-2680.3.2.101

Oatley, K. (2012). *The passionate muse: Exploring emotion in stories.* New York: Oxford University Press.

Oatley, K. (2013). How cues on the screen prompt emotions in the mind. In A. P. Shimamura (Ed.), *Psychocinematics: Exploring cognition at the movies* (pp. 269-284). New York: Oxford University Press.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108*(3), 483–522. doi: 0.1037/0033-295x.108.3.483

Packer, S. (2007). *Movies and the modern psyche.* London: Praeger.

Pandža, K. (2018). The effect of the handheld camera on narration and ocularization in the Dogme 95 film. *AM Journal of Art and Media Studies,* 15, 83-92. doi: 10.25038/am.v0i15.232

Paskow, A. (2008). *The paradoxes of art: A phenomenological investigation*. Cambridge, UK: Cambridge University Press.

Paterson, H.M., Pollick, F. E., & Sanford, A. J. (2001). The role of velocity in affect discrimination. *Proceedings of the Cognitive Science Society*, 23. Retrieved from: https://escholarship.org/uc/item/3191m9bh

Pek, J., Wong, A. C. M., & Wong, O. C. Y. (2017). Confidence intervals for the mean of non-normal distribution: Transform or not to transform. *Open Journal of Statistics*, *7*, 405. doi: 10.4236/ojs.2017.73029

Pinchasov, B. B., Shurgaja, A. M., Grischin, O. V., & Putilov, A. A. (2005). Mood and energy regulation in seasonal and non-seasonal depression before and after midday treatment with physical exercise or bright light. *Psychiatry Research, 94*(1), 29-42. doi: 10.1016/s0165-1781(00)00138-4

Plantinga, C. (2009). *Moving viewers: American film and the spectator's experience*. University of California Press.

Plutchik, R. (1991). *The emotions.* Lanham, MD: University Press of America.

Polking, K. (Ed.). (1990). *Writing A to Z: The terms, procedures, and facts of the writing* *business defined, explained, and put within reach.* Cincinnati, OH: Writer's Digest Press.

Pollatos, O., Laubrock, J., & Wittmann, M. (2014). Interoceptive focus shapes the experience of time. *PLOS ONE*, 9(1), e86934. doi: 10.1371/journal.pone.0086934

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology, 17*(3), 715-734. doi: 10.1017/S0954579405050340

Pronin, E. & Jacobs, E. (2008). Thought speed, mood and the experience of emotion. *Perspectives on Psychological Science, 3* (1), 461-85. doi: 10.1111/j.1745-6924.2008.00091.x

Pronin, E. & Wegner, D.M. (2006). Manic thinking: Independent effects of thought speed and thought content on mood. *Psychological Science, 17* (9), 807-13. doi: 10.1111/j.1467-9280.2006.01786.x

Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin*, *140*(5), 1281–1302. doi: 10.1037/a0037004

Richmond, L. L., Gold, D. A., & Zacks, J. M. (2017). Event perception: Translations and applications. *Journal of Applied Research in Memory and Cognition*, *6*(2), 111–120. doi: 10.1016/j.jarmac.2016.11.002

Rickard, N. S. (2004). Intense emotional responses to music: A test of the physiological arousal hypothesis. *Psychology of Music, 32*(4), 371-388. doi: 10.1177/0305735604046096

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*(1), 169-192. doi: 10.1146/annurev.neuro.27.070203.144230

Rooney, B., & Bálint, K. E. (2018). Watching more closely: Shot scale affects film viewers' theory of mind tendency but not ability. *Frontiers in Psychology, 8,* 2349. doi: 10.3389/fpsyg.2017.02349

Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision, 7*(2), 1–22. doi: 10.1167/7.2.17

Rozin, P. (1999). Food is fundamental, fun, frightening, and far-reaching. *Social Research*, *66*(1), 9–30.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178. doi: 10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 145-172. doi: 10.1037/0033-295x.110.1.145

Salt, B. (1992). *Film style and technology: History and analysis* (2nd ed.). London: Starwood.

Salt, B. (2006). *Moving into pictures*. London: Starwood.

Salt, B. (n.d.). Cinemetrics. Retrieved from http://www.cinemetrics.lv/satltdb.php.

Sawada, R., & Sato, W. (2015). Emotional attention capture by facial expressions. *Scientific Reports, 5,* 14042. doi: 10.1038/srep14042

Schacter, D. L., Gilbert, D. T., Wegner, D. M., & Nock, M. K. (2014). *Psychology* (3rd ed.). New York, NY: Worth Publishers.

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 69*(5), 379–399. doi: 10.1037/h0046234

Schaefer, A., Nils, F., Sanchez, X., & Philippot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition & Emotion, 24*(7), 1153-1172. doi: 10.1080/02699930903274322

Schindler, S., Zell, E., Botsch, M., & Kissler, J. (2017). Differential effects of face-realism and emotion on event-related brain potentials and their implications for the uncanny valley theory. *Scientific Reports*, *7*, 45003. doi: 10.1038/srep45003

Schmidt, A, & Thews, G. (1989). Autonomic nervous system. In W. Janig (Ed.), *Human Physiology* (2nd ed.), 333-370. New York: Springer-Verlag.

Shermer, M. (2008). Patternicity: Finding meaningful patterns in meaningless noise. *Scientific American.* Retrieved from https://www.scientificamerican.com/article/patternicity-finding-meaningful-patterns/

Shimamura, A. P., Cohn-Sheehy, B. I., & Shimamura, T. A. (2014). Perceiving movement across film edits: A psychocinematic analysis. *Psychology of Aesthetics, Creativity, and the Arts, 8*(1), 77–80. doi: 10.1037/a0034595

Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., & Yang, X. (2018). A review of emotion recognition using physiological signals. *Sensors, 18*(7), 2074. doi: 10.3390/s18072074

Sinkeviciute, V. (2017). “It's just a bit of cultural […] lost in translation”: Australian and British intracultural and intercultural metapragmatic evaluations of jocularity. *Lingua, 197*, 50-67. doi: 10.1016/j.lingua.2017.03.004

Smith, G. M. (2003). *Film structure and the emotion system*. Cambridge, UK: Cambridge University Press.

Smith, T. J. (2012). The attentional theory of cinematic continuity. *Projections*, *6*(1), 1–27. doi: 10.3167/proj.2012.060102

Smith, T. J. (2013). Watching you watch movies: Using eye tracking to inform cognitive film theory. In A. P. Shimamura (Ed.), *Psychocinematics: Exploring cognition at the movies*. (pp. 165-191). New York: Oxford University Press.

Smith, T., & Henderson, J. (2008). Attentional synchrony in static and dynamic scenes. *Journal* *of Vision*, *8*(6), 773–773. doi: 10.1167/8.6.773

Smith, T. J., Levin, D., & Cutting, J. E. (2012). A window on reality: Perceiving edited moving images. *Current Directions in Psychological Science*, *21*(2), 107–113. doi: 10.1177/0963721412437407

Snibbe, S. S., & Raffle, H. S. (2009). Social immersive media: Pursuing best practices for multi-user interactive camera/projector exhibits. *Proceedings of the 27th International Conference on Human Factors in Computing Systems*. doi: 10.1145/1518701.1518920

Speer, N. K., Swallow, K. M., & Zacks, J. M. (2003). Activation of human motion processing areas during event perception. *Cognitive, Affective & Behavioral Neuroscience*, *3*(4), 335–345. doi: 10.3758/cabn.3.4.335

Sternbach, R. A. (1962). Assessing differential autonomic patterns in emotions. *Journal of Psychosomatic Research, 6*(2), 87-91. doi: 10.1016/0022-3999(62)90059-4

Sternberg, R. J., Grigorenko, E. L., & Kidd, K. K. (2005). Intelligence, race, and genetics. *The American Psychologist*, *60*(1), 46–59. doi: 10.1037/0003-066X.60.1.46

Tan, E. S. (1996). *Emotion and the structure of narrative film: Film as an emotion machine*. Hillsdale, NJ: L. Erlbaum Associates.

Tan, E. S. (2018). A psychology of the film. *Palgrave Communications,* 4, Article 82. doi: 10.1057/s41599-018-0111-y

Tan, S-L., Spackman, M. P., & Bezdek, M. A. (2007). Viewers' interpretations of film characters' emotions: Effects of presenting film music before or after a character is shown. *Music Perception, 25*(2), 135-152. doi: 10.1525/mp.2007.25.2.135

Tanaka, J. W., Kaiser, M. D., Butler, S., & Le Grand, R. (2012). Mixed emotions: Holistic and analytic perception of facial expressions. *Cognition and Emotion, 26*(6), 961-977. doi:

10.1080/02699931.2011.630933

Tarvainen, J., Westman, S., & Oittinen, P. (2015). The way films feel: Aesthetic features and mood in film. *Psychology of Aesthetics, Creativity, and the Arts, 9*(3), 254-265. doi: 10.1037/a0039432

Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, *7*(14), 4.1-17. doi: 10.1167/7.14.4

Taylor, G. J., & Bagby, R. M. (2012). The alexithymia personality dimension. In T. A. Widiger (Ed.), *The Oxford handbook of personality disorders* (648-673). New York: Oxford University Press.

Thompson, K. (1999). *Storytelling in the new Hollywood*. Cambridge, MA: Harvard University Press.

Tishkoff, S. A., & Kidd, K. K. (2004). Implications of biogeography of human populations for “race” and medicine. *Nature Genetics*, *36*(11), S21-27. doi: 10.1038/ng1438

Trautmann, S. A., Fehr, T., & Herrmann, M. (2009). Emotions in motion: Dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. *Brain Research,* 1284, 100-115. doi: 10.1016/j.brainres.2009.05.075

Tseng, P.H., Carmi, R., Cameron, I. G. M., Munoz, D. P., & Itti, L. (2009). Quantifying center bias of observers in free viewing of dynamic natural scenes. *Journal of Vision*, *9*(7), 4. doi: 10.1167/9.7.4

Valdez, P., & Mehrabian, A. (1994). Effects of color on emotions. *Journal of Experimental Psychology, 123*(4), 394–409. doi:10.1037/0096-3445.123.4.394

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 43A*(2), 161–204. doi: 10.1080/14640749108400966

van der Velde, J., Servaas, M. N., Goerlich, K. S., Bruggeman, R., Horton, P., Costafreda, S. G., & Aleman, A. (2013). Neural correlates of alexithymia: A meta-analysis of emotion processing studies. *Neuroscience & Biobehavioral Reviews, 37*(8), 1774-1785. doi: 10.1016/j.neubiorev.2013.07.008

van Laer, T., de Ruyter, K., Visconti, L. M., & Wetzels, M. (2014). The extended transportation-imagery model: A meta-analysis of the antecedents and consequences of consumers' narrative transportation. *Journal of Consumer Research*, *40*(5), 797–817. doi: 10.2139/ssrn.2033192

Visch, V. T., Tan, E. S., & Molenaar, D. (2010). The emotional and cognitive effect of immersion in film viewing. *Cognition & Emotion, 24*(8), 1439-1445. doi: 10.1080/02699930903498186

Voelkle, M. C., Ebner, N. C., Lindenberger, U., & Riediger, M. (2014). A note on age differences in mood-congruent vs. mood-incongruent emotion processing in faces. *Frontiers in Emotion Science, 5*, 635. doi: 10.3389/fpsyg.2014.00635

Vuilleumier, P., Schwartz S. (2001). Emotional facial expressions capture attention. *Neurology, 56*(2), 153-158. doi: 10.1212/wnl.56.2.153

Walker, E. (2010). Hearing the silences (as well as the music) in Michael Haneke's films. *Music and the Moving Image, 3*(3), 1-11. doi: 10.5406/musimoviimag.3.3.0015

Walton, K. L. (1993). *Mimesis as make-believe: On the foundations of the representational arts*. Cambridge, MA: Harvard University Press.

Wang, C. A., Baird, T., Huang, J., Coutinho, J. D., Brien, D. C., & Munoz, D. P. (2018). Arousal effects on pupil size, heart rate, and skin conductance in an emotional face task. *Frontiers in Neurology*, *9*, 1029. doi:10.3389/fneur.2018.01029

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*(2), 219-235. doi: 10.1037//0033-2909.98.2.219

Wertheimer, M. (1912). Experimentelle studien über das sehen von beuegung. *Zeitchrift für Psychologie, 61*, 161-265.

West, S. G., Finch, J. F., Curran, P. J. (1995) Structural equation models with non-normal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues and applications* (56-75). Newbery Park, CA: Sage Publications.

What Is Chiaroscuro in Film? Definition, Techniques, and Film Examples. (2019). *StudioBinder.* Retrieved from https://www.studiobinder.com/blog/what-is-chiaroscuro-definition-examples/

Wise, R. A., & Rompre, P. P. (1989). Brain dopamine and reward. *Annual Review of Psychology*, 40, 191-225. doi: 10.1146/annurev.ps.40.020189.001203

Wittmann, M. (2009). The inner experience of time. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1525), 1955–1967. doi: 10.1098/rstb.2009.0003

Wittmann, M., & Wassenhove, V. (2009). The experience of time: Neural mechanisms and the interplay of emotion, cognition and embodiment. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1525), 1809–1813. doi: 10.1098/rstb.2009.0025

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*(6), 495–501. doi: 10.1038/nrn1411

Woodworth, R. S., & Schlosberg, H. (1960). *Experimental psychology: Revised edition*. New

York: Henry Holt.

Wyman, B., Smith, S., Meyers, D., & Godfrey, M. (2011). Digital storytelling in museums: Observations and best practices. *The Museum Journal, 54*(4), 461-468. doi: 10.1111/j.2151-6952.2011.00110.x

Yavuz, T. T., Claassen, J., & Kleinberg, S. (2019). *Lagged correlations among physiological variables as indicators of consciousness in stroke patients.* Paper presented at the American Medical Informatics Association Annual Symposium, Washington, DC.

Yeshurun, Y., Swanson, S., Simony, E., Chen, J., Lazaridi, C., Honey, C. J., & Hasson, U. (2017). Same story, different story: The neural representation of interpretive frameworks. *Psychological Science*, *28*(3), 307–319. doi: 10.1177/0956797616682029

Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition and Emotion, 24*(1), 3–47. doi: 10.1080/02699930903205698

Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science*, *28*(6), 979–1008. doi: 10.1207/s15516709cog2806_5

Zacks, J. M., Speer, N. K., Swallow, K. M., & Maley, C. J. (2010). The brain's cutting-room floor: Segmentation of narrative cinema. *Frontiers in Human Neuroscience*, *4*, 168. doi: 10.3389/fnhum.2010.00168

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3–21. doi: 10.1037/0033-2909.127.1.3

Zhang, D., Liu, Y., Zhou, C., Chen, Y., & Luo, Y. (2014). Spatial attention effects of disgusted and fearful faces. *PLOS ONE*, *9*(7), e101608. doi: 10.1371/journal.pone.0101608

Filmography

Anderson, P. T. (Director). (2007). *There Will Be Blood*. United States: Paramount Pictures.

Campbell, M. (Director). (2006). *Casino Royale*. United Kingdom: Columbia Picutres.

Cianfrance, D. (Director). (2016). *The Light Between Oceans*. United States: Walt Disney.

Cleary, B. (Director). (2015). *Stutterer*. Ireland: Galway Film.

Coats, A., & Hamou-Lhadj, L. (Directors). (2015). *Borrowed Time*. United States: Quorum Films.

Coppola, F. F. (Director). (1972). *The Godfather*. United States: Paramount Pictures.

Cortés, R. (Director). (2010). *Buried*. Spain: Lionsgate.

Gill, M. (Director). (2013). *The Voorman Problem.* United States: Honlodge Productions.

Haneke, M. (Director). (2005). *Caché.* France: Les Films du Losange.

Haneke, M. (Director). (2007). *Funny Games.* United States: Celluloid Dreams.

Jarmusch, J. (Director). (2016). *Paterson*. Germany: K5 International.

Kahrs, J. (Director). (2012). *Paperman.* United States: Walt Disney.

Kuperberg, C., & Kuperberg, R. (Directors). (2005). *L'émotion Par La Musique*. France: LGM.

Leone, S. (Director). (1966). *The Good, the Bad and the Ugly*. Italy: Produzioni Europee Associate.

López-Gallego, G. (Director). (2011). *Apollo 18*. United States: Dimension Films.

Marshall, N. (Director). (2005). *The Descent.* United Kingdom: Lionsgate.

Myrick, D., & Sanchez, E. (Directors). (1999). *The Blair Witch Project*. United States: Artisan.

Appendix A: Plot Breakdown of the Four Short Films

Borrowed Time

SETUP (0:00-1:39)

Scene 1:

Shots 1-5: (Fade in) The sheriff stands a short distance away from a steep cliff, looking plaintively into the distance. He removes his hat (slow music begins 0:07).

Scene 2:

Shots 6-12: Cross-cutting to a flashback, the sheriff as a young boy rides next to his father (who is carrying a rifle on his back) in a horse-drawn carriage through the desert. The boy looks bored while his father handles a pocket watch. The father winks at the boy then hands him the watch, ruffles the boy's hair, and places his hat on the boy's head. The boy smiles.

Scene 3:

Shots 13-20: Back on the cliff, the sheriff, with a limp, walks slowly toward the cliff's edge. Walking over rubble and debris, he pauses upon an animal skeleton, looks out over the cliff, sighs, and continues walking.

Scene 4:

Shots 21-24: Cross-cutting to the flashback, the boy and his father are pursued by a gunman on horseback (sound amplifies 1:29). The gunman shoots. The boy shouts in surprise and ducks his head.

Scene 5:

Shots 25-26: Back on the cliff, the sheriff continues his slow meander toward the cliff's edge (music intensifies 1:32).

Scene 6:

Shots 27-29: Cross-cutting to the flashback, the father places the reins in the boy's hands (the inciting incident), then leans over the back of the carriage and shoots at their pursuer.

COMPLICATING ACTION (1:39-2:03)

Scene 7:

Shots 30-31: Back on the cliff, the sheriff quickens his pace as he gets closer to the cliff's edge.

Scene 8:

Shots 32-37: Cross-cutting to the flashback, the carriage hits a boulder, and the boy is thrown from the carriage and dragged alongside it as he holds onto the reins.

Scene 9:

Shots 38-39: Back on the cliff, the sheriff's journey toward the cliff's edge becomes more hurried.

Scene 10:

Shots 40-41: Cross-cutting to the flashback, the boy continues to be dragged alongside the carriage. The horse whinnies.

Scene 11:

Shots 42-44: Back on the cliff, the sheriff makes it to the cliff's edge and drops to his knees.

Scene 12:

Shots 45-48: Cross-cutting to the flashback, the carriage topples over. The father is flown from its interior. He rolls over a cliff (music and sound cease 2:00).

DEVELOPMENT (2:03-2:54)

Scene 12 (continued):

Shots 49-69: The boy recovers (music starts back up 2:34) and runs to the cliff's edge. He looks over to see his father clinging to the rockface. The father shouts for the boy to give him his hand. The boy attempts to pull his father back up, but he is too far away to properly grip his father's hand. The father slips from the boy's feeble grasp and once again clings to the rock face. A wide shot reveals that this cliff is the same one that the adult sheriff is walking toward in the parallel sequence.

Shots 70-72: The father hands the boy his rifle to use as leverage (the point of no return).

CLIMAX (2:54-5:30)

Scene 12 (continued):

Shots 73-93: As the father grips the barrel of the rifle, the boy attempts to pull him up by the stock. The father makes some progress, coming close to the cliff's edge, but then loses his grip. The boy quickly grabs his father's shirt while continuing to grasp the rifle. The boy's hand slips, and he pulls the trigger. A shot is fired. The pocket watch flies through the air and bounces on the rocky ground, landing some distance behind the boy. It is splattered with blood. (Music stops 3:16 and is replaced by the sound of the ticking pocket watch.)

Shots 94-95: The boy, also splattered with blood, gasps on all fours. Then, a zoom in to the pocket watch as it stops ticking.

Scene 13:

Shots 96-101: (Crossfade into extreme long shot) Back on the cliff, the sheriff slowly rises from his kneeling position and looks over the cliff's edge. He steps to the precipice and looks out into the distance. The wind is blowing. He sways and closes his eyes.

Shots 102-110: Something bright catches the sheriff's attention. He opens his eyes to see the pocket watch, gleaming in the sunset. He loses his balance and slips over the cliff's edge.

Shots 111-112: The sheriff gazes at the stop watch, then grasps the rock face and pulls himself back up over the edge.

Shots 113-122: The sheriff lifts the pocket watch and stares into its dusty face. He wipes the dust away to reveal an old photograph of himself and his father (music starts at 4:48). He stares at his father's face and sobs. He holds the pocket watch to his chest and takes a deep breath (ticking can be heard 5:24). Pan out over the sunset.

Paperman

SETUP (0:00-1:27)

Scene 1:

Shots 1-9: (Fade in) George stands on an elevated train platform (prolog with long take) and on the track behind him a train rushes by (music begins 0:15). Meg runs past George chasing a piece of paper, then returns to stand next to him. Another train rushes by, and the wind slaps another piece of paper into Meg's face (music fades out as train rushes past, 0:31). George peels the paper off Meg's face, she looks at him, then the paper, and laughs. Wondering why she laughs, George sees that her lipstick left a mark on the paper. He too laughs, but Meg has already gotten onto the second train.

Shots 10-14: Having sat down, Meg looks out the window at George (the inciting incident); he looks at her; she looks at him; and he looks at her as the train leaves the platform. He continues to hold the sheet with her lipstick (the metonym).

Scene 2:

Shots 15-22: George is at work, in an office up high in a city building, and his boss drops several reams of paper on his desk. George's window is open, and the metonym almost flies out. At the window, George sees that Meg is interviewing in the building across the way. The window in front of her is open.

COMPLICATING ACTION (1:27-3:05)

Scene 2 (continued):

Shots 23-29: (music begins 1:27) George tries to get Meg's attention, but his efforts are futile. His boss signals at him to get to work. George thinks about what to do.

Shots 30-76: From the stack of paper on his desk, George serially folds then flies several paper airplanes across the street, but still fails to get Meg's attention. His boss chastises him and pointedly closes the window. George continues to fold and fly paper airplanes until he runs out of paper (music stops 2:43). Reaching for paper that isn't there, George loudly knocks items off of his desk. His coworkers stare at him. Only the metonym remains.

Shots 77-84: George notices that Meg has stood up; her interview is over. He carefully folds the metonym to make one last paper airplane, but as he prepares to fly it out the window, a gust a wind sucks it out of his hands.

DEVELOPMENT (3:05-4:17)

Scene 2 (continued):

Shots 85-97: Meg leaves the interview. George's coworkers stare at him, and his boss opens his office door to catch George off task, again (music starts back up then builds 3:11). George catches sight of Meg walking out of the building onto the street below. The boss drops a new stack of paper on George's desk. As the boss walks away, George runs out of the office (the point of no return).

Scene 3:

Shots: 98-102: (music stops 3:38) George dangerously crosses through traffic to the building from which Meg exited, and sees the metonymic airplane perched on top of a mailbox (music starts 3:53). Dejected, he throws the airplane far into the air.

Shots 103-105: The airplane drifts down into a small alleyway, joining hundreds of the other paper airplanes (music stops 4:10). It lies still on the ground.

CLIMAX (4:17-6:08)

Action Sequence:

Shots 105-112: The airplanes, led by the metonym, start to dance autonomously (new theme music 4:17); they follow George as he walks past the entrance of the alleyway. They surround him and force him back across the street, pushing him through traffic.

Shots 113-124: The metonym flies away, sails past a newsstand, and lands in an arrangement of flowers where Meg stands. Meg notices the metonym just before it flies off again. She pirouettes and follows it (music intensifies 5:11).

Shot 125: Meanwhile, George gets pushed up the stairs to an elevated train platform.

Shot 126: Continuing to follow the metonym, Meg runs up the steps to another train platform.

Shot 127: George continues to get pushed toward the train platform.

Shot 128: Meg chases the metonym around the platform.

Shot 129: George is still pushed by the paper airplanes.

Shots 130-132: Train doors open, and Meg follows the metonym inside. The train begins to leave the platform as Meg continues to chase the metonym.

Shots 133-135: George is pushed onto a train as it leaves the platform. He sits on a bench and is covered with many airplanes. He tries to get up, but the airplanes force him back down. George is frustrated. A mother takes her child away from George's bench.

Shots 136-143: As Meg holds the metonym, both trains arrive at the station. Meg steps onto the platform (music slows 5:45). She beholds the metonym in her hands. Paper airplanes rush underfoot, and Meg looks to the side to see George. He breaks out of his cloak of airplaces, and they approach each other (blackout, music continues).

EPILOG (with credits to the side, fade in 6:06, fade out 6:15)

Shots 144-148: Still shots (with dissolves) show George and Meg sitting, talking, and laughing in a café, with the epochal metonym sitting in front of them.

Stutterer

SETUP (0:00-4:03)

Scene 1:

Shots 1-5: Greenwood places a call with his internet provider, then stutters through an unsuccessful conversation with a customer representative (conveyed through jump cuts). The representative hangs up on him (dial tone) (cut to black, soft music begins 0:24).

Shots 6-11: (Opening title over wide shot of an apartment building in the city) Greenwood is on hold with the internet company as he practices sign language in his apartment. He flips through a book and stares at the phone (music stops 1:03).

Shots 12-24: Greenwood is sitting in bed with his laptop, when a notification sounds. He engages in a conversation over social media with his internet love interest, Ellie. Ellie ends the conversation by telling Greenwood that she has a surprise for him the next day (the inciting incident). Greenwood is intrigued. They say good night (fade to black, music starts back up 2:11).

Scene 2:

Shots 25-28: The next day, Greenwood, a type designer, works alone in his studio.

Shots 29-34: While standing outside smoking a cigarette, Greenwood observes a teenage boy sitting across the street (music fades 2:36). A voiceover describes Greenwood's mental speculation regarding the adolescent's life. Greenwood's thoughts, unlike his speaking voice, are unhindered by his stutter.

Scene 3:

Shots 35-37: Greenwood walks down the sidewalk, creating a poem about music in his head. He arrives at his father's house.

Shots 38-49: Greenwood and his father play a board game together, and Greenwood recites the poem out loud to his father. He falters over the words. His father listens patiently until Greenwood is done, then tells Greenwood that he likes the poem.

COMPLICATING ACTION (4:04-7:03)

Scene 4:

Shots 50-58: (Wide shot of Greenwood's apartment building) Back on hold with his internet provider, Greenwood is working on typeface when he hears a notification on his laptop. He goes to his laptop to find that Ellie has sent him a message requesting to meet him in person for the first time (music begins softly and builds 4:26). Greenwood pauses, then slowly closes his laptop (fade to black).

Scene 5:

Shots 59-61: The next day, while chopping vegetables at his father's house, Greenwood overhears his father on the phone with his internet provider, chastising the company for failing to communicate with his son.

Shot 62: Greenwood and his father sit together on the couch, eating a meal and watching television.

Scene 6:

Shots 63-72: While riding the train, Greenwood observes an attractive woman sitting nearby. Again, a voiceover describes Greenwood's mental speculation surrounding the woman's life (music fades 6:01).

Scene 7:

Shots 73-78: Greenwood looks through photos of Ellie on social media, then begins to type out a message making excuses for why he can't meet her. He stops himself and deletes the message (fade to black).

Scene 8:

Shots 79-82: (Wide shot over the city) The next day, Greenwood is standing outside smoking a cigarette. A woman approaches him to ask for directions. He signs to her, signaling that he can't speak. After she walks away, a voiceover reveals that he fluently provides the directions is his head.

DEVELOPMENT (7:03-9:13)

Scene 9:

Shots 83-85: Back in his studio, Greenwood is distracted (music starts back up 7:04). He opens his laptop and sends Ellie an apologetic message requesting to meet with her if she is still available (the point of no return) (music fades 7:43).

Scene 10:

Shots 86-101: Greenwood is sitting at a bus stop checking his phone when his attention is caught by a man shouting angrily at his girlfriend. The man then gets physically aggressive. Greenwood confronts him, and the man turns to shout obscenities at Greenwood. Cut to black.

Scene 11:

Shots 102: Greenwood sits in his apartment with a bandaged nose (music starts back up 8:34).

Shot 103: Greenwood works distractedly in his studio.

Scene 12

Shots 104-106: Greenwood watches television with his father. He checks his phone to find that he has no new messages from Ellie.

Scene 13:

Shots 107-111: Back at his apartment, Greenwood sits in his bedroom, despondent. He checks his laptop for messages from Ellie. There are none. He then looks in the mirror, staring at his bandaged face while, in a voiceover, he mentally and disparagingly describes his own life—in the same manner that he mentally speculates about the lives of others (music fades 9:12).

CLIMAX (9:13-12:10)

Scene 14:

Shots 112-113: Greenwood lies in bed awake. He hears a notification from his laptop. He rushes to check his messages, and stares at the screen in disbelief (music starts back up 9:41).

Scene 15:

Shots 114-118: The next morning, while brushing his teeth and getting dressed in preparation to meet Ellie, in a voiceover Greenwood mentally rehearses what he will say to her.

Scene 16 (Greenwood embarks on a meandering and anxiety-ridden journey to the location where he has arranged to meet Ellie. From shots 119-142, a voiceover of Greenwood's inner thoughts continues, becoming increasingly jumbled. This sequence is marked by several jump cuts):

Shots: 119-121: Greenwood carries flowers as he walks down the sidewalk.

Shots 122-124: Greenwood waits at a bus stop, where he throws the flowers in a trash can.

Shots 125-128: On the bus, Greenwood removes his tie.

Shots 129-133: Greenwood purchases a book from a bookstore, gift wraps it, then carries it along with him.

Shots 134-142: The sun is setting as Greenwood walks through a park (music intensifies 10:30). He veers from the park, walks to a road, and looks across the street.

Shots 143-145: Ellie stands on the sidewalk, looking down the street (music softens 10:51). Greenwood stares at her. Ellie continues to look down the street, when a man sitting behind her stands and leans toward her. The man taps Ellie on the shoulder and asks her a question. She signs to him, and he apologizes and sits back down.

Shots 146-157: Greenwood continues to stare at Ellie, awestruck by her use of sign language. Ellie turns back toward the street and catches sight of Greenwood (music rises 11:26). Tenderly, they hold each other's gaze for several seconds. Ellie smiles, then waves. Greenwood waves back. Ellie signs to Greenwood, inviting him to cross the street. Greenwood smiles, then looks both ways (cut to black, music continues into credits).

The Voorman Problem

SETUP (0:00-2:58)

Scene 1:

Shots 1-5: (Theme music starts 0:00 and stops abruptly 0:26) Dr. Williams walks down the sidewalk. He arrives at the entrance of a prison and pushes the buzzer. He introduces himself and explains that he is there to see Governor Bentley. No one responds. Faint chanting can be heard through the speaker. A buzzer sounds and Dr. Williams enters.

Scene 2:

Shots 6-22: In Governor Bentley's office, the governor invites Dr. Williams to sit down, and then pours liquor into his coffee cup. Dr. Williams sits. Governor Bentley explains that he would like Dr. Williams to examine Voorman, a prisoner in isolation who believes that he is a god (the inciting indident). Governor Bentley further explains that the chanting Dr. Williams heard as he entered the prison was the "song of Voorman." Dr. Williams asks about Voorman's conviction. Governor Bentley tells Dr. Williams that the computer systems crashed last year (implying that Voorman's crime is unknown). When Dr. Williams asks Governor Bentley how he will know when to release the prisoners, Governor Bentley responds, "Release?"

COMPLICATING ACTION (2:58-5:53)

Scene 3:

Shots 23-42: Dr. Williams walks into a room in which Voorman is seated and restrained in a straitjacket. Dr. Williams settles himself at a table several feet from Voorman. Dr. Williams asks Voorman how long he has believed himself to be a god; Voorman proceeds to answer Dr. Williams's questions in riddles. He claims that he created the universe nine days ago, as well as the evidence suggesting that the universe is older than nine days. Voorman also claims to have created Dr. Williams and his memories. Voorman further claims to carry the responsibility of imagining every last atom or it all goes "poof." Then, Voorman proposes an experiment to verify his claim: he will make Belgium disappear.

DEVELOPMENT (5:53-7:35)

Scene 4:

Shots 43-48: At home, Dr. Williams is writing notes while his wife serves him wine and asks what is worrying him. Dr. Williams explains that, in their country, there is very little difference between prison and the asylum. He explains that a prisoner named Voorman has promised to make Belgium disappear by teatime. Dr. Williams's wife asks if Belgium is another prisoner. Dr. Williams clarifies that he means Belgium, the country. His wife asks if he is joking. Dr. Williams retorts that he doesn't joke about his work.

Shots 49-57: Dr. Williams walks over to a box and pulls out a book (the point of no return). He flips through it and stares at the page in front of him in disbelief. A close-up shot reveals that he is inspecting a map from which Belgium has indeed disappeared (and is replaced with "Walloon Lagoon.") He says, repeatedly, to himself, that "this cannot be."

CLIMAX (7:35-11:03)

Scene 5:

Shots 58-64: Back at the prison, Dr. Williams meets with Voorman, who greets Dr. Williams and asks if Belgium will appear in his session notes. Voorman asks if Dr. Williams believes he is a demon who should be exorcised. Dr. Williams asks Voorman if he thinks he should be. Voorman shrugs.

Shots 65-71: Dr. Williams asks Voorman what a god is doing restrained in a prison. Voorman asks, sarcastically, if Dr. Williams supposes that he should be playing golf instead. Dr. Williams then asks Voorman how he spends his time; Voorman replies that he seeks amusement in humans, and that is why he created humans with imaginations—so that they could cook up new ways to entertain him. He further claims that he finds war amusing, which he chalks up to boredom.

Shots 72-84: Voorman then baits Dr. Williams by claiming that the doctor's intentions were to label him as either a faker or a lunatic, but that he instead ended up proving Voorman's "omniscient divinity." Dr. Williams retorts that nothing of the sort has been proven. Voorman then tells Dr. Williams that he has decided to switch places with him: Dr. Williams can be the omniscient god while Voorman seduces Dr. Williams's wife. Dr. Williams responds by calling Voorman a sick man.

Shots 85-90: In an extreme close-up shot, Voorman begins whistling, and then is revealed to be in Dr. Williams's place, sitting in his chair and wearing his suit. Dr. Williams is in Voorman's chair, restrained by a straitjacket. Dr. Williams, confused, struggles to get his bearings. He fights violently against the straitjacket while Voorman stands to leave. Dr. Williams shouts for the guards. Voorman tells Dr. Williams that if he cannot discuss things like a rational adult, then he has no choice but to end the interview (the "Song of Voorman" fades in, mixed with the theme music, and rises 10:33).

Shots 91-95: Dr. Williams shouts louder. Voorman promises that Dr. Williams will "soon learn the ropes." As Voorman walks out the door with Dr. Williams's belongings, he instructs Dr. Williams to watch North Korea. After Voorman leaves, Dr. Williams shouts into the empty room, with increasing desperation, that *he* is the real Dr. Williams (cut to black).

Appendix B: Participant Demographics

Geographical Region: Sub-Saharan Africa (n = 57)

- Mean age: 32.53
- Gender: 32% Female; 68% Male
- Countries of residence:
  - Kenya (16)
  - Nigeria (15)
  - South Africa (14)
  - Namibia (3)
  - Ghana (2)
  - Zimbabwe (2)
  - Cameroon (1)
  - Ethiopia (1)
  - Ivory Coast (1)
  - Mauritius (1)
  - Rwanda (1)
- Racial self-identification of participants:
  - Black (38)
  - White (11)
  - South Asian (2)
  - Mixed Race (1)
  - Other (5)

Geographical Region: United States & Canada (n = 56)

- Mean age: 35.20
- Gender: 39% Female; 59% Male; 2% Other
- Countries of residence:
  - United States (55)
  - Canada (1)
- Racial self-identification of participants:
  - White (49)
  - Mixed Race (4)
  - East Asian (2)
  - Black (1)

Geographical Region: South Asia (n = 51)

- Mean age: 29.63
- Gender: 20% Female; 80% Male
- Countries of residence:
  - India (50)
  - Pakistan (1)
- Racial self-identification of participants:
  - South Asian (46)
  - White (2)
  - Black (1)

    - Southeast Asian (1)
    - Other (1)

Geographical Region: East Asia (n = 50)

- Mean age: 31.10
- Gender: 36% Female; 62% Male; 2% Unspecified
- Countries of residence:
    - Japan (33)
    - Vietnam (5)
    - China (4)
    - Hong Kong (3)
    - South Korea (3)
    - Mongolia (1)
    - Taiwan (1)
- Racial self-identification of participants:
    - East Asian (17)
    - White (13)
    - Southeast Asian (12)
    - South Asian (3)
    - Mixed Race (2)
    - Black (1)
    - Other (2)

Appendix C: Bootstrapped Mean Ratings for Core Stimuli

Core Stimuli	Facial Expression	Geographical Location of Raters				
		Africa	N. America	South Asia	East Asia	Aggregate
Black Male	Neutral	6.27 [5.93, 6.61]	5.89 [5.52, 6.25]	5.67 [5.35, 5.98]	6.32 [6.06, 6.58]	**6.09 [5.93, 6.25]**
	Anger	6.23 [5.93, 6.53]	5.83 [5.50, 6.16]	5.62 [5.28, 5.96]	5.84 [5.44, 6.24]	**5.94 [5.76, 6.11]**
	Disgust	5.98 [5.61, 6.35]	5.92 [5.54, 6.29]	5.44 [5.05, 5.82]	6.01 [5.68, 6.34]	**5.89 [5.71, 6.07]**
	Fear	6.33 [6.00, 6.65]	6.08 [5.79, 6.36]	5.71 [5.41, 6.00]	5.95 [5.58, 6.32]	**6.08 [5.92, 6.23]**
	Happiness	6.47 [6.19, 6.74]	5.90 [5.54, 6.25]	5.73 [5.45, 6.00]	6.16 [5.80, 6.52]	**6.13 [5.97, 6.28]**
	Sadness	6.26 [5.95, 6.56]	5.97 [5.63, 6.30]	5.55 [5.24, 5.86]	6.17 [5.82, 6.52]	**6.05 [5.88, 6.21]**
	Surprise	6.58 [6.39, 6.77]	5.91 [5.59, 6.23]	5.68 [5.33, 6.02]	6.06 [5.70, 6.42]	**6.11 [5.95, 6.26]**
	Aggregate	**6.37 [6.25, 6.48]**	**5.99 [5.86, 6.11]**	**5.66 [5.54, 5.78]**	**6.14 [6.01, 6.27]**	6.05 [5.99, 6.11]
Black Female	Neutral	6.26 [6.02, 6.49]	5.90 [5.59, 6.21]	5.42 [5.06, 5.77]	6.01 [5.62, 6.40]	**5.95 [5.79, 6.11]**
	Anger	5.82 [5.47, 6.16]	5.56 [5.18, 5.93]	5.11 [4.77, 5.45]	5.74 [5.34, 6.14]	**5.61 [5.42, 5.79]**
	Disgust	5.82 [5.46, 6.18]	5.89 [5.57, 6.21]	5.43 [5.08, 5.78]	6.17 [5.84, 6.50]	**5.87 [5.69, 6.04]**
	Fear	5.97 [5.67, 6.26]	5.76 [5.41. 6.11]	5.22 [4.84, 5.59]	5.54 [5.08, 6.00]	**5.68 [5.50, 5.86]**
	Happiness	6.12 [5.83, 6.40]	5.73 [5.38, 6.07]	5.41 [5.06, 5.75]	6.04 [5.72, 6.36]	**5.87 [5.70, 6.03]**
	Sadness	6.13 [5.83, 6.42]	5.82 [5.48, 6.16]	5.39 [5.02, 5.75]	6.13 [5.82, 6.44]	**5.92 [5.75, 6.09]**
	Surprise	6.15 [5.88, 6.42]	5.64 [5.29, 5.98]	5.39 [5.04, 5.73]	5.88 [5.45, 6.30]	**5.82 [5.64, 5.99]**
	Aggregate	**6.09 [5.98, 6.20]**	**5.80 [5.67, 5.93]**	**5.37 [5.23, 5.50]**	**6.00 [5.85, 6.14]**	5.83 [5.76, 5.89]
White Male	Neutral	6.08 [5.77, 6.39]	5.60 [5.27, 5.93]	5.21 [4.86, 5.55]	6.01 [5.68, 6.34]	**5.77 [5.60, 5.94]**
	Anger	5.52 [5.11, 5.93]	5.37 [5.04, 5.70]	4.85* [4.53, 5.16]	5.59 [5.18, 6.00]	**5.36 [5.17, 5.55]**
	Disgust	5.34 [4.93, 5.74]	5.19 [4.82, 5.55]	4.71* [4.33, 5.08]	5.07 [4.60, 5.54]	**5.12 [4.91, 5.32]**
	Fear	6.13 [5.84, 6.42]	5.33 [4.96, 5.70]	4.98* [4.65, 5.31]	5.88 [5.50, 6.26]	**5.63 [5.45, 5.81]**
	Happiness	6.14 [5.86, 6.42]	5.52 [5.16, 5.88]	5.06 [4.73, 5.39]	5.73 [5.38, 6.08]	**5.60 [5.43, 5.77]**
	Sadness	5.79 [5.44, 6.14]	5.49 [5.16, 5.82]	5.13 [4.82, 5.43]	5.79 [5.42, 6.16]	**5.60 [5.43, 5.77]**
	Surprise	5.91 [5.56, 6.26]	5.55 [5.21, 5.88]	5.05 [4.73, 5.37]	5.91 [5.58, 6.24]	**5.65 [5.48, 5.82]**
	Aggregate	**5.90 [5.77, 6.03]**	**5.47 [5.34, 5.60]**	**5.03 [4.90, 5.16]**	**5.77 [5.63, 5.91]**	5.56 [5.49, 5.62]
White Female	Neutral	6.13 [5.84, 6.42]	5.66 [5.30, 6.02]	5.47 [5.18, 5.75]	6.04 [5.70, 6.38]	**5.88 [5.71, 6.04]**
	Anger	5.69 [5.30, 6.07]	5.59 [5.25, 5.93]	4.86* [4.51, 5.20]	5.78 [5.40, 6.16]	**5.54 [5.35, 5.72]**
	Disgust	5.25 [4.81, 5.68]	5.48 [5.16, 5.80]	4.87* [4.51, 5.22]	5.30 [4.86, 5.74]	**5.27 [5.07, 5.46]**
	Fear	5.98 [5.68, 6.28]	5.63 [5.25, 6.00]	5.33 [5.00, 5.65]	5.79 [5.40, 6.18]	**5.74 [5.56, 5.91]**

	Happiness	5.97 [5.67, 6.26]	5.70 [5.36, 6.04]	5.44 [5.14, 5.73]	5.74 [5.38, 6.10]	**5.75 [5.58, 5.91]**
	Sadness	5.99 [5.65, 6.32]	5.55 [5.21, 5.88]	5.20 [4.86, 5.53]	6.09 [5.78, 6.40]	**5.75 [5.58, 5.91]**
	Surprise	6.08 [5.79, 6.37]	5.45 [5.05, 5.84]	5.19 [4.86, 5.51]	5.62 [5.22, 6.02]	**5.63 [5.45, 5.81]**
	Aggregate	**5.93 [5.80, 6.05]**	**5.62 [5.49, 5.75]**	**5.22 [5.10, 5.34]**	**5.83 [5.68, 5.97]**	**5.66 [5.59, 5.72]**
South Asian Male	**Neutral**	5.57 [5.19, 5.95]	4.74* [4.30, 5.18]	5.73 [5.45, 6.10]	5.49 [5.08, 5.90]	**5.43 [5.22, 5.63]**
	Anger	5.53 [5.18, 5.88]	4.53* [4.05, 5.00]	5.40 [5.04, 5.75]	5.01 [4.60, 5.42]	**5.16 [4.95, 5.36]**
	Disgust	5.11 [4.68, 5.81]	4.66* [4.23, 5.09]	5.15 [4.78, 5.51]	5.04 [4.62, 5.45]	**5.02 [4.81, 5.22]**
	Fear	5.48 [5.14, 5.81]	4.78* [4.34, 5.21]	5.60 [5.26, 5.94]	5.46 [5.06, 5.86]	**5.36 [5.16, 5.55]**
	Happiness	5.50 [5.12, 5.88]	4.66* [4.18, 5.13]	5.63 [5.31, 5.94]	5.52 [5.16, 5.88]	**5.28 [5.14, 5.42]**
	Sadness	5.61 [5.29, 5.93]	4.80* [4.34, 5.25]	5.71 [5.35, 6.06]	5.70 [5.38, 6.02]	**5.48 [5.29, 5.67]**
	Surprise	5.51 [5.11, 5.91]	4.65* [4.20, 5.09]	5.84 [5.55, 6.12]	5.68 [5.34, 6.02]	**5.44 [5.24, 5.63]**
	Aggregate	**5.51 [5.37, 5.65]**	**4.73* [4.56, 4.90]**	**5.63 [5.50, 5.75]**	**5.46 [5.32, 5.60]**	**5.33 [5.25, 5.40]**
South Asian Female	**Neutral**	4.44* [3.97, 4.90]	4.48* [4.07, 4.88]	5.58 [5.22, 5.94]	5.00 [4.56, 5.44]	**4.88* [4.66, 5.09]**
	Anger	4.44* [3.98, 4.90]	4.24* [3.84, 4.63]	5.03 [4.67, 5.39]	4.57* [4.20, 4.94]	**4.59* [4.38, 4.79]**
	Disgust	4.25* [3.81, 4.68]	4.31* [3.89, 4.73]	4.37* [3.96, 4.77]	4.30* [3.90, 4.70]	**4.32* [4.11, 4.53]**
	Fear	4.66* [4.21, 5.11]	4.04* [3.63, 4.45]	5.32 [4.98, 5.65]	4.69* [4.30, 5.08]	**4.68* [4.47, 4.89]**
	Happiness	4.73* [4.26, 5.19]	4.18* [3.77, 4.59]	5.33 [4.98, 5.67]	5.26 [4.94, 5.58]	**4.88* [4.67, 5.08]**
	Sadness	4.70* [4.28, 5.12]	3.99* [3.57, 4.41]	5.60 [5.26, 5.94]	4.77* [4.38, 5.16]	**4.78* [4.57, 4.99]**
	Surprise	4.69* [4.21, 5.16]	4.18* [3.73, 4.63]	5.29 [4.92, 5.65]	5.01 [4.70, 5.32]	**4.80* [4.58, 5.01]**
	Aggregate	**4.59* [4.41, 4.76]**	**4.22* [4.06, 4.38]**	**5.25 [5.11, 5.39]**	**4.84* [4.69, 4.98]**	**4.71* [4.63, 4.79]**
East Asian Male	**Neutral**	6.15 [5.88, 6.42]	5.76 [5.43, 6.09]	5.33 [4.96, 5.69]	6.12 [5.82, 6.42]	**5.89 [5.72, 6.05]**
	Anger	6.12 [5.87, 6.37]	5.91 [5.64, 6.18]	5.07 [4.69, 5.45]	5.64 [5.26, 6.02]	**5.74 [5.58, 5.90]**
	Disgust	5.94 [5.61, 6.26]	5.84 [5.50, 6.18]	5.07 [4.71, 5.43]	5.36 [4.94, 5.78]	**5.62 [5.43, 5.80]**
	Fear	6.06 [5.77, 6.35]	5.66 [5.30, 6.02]	5.23 [4.86, 5.59]	6.04 [5.78, 6.30]	**5.80 [5.63, 5.96]**
	Happiness	6.02 [5.68, 6.35]	5.66 [5.36, 5.96]	5.61 [5.29, 5.92]	6.10 [5.82, 6.38]	**5.89 [5.74, 6.04]**
	Sadness	6.19 [5.97, 6.40]	5.85 [5.55, 6.14]	5.25 [4.92, 5.57]	6.20 [5.92, 6.48]	**5.92 [5.77, 6.06]**
	Surprise	5.95 [5.63, 6.26]	5.76 [5.38, 6.13]	5.42 [5.08, 5.75]	5.75 [5.36, 6.14]	**5.78 [5.60, 5.95]**
	Aggregate	**6.12 [6.01, 6.22]**	**5.82 [5.69, 5.94]**	**5.32 [5.19, 5.45]**	**5.95 [5.82, 6.07]**	**5.82 [5.75, 5.88]**
East Asian Female	**Neutral**	5.14 [4.74, 5.54]	4.99* [4.61, 5.36]	5.13 [4.77, 5.49]	5.75 [5.42, 6.08]	**5.27 [5.08, 5.46]**
	Anger	5.26 [4.84, 5.68]	4.93* [4.55, 5.30]	4.90* [4.55, 5.24]	5.00 [4.52, 5.48]	**5.06 [4.85, 5.26]**

	Disgust	5.20 [4.83, 5.56]	4.99* [4.61, 5.36]	4.96* [4.63, 5.29]	4.94* [4.50, 5.38]	**5.06** **[4.86, 5.25]**
	Fear	4.98* [4.56, 5.40]	5.20 [4.88, 5.52]	5.06 [4.77, 5.35]	5.21 [4.78, 5.64]	**5.14** **[4.95, 5.32]**
	Happiness	4.91* [4.49, 5.32]	4.79* [4.41, 5.16]	5.08 [4.75, 5.41]	5.15 [4.70, 5.60]	**5.00** **[4.80, 5.20]**
	Sadness	5.26 [4.84, 5.67]	5.13 [4.75, 5.50]	5.16 [4.85, 5.47]	5.36 [4.94, 5.78]	**5.26** **[5.07, 5.45]**
	Surprise	5.02 [4.59, 5.44]	4.82* [4.43, 5.20]	5.01 [4.69, 5.33]	5.23 [4.82, 5.64]	**5.05** **[4.85, 5.24]**
	Aggregate	**5.16** **[5.01, 5.31]**	**5.00** **[4.86, 5.14]**	**5.08** **[4.95, 5.20]**	**5.27** **[5.11, 5.43]**	**5.13** **[5.06, 5.20]**