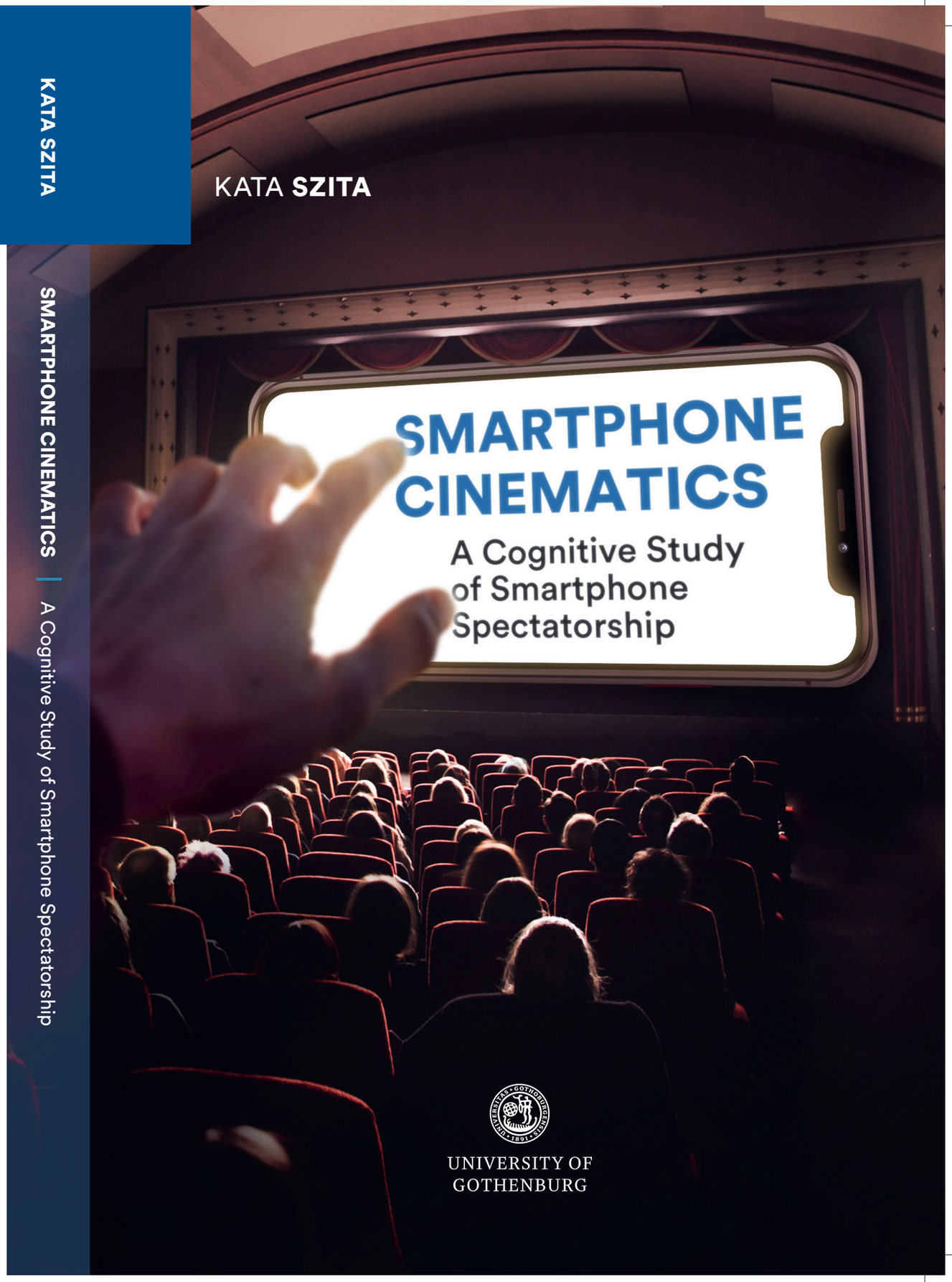
SMARTPHONE CINEMATICS

KATA SZITA

Smartphone spectatorship is not simply *à la mode* in contemporary media consumption; it signals a pervasive cultural and social shift that has followed upon the quickly expanding terrain of digital technology and digital industry. What secured the smartphone's legacy in terms of spectatorship is that it is always ready at hand or sitting in one's pocket, and grants access to movie and video content anytime, anywhere. But this nonstop access leaves its mark on consumption patterns, viewing duration, and the overall viewing experience. The smartphone's small size and the constant control over the screen position and content, though they facilitate spontaneity, are perhaps the most notable factors responsible for reducing consistent attention to the screened content and inducing attention shifts between a narration, interface, and the surrounding, often unenclosed, space.

Active bodily control places smartphones at the center of attention of a technologically and cognitively equally flavored research project that balances cinema-centered and tech-oriented views to approximate the interests of content and hardware providers to spectator psychology and a sustainable, ethical, and pleasurable moving-image consumption. By combining media-theoretical, cognitive, phenomenological, and behavioral approaches, this volume examines interactive watching, usability, changes in sensory perspectives, consumption, and participation. These areas of inquiry expand the limited insight into mobile media and moving images and illuminate the plasticity and personal quality of spectatorship.



Smartphone Cinematics

Smartphone Cinematics

A Cognitive Study of Smartphone Spectatorship

Kata Szita

UNIVERSITY OF GOTHENBURG

ISBN 978-91-85974-23-8

Doctoral dissertation in Film Studies, Department of Cultural Sciences, University of Gothenburg.

Cover design: Dániel Buzás

Printed by BrandFactory, Kållered, 2019

Abstract
Title: Smartphone Cinematics: A Cognitive Study of Smartphone Spectatorship
Author: Kata Szita
Language: English with Swedish summary
Department: Department of Cultural Sciences, University of Gothenburg
ISBN 978-91-85974-23-8

This thesis discusses the wide-spread, yet scarcely researched area of smartphone spectatorship and provides a theoretical, empirical, and methodological contribution to the study of digital media and moving-image culture. The peculiarity of smartphone spectatorship lies in its pervasiveness; the fact that it bears little institutional, social, temporal, or spatial constraints. By combining media-theoretical, cognitive, phenomenological, and behavioral approaches, this thesis examines interactive watching, usability, changes in sensory perspectives, film consumption and participation, as well as spectatorship in unenclosed spaces.

The first part of this thesis includes a theoretical overview of smartphone spectatorship explored through the lens of multisensory and interactive viewing. This overview begins with an investigation of the modes smartphones remediate the temporal, spatial, and cultural frameworks of earlier screening platforms in order to define smartphone-related practices which include the options and motivations for interactions. Interactive watching on smartphones pertains to adjustments performed by moving the screen to change the visual angle or by executing manipulations on screen content through touchscreen interaction. The need for interactions is attributed to both external distractions and internal motivations which effectuate the spectator's encounter with a personalized, yet often fragmented, narrative.

The second part uses a combined method: it contains a theoretical discussion of perception and narrative comprehension regarding smartphone spectatorship, followed by a behavioral experiment measuring the effects of screen size and unenclosed viewing spaces. The empirical methodology chosen involves a combination of eye tracking and electrodermal activity measures, subjective presence ratings, and a performance test measuring narrative comprehension. This experiment has two areas of investigation. First, it compares viewers' physiological responses, engagement, and comprehension related to smartphones versus large screen use in environments with and without external distractions. Second, it measures what attributes of distracting stimuli affect the likelihood of attention oscillations between a movie and the surrounding space.

Smartphone spectatorship is demonstratively based on spectators' medium- and platform-specific knowledge and their bodily and mental presence. These factors not only mobilize the cognitive processes for accessing and assembling a coherent narrative, they also inform the means, time, and degree of interactions. The overall findings of this dissertation present that the smartphone's small size can be compromised by screen positioning and momentary adjustments performed in favor of an uninterrupted viewing experience. Nevertheless, smartphone viewers remain susceptible to distraction depending on attributes of external stimuli.

KEYWORDS: smartphone, screen, spectatorship, viewer behavior, cognition, movie narration, eye tracking, electrodermal activity.

Acknowledgements

My sincere gratitude extends to all of those who made developing my ideas into this book possible. You all taught me a lot and gave great inspiration to my work.

I am indebted to my supervisors, Mats Björkin, Mats Jönsson, and Lisa Åkervall, for letting me embark on this wonderful journey and leading me during the past half a decade. I am not least grateful for the help of my colleagues at the Department of Cultural Sciences, including Ola Stockfelt, who gave insightful comments on sound, and Alf Björnberg and Karin Wagner, who have always made sure that doctoral students have all they need for going through one of their first big challenges as researchers. *Tack* goes to Britt Klintenberg for her help with Swedish texts. And *tack*, Gunilla Zachau, for ensuring not only a great working environment but also for providing incredible support during tough times.

Kiitos paljon, Pia Tikka and Lily Díaz-Kommonen, for inviting and welcoming me at Aalto Media Lab to carry out my experimental work in 2016–2017; and Kimmo Alho from the University of Helsinki and Tapio Takala from Aalto University for their suggestions about experimental methods. A special thanks goes to Veli-Matti Saarinen at Aalto Behavioral Laboratory for his work in preparing the not-always-uncomplicated experimental setup and for teaching me a great deal about the technical side of behavioral experiments. For the financial support for completing this experiment, I owe thanks to Holger och Thyra Lauritzens stiftelse and Stiftelsen Lars Hiertas Minne.

I want to extend my deepest thanks to Brendan Rooney for his incredible guidance in analyzing the experimental results and for inviting me to University College Dublin in 2018. And thank you, Åsa Wengelin, Kajsa Hansen Yang, Katalin Bálint, and Vera Szekér for giving me a thread in the labyrinth of statistics.

I am especially grateful to my final seminar opponent, Miklós Kiss, who provided reassurance and immense guidance before the last couple of months of my PhD and to all those amazing colleagues, who participated in the seminar.

A great deal of my work was fueled by input from colleagues who attended my talks at various conferences, courses, and departmental events from Melbourne to Ithaca, from Gorizia to Umeå. Thanks go to my doctoral colleagues, as well as the members of the Society for Cognitive Studies of the

Moving Image, the European Network for Cinema and Media Studies, the Swedish Cognitive Science Society, and many more.

This research would have never been possible without my wonderful friends in Hungary, Sweden, and Finland. From the University of Gothenburg, I am especially grateful to Valeria Villegas Lindvall for her positive attitude and never-ending support during all those years of our friendship; Netta Hibsher for her valuable comments on my text and for taking my mind off of work even when I didn't think it was necessary; Christine Sjöberg, for her always spot-on suggestions; Alla Rybina for always being there and making me laugh; and Ricky LaBontee for being an amazing language counsellor. In the end of the list from the University, my deepest gratitude goes to Olga Nikolaeva, who was always ready to brainstorm, but has been an even greater support as friend, mental counsellor, roommate, and many other things from the very first day.

I am grateful to the entire Servinkuja alumni, Annika Stuke, Heeryung Shin, Mareike Monschein, and Sanja Šćepanović, for supporting me all the way and for always keeping their doors open in Finland or wherever life takes us. Thank you all for your help with my experiment, and Heeryung, I will never forget your ceaseless support when I spent days and nights in the lab. Thanks go also to Marien Bouma for being an eager typo hunter; to Claire Howlin for her helpful suggestions; to Dani Buzás for the stunning cover design; as well as to the wonderful singers of WiOL, Wene-WiOL, and Chalmers Sångkör for providing an amazing soundtrack to these years.

Heartfelt thanks go to my family for always standing by my side, no matter what goals I pursue. Thank you for sharing your interest in films and your passion and knowledge. Anyu, Apu, Nagyi, Nóra, és Csongor, köszönök mindent!

Table of Contents

Figures

Tables

Introduction

A Trailer for Film Experience and Mobile Media

Viewers' bodies enliven moving images. They enlivened them when still images flickered on the spinning discs of Muybridge's zoopraxiscope or were observed inside a chamber through the pinhole of Edison's kinetoscope. They enlivened them when sequential images appeared on the walls of public screening rooms and viewers mistook the screened environment for their own. They also enlivened them when cinemas introduced synchronized sound and surround sound; and when not only sound, but even images started to surround viewers at 3D and IMAX screenings and subsequently in virtual reality chambers. Along the way, personal devices appeared and borrowed cinema's image- and sound-projecting capacities to relocate them into new personal and public environments. Now, screens are part of many aspects of everyday life: they occupy a wide range of spaces calling attention to their technological and cultural influence in shaping storytelling, production, consumption, and viewing experiences.

As moving images, screening media, and narrative strategies have evolved, interest has likewise risen in practical, economic, and scholarly terms to introduce, promote, and study these new trajectories: attention was directed to home entertainment (e.g., television, home video), extra-cinematic screenings[1] for commercial and educational purposes, and the hitherto-unseen personalization and customization of screens and content. This dissertation freezes and scrutinizes a moment in technological and commercial development, focusing upon a novel trend in movie consumption: smartphone spectatorship.

[1] Screenings outside cinema spaces.

My first-ever mobile phone was an Alcatel One Touch Max from the early 2000s. It had a three-line alphanumeric screen; the only pictorial representation it could display was a combination of punctuation marks from which one could create simple images, such as faces or early versions of emojis. It could not play music, nor could it record sound. It could not connect to internet networks, not to mention engage in video streaming. Yet, this model was still on the market when developments of one of the first smartphones, the iPhone, began before its release in 2007. The iPhone introduced full-graphical interface and touchscreen navigation and proclaimed what is taken for granted in todays' smartphones: Apple's "i" that stands for internet, individual, instruct, inform, and inspire—not least in the form of moving images.

A smartphone is a cellular phone that, besides telecommunication, includes computer functions, provides wireless internet access, and runs downloaded applications which are controlled using a touchscreen. In this dissertation, smartphones are analyzed, in addition, according to their qualities as mobile and portable devices, which require constant bodily connection to position the screen and enable it to be used in a broad scope of environments. What secured the smartphone's legacy in terms of spectatorship is precisely that it is always ready at hand or sitting in one's pocket, and grants access to movie and video content anytime, anywhere. But this nonstop access leaves its mark on moving-image consumption patterns, viewing duration, and the overall viewing experience. The smartphone's small size and the constant control over the screen position and content, though they facilitate spontaneity, are perhaps the most notable factors responsible for reducing consistent attention to the screened content and inducing attention shifts between a narration, the device, and the surrounding, often unenclosed, space.

Smartphones and smartphone spectatorship borrow a great deal from cinema and post-cinematic screens, but synthesize these features with continuous interaction between the viewer's body, the screen, and the interface. This synthesis is the result of the smartphone not being a dedicated screening apparatus: whereas its screen and software are capable of streaming, playing, and visualizing moving images, the sensory presence of a surrounding environment and the countless other parallel running applications on the device cause distractions. In this regard, smartphones are analogous to other portable smart devices, such as tablets and phablets, which may have slightly larger screens but

users are still subjects to distractions.[2] These devices have similar features—including handheld usage and touch control—and the same opportunities and drawbacks as smartphones do.

Smartphone spectatorship is not simply *à la mode* in contemporary media consumption; it signals a pervasive cultural and social shift that has followed upon the quickly expanding terrain of digital technology and digital industry. It incorporates personal tools with social behavioral norms and consumption patterns (see Chapter One), communication devices with film and video art (see Chapter Two), and narrative elements with integrated multisensory information originating from screen and surroundings (see Chapters Three and Four). The reasons for choosing smartphones to illustrate the novel realms for viewing experiences outside the dedicated screening spaces and their advancement to interactivity and personalization are manifold. Besides their indisputable ubiquity, smartphones provide a platform for continuous innovation when it comes to screening moving-image content, both on a hardware and software level. Moreover, they simultaneously integrate the features of portable, touchscreen-operated, and internet-connected multimedia devices. Possessing both portable design and touchscreen interaction, smartphones challenge the sovereignty of stationary screens and that of media devices designed with keys and controllers with their customizable user interface and control mechanisms. Viewers' bodies enliven moving images on smartphones through the features of the human body: the eyes, the ears, the hands, and fingertips.

Active bodily control places smartphones at the center of attention of a technologically and cognitively flavored research project that balances cinema-centered and tech-oriented views. This approach is fundamental in that it approximates the interests of content and hardware providers to spectator psychology and a sustainable, ethical, and pleasurable moving-image consumption. Smartphones are suitable for modeling user behavior and technology adoption, which are the central premises of this dissertation. In it, I lay the foundations for a potential methodological framework for studying the moving-image and media palette for long term uses in education, health, and entertainment. The choice of embedding new media and screening technologies into a cognitive and neuroscientific discourse, in other words, studying spectator behavior, serves the purpose of understanding reception in the abundance of audiovisual exposure.

[2] These categories overlap to a great extent: smartphones of the late 2010s often have screens larger than five inches and tablets' screen size (such as the iPad mini or the Amazon Fire 7) starts at around seven inches.

Smartphone spectatorship is only the first step. Its thorough research is fundamental to embracing various theoretical concepts and applying them to media-technological advancements of the present and near future. These advancements include, but are certainly not limited to, medium-specific adaptive narrative strategies, content generation and presentation by artificial intelligence, and smart storytelling tailored to a consumer's habits or bodily features. More specifically, these trends point toward wearable media players (such as iTV Goggles), media experience in games, interactive and 360 degrees' movies, augmented or virtual-reality environments, and immersive entertainment systems, such as Netflix VR, Cmoar VR Cinema by Oculus, or KLM's in-flight VR system. The common ground among the paradigm prototypes and technological trajectories listed here hallmark the convergence of moving-image narration into interactive platforms, to which smartphones serve as a relevant and current example.

Film Theory Research and Smartphones

Studying the role of body and mind in spectatorship has long been part of film studies' discourse. Münsterberg's (1916/2014) writing on the "photoplay" explored basic theses of movie perception in the early days of cinema. Behaviorist theoreticians, such as Sergei Eisenstein, experimented with the relationships between storytelling and comprehension beginning in the first half of the twentieth century. By engaging in rational and analytical inquiries of moving images in the 1980s and 1990s, David Bordwell and Noël Carroll, among others, paved the way for theoretical and experimental research in cognitive film studies. Applying social and behavioral psychology sheds light on audience dynamics, attention, and meaning construction, and informs the functional understanding of spectators' affective and embodied presence in front of screens.

The formal and ontological independence of moving images and moving-image presentation equally galvanizes spectator-centric, structural, and technological discourses, which discourses have played significant role in defining the trajectories of the film studies discipline. Consequently, my work cannot circumvent this threefold approach, yet will use analytical and critical approaches in the pursuit of a broader understanding of spectators and spectatorial behavior with regard to smartphones. With the aim of exploring the psychological and physiological effects of spectatorship, I combine the theoreti-

cal bases and findings of social and behavioral studies with that of film studies and the field of digital media and mobile communication. This combination seems necessary in order to expand existing scholarly insights on usability and media consumption and behavioral analyses of moving-image spectatorship.

Cognitivist approaches to studying moving images represent a substantial commitment to spectator-centric investigations of comprehension and emotional reactions with the aim of maintaining a strong connection to scientific validation as well as to objective and observable evidence. In fact, it was the poignant critique of the dominance of schools advocating psychoanalytic and semiotic theories in the 1970s and 1980s that catalyzed the integration of audience research into cognitive theories, intertwined with behavioral, experimental, and neuroscientific discourses. However, as Ted Nannicelli and Paul Taberham (2014) point out, instead of contributing to the broad field of cognitive science, cognitive film studies specifically focus on investigations of the spectator's mind and the psychological and physical effects of movie content and screen media. In the introduction to *Cognitive Media Theory*, Nannicelli and Taberham identify the relevance of the discipline in methodological and theoretical terms, where the former refers to scientific, evidence-based research and the latter to naturalistic objectives.

Cognitivists gradually progressed in the footsteps of film semiotics (e.g., Mitry, trans. 1997; Mitry, trans. 2000) and turned from a structural view of movies and spectatorship (Bordwell, 1985) to more specific inquiries. These inquiries include the spectator's body and mind, that is, spectatorship and the affective connection to filmic representation as defined by Gregory Currie (1995); human evolution and spectatorship in Torben Grodal's (2009) bioculturalist approach; the effects of storytelling and composition in Arthur Shimamura's (2013) anthology, *Psychocinematics*; relations among emotion, imagination, and stimulation in Jeffrey Zacks's (2015) *Flicker*; and the neuroscientific and biological foundation of aesthetics and film in Murray Smith's (2017) recent contribution. This handful of works are merely examples of what the fusion of psychology, neuroscience, and film studies have to offer. However, along with other studies, they demonstrate the importance and dynamic growth of both social and behavioral approaches to examining spectatorship.

According to the foundations of cognitive film studies, making sense of what a spectator sees involves neural processes during which one contextualizes and clusters incoming pieces of information and, using a compressed, individual "database," perceives a personal narrative (see Branigan, 1992). Information

clusters (schemata)[3] are deemed universal (Buckland, 2000), intentional and rational (Bordwell, 1985, 1989), and automatic (Mullarkey, 2009); however, all these approaches purport the role of spatiotemporal associations, social and cultural knowledge, and stylistic elements. Inquiries in cognitive film studies and the need to propose scientifically provable hypotheses not only gave rise to a novel research structure, but also introduced a view which connects the role of screening platforms, film form, and storytelling with the human factor. Thus, the structure of the discipline, which stands upon the building blocks of evidence-based natural and social sciences, has incorporated the study of narratives and narrative structures cultivated by Bordwell and Branigan in order to search answers to how processes of meaning-making occur.

As opposed to ideology-driven understandings of film and film narration (such as those that propel psychoanalytical film studies), a methodological framework of scientific testing makes cognitive approaches suitable for studying the links between spectators and smartphones and the plasticity of movie or video screenings. This last point is critical: smartphone spectatorship involves a great deal of active bodily presence which, via a small portable screen, an interactive interface, and a touch-driven input platform, shape the visual and sonic delivery of the content. Apart from the embodied presence of a spectator and her mental activities, smartphone spectatorship is embedded in muscle movements which extend the senses. Cognitivist approaches, in addition, have room for empirical validation and drive discussions of film and audiovisual media toward a long-term dialogue which, for mobile media experience, I intend to commence here.

Even though the key areas of this investigation are grounded in cognitivist notions, I make use of other theoretical schools that encompass spectatorship and spectatorial behavior. One of these is semio-pragmatics, a theoretical movement hallmarked by the works of Francesco Casetti and Roger Odin (Odin & Casetti, 1990; Odin, 1995), which focus on social practices of spectatorship. Semio-pragmatics has roots in an object-centric approach, according to which the spectator as research subject only exists as a social entity in the context of movie spectatorship. This means that the social disposition of any spectatorial activity is only interpretable from the movie narration's point of view. This approach is efficacious for identifying a broad contextual framework generated by the narration, the screening platform, and other sensory and

[3] Schemata are behavioral or thought patterns that systematize bits of information and structure them in clusters of meaning, synthesizing preconceived knowledge, expectations, and perceptions (Bordwell, 1985).

cultural factors that the spectator accesses. It is also useful for establishing a formula for a comparative study between smartphones and other media that would yield conclusions regarding usability and consumption. Although rejected by some cognitivists (Bordwell, 1985; Buckland, 2000), the fundamental premises of semio-pragmatics are flexible enough to contribute to this research. In my theoretical discussion, I use it to investigate the spectator's role in comprehending narrative information and accordingly adjusting their behavior, such as attention, body posture, or social interactions. This is somewhat similar to how semio-pragmatics has been applied to the specific case of mobile spectatorship in recent works by Odin (2012), Casetti (2012), and Casetti and Sampietro (2012).

The fundamental premises of semio-pragmatics and cognitive film semiotics, which define human experience based on the interpretation of signs, nevertheless paint an incomplete picture when it comes to the neural and biological effects of smartphone spectatorship. This is a result of an internal contradiction: semiotics assumes the existence of a system (a quasi-language, if you like) through which a wide range of linguistic, behavioral, and contextual patterns become accessible, but by doing so, inherently overlooks one of the elements in the spectator–content–medium trio. Focusing on spectators as biological or social entities disregards deviations in film form (e.g., avant-garde films) in favor of the institutional constraints of spectatorship (e.g., cinematic conventions). Likewise, acknowledging the plasticity of content by assigning meaning to sound and visual effects disregards the role of the screening medium. In fact, this latter case demonstrates that analyses of screening platforms using exclusively film semiotics seldom point beyond observations about stationary screens and a normative, constructed spectator. Or so it seems; changes in the visual and sonic disposition of the screening apparatus is somewhat discredited even in the above-mentioned research on mobile spectatorship by Casetti and Odin.

This deficiency can be remedied by employing an ecological view. An approach that declares an object's function through the integration of user and forms of interaction established by the Gibsonian theory of affordances (Gibson, 1977, 1979) informs the dynamic physical and perceptual relationships that exist between the smartphone and the spectator watching film or video content. Whereas Gibson's early proposals have faced opposition from advocates of the cognitive studies of aesthetics, the theory of affordances works well with the points according to which smartphone usership (including spectatorship) is about perceiving the possibility of actions and interactions. This line of thought based on affordance theory suggests that one understands information either by

recalling previous encounters with similar elements and the outcome of these encounters or by drawing conclusions from the given object's or action's attributes presented explicitly or implicitly (Bordwell, 1989). This is true for narrative comprehension, too: for instance, a gun projects the possibility of killing or causing injury and a kiss projects love. When viewers see the melting celluloid and the eddying smoke and flames in *Cinema Paradiso* (Tornatore, 1988), they project the outcome of the fire. In *The Wizard of Oz* (Fleming, 1939), when they see Dorothy clicking the heels of her bright ruby slippers, they know the shoes are about to take her home. Viewers perceive the function of objects, their use, and the potential outcomes by adopting the scenario in which a character appears to be involved. This is the result of a learning process of narrative functions and of audiovisual storytelling. Such point is a prominent one; learning as the underlying factor in spectatorship has been approached from various perspectives, including narrative composition (Schwan & Ildirar, 2010; T. J. Smith, 2012; Hochberg & Brooks, 1978/2017), cinematic strategies (Kraft, Cantor, & Gottdiener, 1991; Prince, 1996), event or action segmentation (Levin & Varakin, 2004; Zacks, Speer, Swallow, & Maley, 2010; Zacks, April, 2010), and medium-specific learning and comprehension (Messaris, 1994; Troscianko, Meese, & Hinde, 2012).

Works on narratology and human cognition (for instance, by Bordwell, 1985; Branigan, 1992; Grodal, 1999, 2009) address spectatorship in terms of predicting upcoming actions. These inquiries seek answers about the neural processes of subjective experience, suggesting that the mechanisms of recognition and projection outlined above are the results of the spectator's embodied presence in fictional narratives. Vittorio Gallese (2005) offers evidence of this in his work on the mirror neuron paradigm, which introduces the neural processes responsible for intaking, recognizing, and copying others' actions and reactions. In other words, neural connections are activated for specific actions when one observes another agent performing those actions. Perception of an action activates the brain areas responsible for that action, even if the action is only communicated verbally: in an experiment in which subjects were instructed to maintain a consistent posture, they were more likely to fail when they heard a voice explaining a motor action, such as standing up and going into another room (Kosonogov, 2011). Gallese and colleagues apply the mirror neuron paradigm to the specific case of narrative comprehension by introducing embodied simulation (among others, in Gallese, 2005; Gallese & Sinigaglia, 2011; Gallese, 2014). According to the theory of embodied simulation, a spectator perceives narrative actions and diegetic objects as if they are happening or located in her

own physical surroundings. In addition, a spectator is capable of anticipating upcoming and inferring past actions by "mirroring" characters' mental state.

Embodiment as a strategy for constructing meaning is widely addressed in film and media theoretical and experimental research (among others, Marks, 2000; Sobchack, 2004; D'Aloia, 2012a; Gallese & Guerra, 2012; Coëgnarts, 2017; Kiss & Willemsen, 2017). Whereas Laura Marks and Vivian Sobchack treat embodiment as the affective presence of the spectator's body in relation to a visually and sonically depicted space, Kiss and Willemsen propose it as part of a problem-solving activity. Although the former approach employs phenomenology and the latter embraces a clearer cognitive mindset, the two coincide in analyses of narrative immersion. Both cognitive and phenomenological schools accept the existence of mental presence in an imagined (fictional) space, which is, again, a fundamental premise of the above-mentioned neuroecological and neurophenomenological theses of narrative comprehension illustrated by the embodied simulation theorem, which serves as a base for understanding the smartphone spectator's embodied involvement in shaping narrative presentation.

In the case of smartphone spectatorship, the matter of space points beyond the attributes of the diegetic space. Similar to other screening platforms, smartphone spectatorship equally—if not surpassingly—concerns the geometrical outline of the screening space and the screen's positioning in that space. The screen as an image container is discussed by many, including Bordwell (cognitivism and narratology) and Gibson (ecological psychology), as well as in various entries in Dominique Chateau and José Moure's (2016) *Screens* anthology, such as those by Christie, Lefebvre, Toddi, and to some extent Huhtamo, who inspects the "archeology of the screen." Simon Lefebvre's account is of particular interest here. He introduces a special case which lacks the surface that divides diegetic from physical space: three-dimensional screenings, where the spheres of coexistence of physical bodies and objects with diegetic ones fuse not only in auditory, but also in visual perception. This understanding of the screen (or its lack) helps to bridge the gap between the strict structural outline of movie spectatorship that is familiar from cinema and other inherently stationary screens and the plasticity of smartphone screens and spectatorship. Lefebvre's conclusion is relevant, primarily because his account describes the simultaneously immersive and distancing effects of 3D screenings, similar to one of the central points of my study regarding spectators' bodily involvement in smartphone spectatorship. He implies that the physical appearance of a 3D movie's image and sound propagates within the cinematic space

surrounding spectators, generating a lifelike composition of objects in space. But since the experience is incompatible with cinematic screenings in general, the additional dimension (and information which makes it similar to real-world experiences) nevertheless distracts the viewer from being immersed in the narrative environment. Viewing content on smartphones possesses an additional distracting factor: changes in perspective (the screen's position and distance from one's eyes) and interactions of other sorts with the footage through image and sound settings is dependent upon and carried out by the spectator.

The question of immersion, in other words, transportation (Gerrig, 1993; A.-L. Cohen, Shavalian, & Rube, 2015) or absorption (Bálint & Tan, 2015), reflects on the state of feeling present in the sphere of narrative actions. Although it has been widely studied with regard to audiovisual media, such as film and virtual reality chambers (see Burch, 1979; Bordwell, 1985; Lombard & Ditton, 1997; M. Slater & Wilbur, 1997; Mestre, 2005; Visch, Tan, & Molenaar, 2010), the sensation of social, emotional, and perceptual presence lacks a straightforward applicability to scenarios where momentary control is in the spectator's (listener's, observer's) hand. This includes cases when the intensity, size, and other attributes of the stimulus source change under the spectator's control. Smartphones illustrate such cases, as they, unlike stationary screens, enable perspectival alterations in the form of immediate and momentary reactions to a movie or external stimuli. The factor in play here is the hardware design, more specifically the size, the weight, and the touch-sensitive screen of the device which together contribute to portability and mobility, and immediate options for content control.

Bodily control is an important feature of handheld multimedia appliances. Heidi Rae Cooley (2004) approaches handheld devices (in her terms, mobile screenic devices or MSDs) as biomechanical entities, claiming that touch and kinesthetic control are closely connected to the ergonomics of the device and the physical capacities of the user's hands and arms. Cooley's hypothesis concerns social learning and obtaining the skills for automatic interactions with mobile media tools. In a corresponding work, only about portable game consoles, Nanna Verhoeff (2009, 2012) uses a practice-based, historical angle to examine Nintendo DS game consoles. She claims that the DS and any other such device for personal use convey historical and social status and, consequently, ought to be studied in the context of functionality. According to Verhoeff, the scholarly eye cannot fail to identify these tools as platforms for information input and output, which view rewards scholars with knowledge of the terms, habits, and economic values of usage. Although products in the DS family have been on the

market since 2004, their similarity to smartphones is restricted to their multimedia capability, graphical user interface, and haptic control. Yet, Verhoeff's and Cooley's points are still valuable for defining the epistemological and cognitive grounds for touch interaction with a virtually represented reality, displayed on a smartphone's screen.

Smartphone research is not equally generous. Although mobile devices (cell phones and even portable media players and tablets) have advanced and gained great significance within a fairly short period of time, there are just a handful of scholarly entries on usability in regards to these devices that have been game changers in entertainment. The majority of sources in the field concern the anthropological side of consumption and round out the already wide range of technological, industrial, economic, and information-technological publications. In fact, mobile- and smartphone research is still largely limited to the social and technological patterns of communication, online behavior, and virtual identity (Katz, 2003; Ling, 2004; Goggin, 2006; N. J. Johnson, 2006; Koskinen, 2007; Goggin, 2008; Goggin & Hjorth, 2009). Other findings in the discipline are relevant as historiographic entries about the mobile evolution shedding light on the speed of technological and economic change, as well as the challenges these changes create for users (Katz & Aakhus, 2002; Horst & Miller, 2006; Ling, 2012). In terms of mobile media consumption, the convergence of media into mobile content (e.g., maps, books, games) has been addressed scarcely (see, for instance, Ling, 2004; Jenkins, 2008; Goggin & Hjorth, 2009), but even these stop short of presenting profound theses about smartphone and moving-image content.

When it comes to mobile devices, communication, and media, a number of factors must be taken into consideration. Still, the most critical lack of research concerns inquiry into smartphones as multimedia devices, as film screens, and as film cameras. In his work on mobile phones and spectatorship, Odin (2012) offers an overview of consumer practices, including mobile filmmaking and movie streaming. Though broad in addressing multiple combinations of content and screening platforms (for instance, movies shot on smartphones but screened in cinemas), Odin's work is limited to content and defining the "proper" form of film creation and consumption. Likewise committed to a formalist overview, Paola Voci and Catherine Fowler (Voci, 2010; Fowler & Voci, 2011) discuss the cinematic roots of small-screen storytelling and conclude that mobile spectators engage with movie narration in a similar manner to that in a theatrical screening room. Their conclusion stems mainly from the idea that spectators' knowledge and previous encounters with the pictorial toolbox of cinema help

engaging with movies even in different viewing spaces. The greatest limitation of these works is being largely outdated: since the early 2010s, mobile-phone spectatorship, content access, and content making has changed due to the increasing quality of both display and recording.

Making movies on mobile devices is an emerging trend, thanks to the improving quality of built-in cameras, as well as the multitude of smartphone/mobile/pocket film festivals and "how-to" entries on blogs and in magazines. *Cellphilm* is one of the academic terms for movies shot on mobile phones coined by Jonathan Dockney and Keyan Tomaselli to describe material used in anthropological research with connotations of social awareness, education, activism, and specific aesthetic practices (Dockney & Tomaselli, 2009; Dockney, Tomaselli, & Hart, 2010; MacEntee, Burkholder, & Schwab-Cartas, 2016). Still, my work extends this approach borrowed from cultural sciences into behavioral studies. Although physiological reactions, engagement, and attention are largely neglected in the intersection of smartphone and moving-image research, See-To, Papagiannidis, and Cho (2012) made an attempt to contribute to the field by addressing Csikszentmihalyi's concept of *flow,* side by side with enjoyment in mobile entertainment. Using an approach of marketing, media, and technology, See-to et al. promise an investigation of the relationships between experience (knowledge) and efficiency, attention and engagement, and enjoyment and satisfaction. Their empirical work regarding differences in response to various genres and screens, however, lacks theoretical grounding in film studies (in terms of genres and filmic representation) or cognitive science.

This dissertation only uses content, content making, and theses on mobile entertainment as stepping stones and instead addresses the social, physiological, and bodily aspects of spectatorship. Although the case of watching movies or videos outside the cinema's institutional and behavioral framework, especially on mobile devices, has been somewhat explored, the key points I identify as the foundation of a spectator-centered approach to smartphone spectatorship (attention, engagement, immersion, and interaction) are, so far, unanswered in scholarship.

Smartphones and Spectatorship

Having identified the scholarly deficit in the wide-spread, yet scarcely researched smartphone spectatorship, I aim to apply a more pragmatic

understanding of smartphone-related practices for developing a spectator-centered approach to post-cinematic investigation. Integrating technology and user habits with a theoretical base in spectatorship yields a broad understanding of the development of screen technology, screen culture, and moving-image consumption. Additionally, acknowledging the plasticity of consumption both in terms of content and screening technology, my approach is open to future trends, too. A method exclusively focused on technological advancements would present a canonical overview of smartphones' development grounded in the ecological mapping of social and cultural practices. My spectator- and behavior-centered analysis expands this method by incorporating theories of attention, usability, and learning. This allows for responsiveness to the dynamic processes that include elements from the wide range of screening devices.

Moving-images represent a transmedia phenomenon, a form of representation that, instead of being attached to individual screens, follows fairly universal patterns of storytelling. Screening devices, storytelling strategies (for instance, in fiction or nonfiction works), and functions—whether attached to the products of social, educational, artistic, or entertainment practices—appear in unique combinations. Some of these combinations have created stronger bonds than others and secured disciplinal legacies. Fiction cinema, for instance, is an entity with clear-cut stylistic and cultural elements, the development of which can be (and is!) described through a fairly linear historical curve. Others, such as mobile media, lack a clearly distinguished framework, which can perhaps be attributed to the ever-changing agglomeration of technology, practice, and usability. This evokes a number of questions about the constantly changing system of moving-image media, its evolution, its boundaries, and interplay with spectators.

Although it might seem perilous and controversial, in my study I avoid reflecting on the genealogy of moving images, but, instead, use a historically inspired comparison between two different, but by no means distant, media—cinema and smartphones—as the basis for my arguments. The reason for this is to focus the spotlight on changes in practices and spectatorial behavior without the need to congregate multiple aspects of smartphone spectatorship into existing discourses of mobile communication. This is crucial, not only because smartphone spectatorship thus far lacks an established scholarly framework, but also because its technological, practical, and commercial schemes are way too pliable to be observed only through the mobile-phone discourse.

Movies come to life on the surface of canvases and screens. They are, however, also embodied in the form of various data media, for instance, film reels, videotapes, and digital files which are translated into vision and sound by

a screening apparatus. Two terms, two trajectories must be distinguished here: the data medium and the screening medium (screening platform). Although digitization and the development of screening devices has erased the previously tangible distinction between the two, the distinction is nevertheless essential when discussing smartphones, especially with a mind toward future developments. Smartphones' design justifies this: they are tools that visualize digital information and assign user mechanisms to it, which often appear in reference to other tools. For instance, a calculator application's interface resembles the design of a physical calculator or a calendar's the paper version's outline. Smartphones were perhaps not designed for video watching (or necessarily video making) initially, but due to their physical attributes, they can also function as screens (or cameras). Cultural understandings of smartphones' attributes (e.g., screen size, aspect ratio, camera) elevated the "cinematic" function so that users watch, shoot, and even edit movie clips on their smartphones creating practices of watching on the go, as well as a new moving-image genre parlaying the screen's size, ergonomics, and portability.

My aim is to connect media-theoretical and behavioral overviews of spectatorship to reflect on the cognitive mechanisms that are in play in smartphone spectatorship and map spectators' reactions to technological novelties and deviances. I hypothesize that spectators adapt to different viewing circumstances (for instance, screens, environments, and practices) such that they rely on previously encountered or known schematic elements. These schematic elements, or conceptual frameworks, are signals stored in the brain and originate from storytelling practices on film or from the spectator's physical and mental presence in the real world. These conceptual frameworks, furthermore, spring from social and cultural formulas, derived from social interactions and knowledge of objects and actors, for instance, knowledge of types of people, elements of nature, or urban environments. According to this assumption, sensory information guides attention and behavior and renders incoming information into context-specific or context-unspecific schema clusters. In other words, sensory information is organized according to its relevance in terms of the watched content. In order to test this hypothesis, I investigate three main areas of inquiry that are addressed in more detail below. These areas reflect on the forms of moving-image consumption and the role of narrative comprehension in acquiring medium-specific knowledge and behavior.

Smartphones to Challenge Cinema

As the quantity, quality, and accessibility of media have increased, and technology has come to allow for vast mobility, moving-image consumption and production have rapidly diverged from what is called cinema. The notion of accommodating certain behavioral frameworks emanates from the pervasive habits of contemporary, Western cinemagoers, namely, sitting in a darkened room to focus attention on the screen and the lack of physical or verbal interaction with fellow spectators or the screened movie. Such customs are encoded in the protocols of watching movies in cinematic spaces. I propose that cinema's thriving existence, parallel to various other screening platforms, such as television, home video, and DVD players, secured its position as a "default" platform for moving-image media, whereas the history of home or personal entertainment shows a less canonical curve. Smartphone spectatorship, accordingly, is understood here as an integrated part of cinema's cultural framework in contrast to, for instance, mobile communication or personal entertainment devices. This idea is reasonable due to the relatively short history of mobile phones and portable electronics' rapidly changing landscape.

Many cinematic conventions derive from the screen and screening room: some are over a century old; others were born during the development of screening and sound-projecting technology. These practices and conventions are defined by the physical and phenomenological relationship between the screen and the audience, as well by the relationship among audience members themselves. My first research question concerns the role of cinematic practices in post-cinematic spectatorship: it asks whether cinema and associated behavioral norms are revisited when viewers watch movies on screens without established screening practices. This research question reflects on whether consumption and perceptual patterns when using non-theatrical screening platforms (smartphones, for instance) correspond to cinematic ones or if those platforms establish their own spectatorial frameworks. This sheds light on the roles the screen, sound projection, screening environment, viewing practices, and the spectator's social and physical enactment play. Using this comparison, I inspect the ways in which smartphones challenge the behavioral and consumption conventions of theatrical spectatorship to explore whether conceptual frameworks contour medium-specific behavior and whether such behavior helps comprehension and engagement with narrations.

The Birth of the Smartphone Spectator

Although smartphones' reference to cinema is presumed to manifest in the demand for the viewer's constant attention to film narration, I also expect that conceptual frameworks inform and motivate changes from the protocols for viewing cinema. This means that, due to the smartphone's hardware and software design, non-narrative information received from the viewer's surrounding environment or from the device may distract her, as response to which she alters the screening. Such alterations may be changes to the screen's position (e.g., moving the screen in relation to her eyes), the device (e.g., volume, aspect ratio), or the video player (e.g., freezing a specific image).

In terms of interactions between spectator and smartphone, I distinguish and analyze two fundamental factors which shape spectatorial behavior and, presumably, affect the construction of meaning. The two factors are based on motivating forces and the nature of interaction and originate either from the viewer's personal motivations or her responses to distracting stimuli. Operating with different cognitive processes and outcomes, the two factors influence attention and spectatorial strategies in different ways.

Interactive watching is by no means unprecedented; however, a specific approach to discussing smartphone spectatorship is still required, especially because, besides involving new customs and practices, interaction affects the quantities and types of content intake, as well as the places and times of movie and video consumption. The ubiquity and portability of smartphones, and the simultaneous embodied presence of a smartphone spectator in both online and off-line environments require engaging in a practice-based discourse, which examines spectatorial routines and the social and phenomenological spheres of spectatorship. My second research question is whether cognitive processing of incoming information and its clustering into relevant and irrelevant agglomerations, respectively, inform spectatorial behavior and interactions with devices. This refers to acts of interaction and adjustments available on the grounds of the smartphone's and the various video player applications' usability formulas, as well as the economic values that have an impact on both the commercial and aesthetic features of spectatorship.

Smartphones for Spectatorship

Spectatorship is a composite process of sensing and processing incoming visual and auditory information through previous experience and knowledge. In the

arena of perception, cognition, and social presence, a question emerges regarding narrative comprehension and interpretation, principally, whether spectatorship can be described as the system that assigns a contextual framework for incoming information. In practice, this supposition suggests that when spectators have previous knowledge of given narrative schemata, such as characters or locations (e.g., woman, teacher, city, et cetera), and of screening platforms, such as cinema or smartphones (e.g., public screening room, mobile communication device with various functions), they are able to identify and operate with high-level schematic constructs that aid comprehension of audiovisual narratives. Processing these constructs provides orientation among context-specific or context-unspecific information. For instance, this supports or hinders the comprehension of moving-image narration or even provides valuable orientation points in the spectator's surrounding physical environment that suggest possible interaction strategies.

Considering smartphones' properties (mobile, enable a number of interactions), as well as their related usability and practical formulas, we can imagine a hypothetical scenario, according to which moving-image content is watched in a space where such activities are less typical than in a designated screening room, given the amount of parallel activities and information sources. Such a scenario could take place on a crowded street or in a subway car, where a greater variety of parallel information sources and behaviors are available than in a typical cinema room. The configuration of these sources of sensory and cultural patterns affect narrative comprehension when one is watching movies or videos on a smartphone. By this, I mean that while narrative composition of film guides attention, newly appearing stimuli or tasks in one's surroundings can equally be stimulating, given that such stimuli either induce epistemic associations with the movie narration or remove attention from it. Such a proposition derives directly from the cognitive role of montage: spectators can comprehend the juxtaposition of temporally and spatially disconnected objects in a way that supports associations between disparate sensory and cultural information.

The importance of this theoretical passage is to recognize the elements of spectatorship—especially in the novel realm of smartphone spectatorship—that are responsible for acquiring medium-specific (that is, related to moving-image storytelling) and platform-specific (screen-technological) knowledge. This leads to an investigation concerning spectatorial strategies to specific screens. My third research question concerns the process involved in contextualizing incoming pieces of information, understood as organizing information to create coherent meaning and identify physical and social indices. With this in mind, I

specifically ask whether conceptual frameworks aid adaptation to new media platforms and whether this occurs irrespective of screens. This suggests two possible points of departure. First, that sensory references are connected to concepts previously accessed via mediated or unmediated[4] stimuli. Second, that these references are classified as relevant or irrelevant in relation to a source of information (in our case, a movie narration).

Integrating Theoretical and Empirical Methodology in Spectatorship Research

Smartphones can be described as tools that extend the user's body. However, scholarly discussion of them is in debt of a theoretical and empirical methodology that is able to provide an instrument for testing small-screen spectatorship in unenclosed spaces and the spectator's bodily involvement in interactive watching of smartphones. Therefore, I engage in a theoretical discussion and propose empirical instruments to measure how physical contact with smartphones affects spectatorial experiences. Studying the behavioral and physiological attributes of smartphone spectatorship introduces an elemental framework that attempts to resolve the tension between media-historical premises and cognitive–behavioral domains.

I present a comprehensive overview of viewing experiences based upon cross-disciplinary research involving psychological, neuroscientific, and phenomenological methods to approach viewing activities through the lens of the spectator's body and mind. This amalgam serves as a tool for investigating the cognitive effects of smartphone spectatorship and exploring the possible paths for tailoring content and platforms according to viewers' needs. For this, the best-suited approach is an evidence-based one that seeks causal links between moving-image consumption, decision-making, interactive viewing, attention, emotional reactions, and comprehension.

Following the line of inquiry outlined in the research questions above, the dissertation is divided into three theoretical chapters, which discuss the different aspects of user behavior and serve as foundations for the fourth chapter, which constitutes an experimental verification. The empirical chapter proposes a methodological framework I developed specifically to study physiological

[4] Not mechanically produced using technology.

reactions to movies watched on smartphones with the goal in mind that the methodology and results may serve as points of departure for future empirical studies in the field of movies and immersive screening platforms. Combining a comprehensive theoretical discussion with a corresponding experiment debates prevalent rigorous scholarly trends that build a barricade between theoretical and empirical inquiries, preventing methodological permeability in both academic and practical discourses. Thus, I aim to contribute both theoretically to film and media studies and empirically to the cognitive and behavioral studies of moving images.

My research methodology is grounded in three branches of scholarship which are closely connected to the practices of smartphone spectatorship: first, in cognitive film studies and behavioral analysis of spectatorship; second, in media-theoretical research; and third, in discourses on embodiment and enactive spectatorship. Although these three branches are woven through the entire body of the dissertation, each chapter applies a slightly different methodological approach in order to effectively access the most fundamental principles of smartphone spectatorship, which are the character of haptic and kinesthetic interaction, the sensation of presence and emotional engagement, and that sensation's grounds in prior forms of spectatorship.

Due to the pace of product development and the continuously ascending indicators of mobile video consumption, a detailed description of manufacturers or operating systems and, thus, image and sound quality is beyond the scope of this text. Moreover, I also omit scrutiny of different individual or cultural practices, in favor of exploring a general framework of viewer behavior. My analysis occurs from the perspective of a potential spectator in generalized viewing circumstances. I assume a baseline routine of moving-image consumption in unenclosed spaces (e.g., on a busy street or in transit), where people, objects, and social norms may interfere with viewing activities and prompt changes in screen position, aspect ratio, stimulus intensity, and flow of content.

The term consumption here indicates more than the choice of movie or video material and screening platform. It refers to the act of spectatorship as part of a system with complex social, economic, behavioral, and cognitive elements. Spectatorship is understood as the cognitive processing of dependent and independent multimodal and multisensory information to create a coherent narrative. It follows that spectatorship is a concept that incorporates the action of watching or *spectating* moving-image content with additional sensory modalities, originating from the device, or surroundings. Except for specific cases when I discuss questions such as sound transmission or the role of haptic involvement,

I refer to an integral sensory experience. The role of sound in viewing experience is indisputable, even if sound transmission through a pair of earphones or headphones largely differs from the surround sound of cinema and other screening platforms. Hapticity, though, is a fairly novel instrument born of the intersection of audiovisual content and new media technologies. In the case of smartphones, touch and hand gestures lack exploratory, information-collecting duties. In other words, when gestures prompt interactions with a smartphone's interface or screen content, haptic interactions only provide physical information about the hardware—size, weight, texture, and the like. Haptic interactions play a role in adjusting sensory perspective and content presentation, and in establishing a bridge between the movie, the screening platform, and the sensory organs. The ability to control these factors of spectatorship calls for a screen- and content-centered approach to movies, requiring a comprehensive examination of screen, touch interaction, consumption patterns, and physiological reactions.

The focus areas that orient the structure of the dissertation are the following: (a) locating the ontological constraints of medium-specific spectatorship in relation to the cinematic in Chapter One; (b) identifying the ways in which screen size, mobility, portability, and spectators' corporal involvement affect viewing experiences in Chapter Two; (c) defining patterns of the cognitive processes involved in extra-cinematic viewing in Chapter Three; and (d) determining the effects of smartphones on attention, immersion, and spectators' overall involvement in spectating on smartphones in Chapter Four.

(a) In order to relate smartphone spectatorship and consumption to cinematic viewing, I begin with an approach inspired by media theory and media history. This is a tool frequently used to characterize the technological development of screening media and the social impacts thereof. Drawing conclusions about the evolution of screens and sound systems are suitable for cases which target a comprehensive genealogy. In this case, however, mapping the process leading from cinemas to smartphone video players is not only trivial, it is also misleading. Therefore, instead of drawing a straight timeline, I regard the two as end results of two separate innovative processes, which happened to intersect in today's social and technological disposition of screening media. Such an approach is beneficial for providing the study of viewing practices an essential instrument—that of comparability through the tools of remediation and premediation (Bolter & Grusin, 1999; Grusin, 2004, 2016). The basic theses of this approach include that screening platforms integrate the ca-

pacities of other media, and by doing so, adopt and reinterpret those capacities to suit their own means. This implies that a methodology grounded in comparative analysis of cinema and smartphones connects the two in a bilateral cause-and-effect relation, enabling a closer look at the spectator's role in harnessing and shaping both.

(b) Inspecting the smartphone and its usability, I use an ecological and cognitive approach to understand the mechanisms that guide spectatorial behavior and inform intake of narrative information. Here, I borrow methodological solutions from ecological psychology, cognitive film studies, and cognitivism in general, but assemble them using tools from phenomenology. Such an aggregation is rare in academic discourse (though not unheard of, see, for instance, Bayne & Montague, 2014), but essential for combining a quantitative overview of spectatorship, narrative consumption, and spectators' feeling of presence with a qualitative study of spectatorial involvement. While a cognitive approach is valuable for understanding the effects of screen size, surroundings, and the spectator's previous encounters with movies and different screening platforms, phenomenology illuminates the subjective nature of experiences which presumably drive bodily interventions. The conceptual character of this methodology balances on the border of those conscious and automatic processes involved in voluntary and spontaneous decision-making, which govern mechanical (kinesthetic) and contextual (tactile) interactions with devices and screened footage.

(c) To continue connecting the attributes of the development of screening platforms and spectators, I turn to a close inspection of how medium-specific practices and meaning-making strategies are acquired, making use of the neuroscientific roots of cognitive theory. My goal is to characterize the ways in which sensory inputs (originating both from the screen and surrounding space) affect narrative comprehension. Cognitive film research often justifies its claims by connecting semiotically inspired arguments (e.g., constructivism) to behavioral theses. Still, such a methodological path limits the insights of that research to logical derivations of socially and culturally determined reaction prototypes and neglects a close examination of intentionality and consciousness. In making this statement, I do not intend to criticize theoretical models of cognitive film studies (especially due to their appealing association of the body and the mind), I contend that a granular exploration of embodi-

ment and bodily reactions can benefit a reconstruction of theoretical approaches as empirical models.

(d) A theoretical description of learning medium-specific behavior, particularly regarding sensory, motor, and cognitive involvement in smartphone spectatorship, contributes to an empirical model that is tailored for comparisons between portable and stationary screens. I attempt to outline a comprehensive model of attention, emotional engagement, and narrative understanding using eye tracking, electrodermal activity (skin responses measured on the surface of the skin), self-evaluation survey, and comprehension test. Such a behavioral experiment provides tools for measuring the impact of viewing circumstances: whereas cinema, television, and other screens are mostly located at a fixed position from the viewer, audiovisual content on smartphones is typically accessed in unenclosed spaces with a multitude of activities, social roles, and stimuli that can induce attention oscillations between the movie's diegetic space and the spectator's physical environment. In order to recreate these specificities, I used a two-by-two factorial design to isolate the effects of screen type and distraction. Overall, the experiment was designed to recreate practices connected to smartphones and large, stationary screens, and to analyze whether or not screen size and interruptions affect viewing behavior.

Structural Overview

The theoretical and methodological approaches described above serve as the skeleton for the outline of this dissertation. Each section uses different methodology to provide a comprehensive and detailed overview of smartphone spectatorship—interlacing theory with practice, description, and experimental methods, and movies with cognitive science throughout the dissertation. I thereby ask questions and provide solutions regarding the historical foundations and future trends of new media consumption, the sensory and cognitive aspects of interactive spectatorship, not only for describing spectatorial behavior, but also for potentially identifying medium-specific storytelling instruments.

Opening the discussion with the smartphone screen and its ties to earlier screening media, Part One, "Smartphones and the Spectator's Body," lays down the theoretical ground for spectatorship on smartphones. Its two chapters

focus on the remediation of cinematic references and the attributes of multisensory and interactive viewing, with the goal of identifying innovations and challenges of the new trends of spectatorship. Part One regards viewer behavior from a film and media theoretical perspective and fuses it with phenomenology and cognitive and behavioral psychology. It applies these approaches to the case of smartphones to understand the constructive role of earlier media on the fast-growing (and fast-changing) variety of screening platforms.

The specificities of smartphone spectatorship include the unenclosed (or even undefined) viewing space, the physical contact between the spectator's body and the interactive interface, and the screen's size. As a point of departure, Chapter One, "Remediating Cinema," evaluates these specifics in comparison to theatrical screening rooms and stationary screens, which have defined spectatorship for a good part of the past century. Chapter One uses cinema as a case for a comparative analysis for several reasons. Most importantly, the two are located at opposite poles of a wide range of film-screening platforms popular in the mid- and late 2010s and represent different consumption patterns.

While Chapter One concentrates on smartphones' mediating, behavioral, and social relations to cinema, Chapter Two, "Pocket Cinemas," establishes a view of smartphone as an independent screening platform in terms of interface, interactivity, mobility, and portability. The chapter borrows concepts from haptic media and embodiment, human–computer interaction, and virtual-reality research, as well as from interactive cinema, to establish smartphones' role in contemporary, Western moving-image consumption. It builds upon the findings of the comparative analysis of Chapter One by applying a threefold framework to screening space, bodily involvement, and screen specifications. As a consequence of mobility, smartphones invite users to be physically immersed in creating their viewing experiences by adjusting the screen position. This deviation from the conditions of traditional viewing induces novel perceptual mechanisms and changes to the sensory scope. The touchscreen gives rise to physical interactions by touching and interacting with the interface and content presentation. The primary aim of this chapter is to analyze the smartphone's role as a film- and video-playing apparatus, as well as its flexibility, tangibility, and immediacy. These practices create a specific type of bodily and cognitive involvement and blur the boundaries between public and private, narrative and reality, watching and "non-watching."

The chapters of Part Two, entitled "Technology Adoption and the Physiological Effects of Smartphone Spectatorship," shift the center of discussion toward the specific domain of perception and cognition. In order to

connect the theoretical components of cognitive film studies and neurocinematics to experimental validation, Part Two is divided according to the used methodological solutions. Chapter Three, "Neurocinematics and Portable Screens" explores how spectatorship has evolved into a skill, what cinema experience had to do with it, and how this skill is applied to new media platforms. Also, Chapter Three describes the novelties smartphones offer to viewers, the way these novelties are acquired, and contemplates whether such properties will define film consumption in the future. This chapter employs a theoretical framework of film spectatorship and spectator behavior to develop notions of acquiring medium-specific viewing strategies. This means the skills needed to identify and interpret narration and narrative cues, which requires a spectator's perceptual (sensory) abilities, textual understanding, and familiarity with the screening platform. These include the effects of habits and previous knowledge of certain media platforms, and the viewer's focused attention on visual and sonic stimuli.

Chapter Three serves as a link between the preceding theoretical discussions and the empirical study, while Chapter Four, "The Physiological Effects of Smartphone Spectatorship," presents the methodological means for experimental testing and the results of the empirical validation. The experiment aims to further deepen understandings of smartphones' role in reorganizing consumption of a wide range of audiovisual content in relation to hardware design. Using biometric measures and surveys, the chapter returns to and examines the effects of screen size, physical connection between the user's body and smartphone, and viewing environments on attention, engagement, and sensations of presence. It, moreover, closes the dissertation with the intention of providing a new path for studying portable movie screens, which can serve as a stepping stone for optimizing screening devices as well as content made for small-screen distribution. As portability and flexibility become increasingly important, trends point in the direction of a broadening market not only for smartphone film consumption and on-demand streaming, but also for mobile filmmaking.

PART 1 Smartphones and the Spectator's Body

Chapter I.

Remediating Cinema

Cinematic and Extra-Cinematic Viewing Experiences

In the history of moving images, cinema has defined both the physical and social conventions of spectatorship. As one of the first screening media to set photographic images into motion and present them to a large audience all at the same time, cinema has endured the course of technological transitions and configured narrative formulae, as well as the related cultural and technological *modus operandi*. The material and ideological legacy of cinema linked the particulars of mediated reality to collective spectatorial experience.

Starting from the early nineteenth century, numerous apparatuses attempted to establish the grounds of spectatorship by gradually settling the technological and social foundations of copresence, multimodality, and filmic narration (see the Introduction to Münsterberg's work, *The Photoplay*, 1916/2014). Later, during the interwar and post-World War II periods and the emergence of national schools of neorealism, novel narrative conventions reformed both the spatial and temporal frame of moving-image consumption and straightened cinema's hegemony. Cinematic entertainment eradicated its role as a background spectacle for social gatherings and the customs of spectatorship became chiefly defined by the need to devote full attention and time to a story being told. Early cinema (as well as pre-cinematic theatrical entertainment) "preserved a perceptual continuum between the space/time of the theater and the illusionist world on screen." Then, the new trends of classical cinema separated the "screen and theater space with its regime of absence and presence and its discipline of silence, spellbound passivity, and perceptual isolation" (Hansen, 1995, p. 139). Cinema as an equally social, cultural, and artistic phenomenon

reinforced participation with newly appearing forums of quasi-experts and cinephiles, the empathic advocates of cinema's formal hegemony, who were, then, opposed by the agents of a broadening television and then video culture. Video not only democratized moving images, but also reinterpreted all aspects of film consumption—including discussions, trade, age limits, and even spectatorship itself.

Now, over a century after the first historical cinematic screenings, the theatrical moving-image culture is less prevalent. Movies are abundant and widespread, screening appliances are ubiquitous, much of them are mobile, some even lack physical data media. Contemporary screening platforms no longer delimit movie consumption into a physical or social framework; they instead foreground bodily and mental associations between viewer and content. Smartphones serve as suitable examples of such trends: they are portable, their screen touchable, and their image digital. Moreover, the viewing environment is not physically restricted, and the moving-image content is streamed without being confined into a discrete data medium. Despite these distinct technological properties, smartphones integrate functions that belong to the broader scope of cinematics. They are used to create, access, consume, distribute, and evaluate audiovisual content, thereby incorporating and remediating the meanings, purposes, and practices of cinema.

Stages of technological and formal development defined cinematic spectatorship as focused and static; the state primarily associated with engaging with moving images. But despite cinema's dominance, the increasing numbers of smartphones and other new extra-cinematic screens challenge immobility and compel spectators to take part in shaping their own audiovisual experiences. These screening platforms compete with the vivid imagery and sound of theater screenings. Beyond this, they allow for, even demand, a great deal of user involvement—unlike cinema, which lacks parallel channels and minimizes parallel activities. Another crucial factor that distinguishes smartphones from traditional cinematic experiences is the spectator's embodied presence: cinematic conventions attempt to obscure and reduce awareness of a physical apparatus by eradicating corporal interaction with the screen in favor of transmitting and preserving artistic instruments. By contrast, smartphones define the spheres of interaction along the intersection of the screen and the viewer's body.

As a communication device, cinema conveys audiovisual narratives through the language of film to an audience within the ontological constraints of the screening space and time. In a similar manner, audiovisual content on

smartphones could be described as a sovereign system of form and language that is optimized for the spatial, temporal, social, and technological standards of the device. While such reasoning applies to the cultural phenomenon that floods video sharing sites with user-generated content, it would hardly present the entire picture. It severely neglects the fact that a significant amount of moving-image content (e.g., films, series, television programs) consumed on portable devices is created using the cinematic or televisual standards of storytelling and produced for a wide variety of screens. This means that, although content production established a set of new technology-conforming or technology-driven techniques (for instance, vertical image capture and aspect ratio suitable for smartphone displays), the changing spheres of movie spectatorship has so far had little impact on audiovisual storytelling in general. Yet, I argue that digital and portable media tools present new economic and practical values that can affect the aesthetic features of movies, as well as the interplay of professional and non- or semi-professional content providers.

It nevertheless seems that professionally made moving-image content is apt to follow traditions developed before the era of new media and only sporadically reveal tendencies to optimize for newly emerging screens. Evidence of this can be found in the now fairly rich smorgasbord of feature and short films shot by renowned directors with smartphones; and not least in the fact that many of these films tend to hide the presence of a phone camera behind cinematic storytelling techniques. Meanwhile, the economic benefits and the novelty factor of mobile filmmaking are often widely proclaimed. At the end of the next chapter, I return to the question of smartphone films in a comprehensive study of the challenges of audiovisual representation, especially within the tide of user-generated content. By way of introduction now, it is nonetheless important to note that smartphones used as cameras reduce filmmaking budgets; digital distribution increases audience numbers; and online streaming results in cheaper access for viewers. In addition, films that are shot and/or distributed on mobile platforms circulate without carrying the weight of industrial and commercial procedures, and target the momentary attention of a single spectator. Though, as Voci and Fowler (Voci, 2010; Fowler & Voci, 2011) noted nearly a decade ago, the apparent logic that underpins both the formal and modal ingredients of moving-image storytelling even on mobile devices sprouts from the roots of cinematic references.

Storytelling schemes mostly remain intact, but producers and distributors already pronounce and praise smart- or portable-device optimization in independent filmmaking. Setting an example, *The Silver Goat* by Aaron Brookner

(2011) was proclaimed the first film ever made specifically for viewing on portable devices. A romantic drama about an actress and her maneuvers between family and theater, *The Silver Goat* uses the artistic world as a backdrop—perhaps no coincidence considering the way the premier campaign connected art with personal realms. Attempting to prevent the film from sinking into the overabundance of Hollywood blockbusters, the release campaign followed the newest trends in movie consumption and spectatorship: *The Silver Goat* premiered in the United Kingdom through an iPad application aboard a double-decker bus traversing the film's shooting locations in Central London, with a red-carpet moment at a bus stop on the South Bank (see Figure 1). Its subsequent international release was also conducted through the application, which, by then, was available for multiple mobile platforms (Tagholm, 2012).

Figure 1. Premier of *The Silver Goat* in London, 2012. Image reprinted with permission from Third City.

Although the PR touch is evident, such a release strategy is not only significant with regard to publicity and the creative ways in which independent filmmakers make use of technology and social practices to reduce filmmaking and distribution costs. It also reflects on the ways cinematic viewing customs and consumption strategies are challenged and extended by handheld screens. Along with other types of mobile smart devices, the iPad used at the premier of *The Silver Goat* is portable and personal yet still public: it can be taken outside of a designated room to a bus, to the open space of London. By referencing cinema (the ordinary space for film premieres), the iPads turned the iconic red Routemaster into a screening room containing simultaneously public and private viewing experiences: private in the sense that each viewer had a slightly different encounter with the content on their own screening device, while publicly sharing the same space. The app's release also supported this duality:

although promising unique and intimate engagement with the movie, access included geotagged information and download statistics (Tagholm, 2012).

Figure 2. Stills from the movie trailer for *The Silver Goat*. Screenshots from video material. Images reprinted with permission from Third City.

The Silver Goat marks another equally crucial matter, which I attribute chiefly to mobile distribution and less to handheld recording devices (the movie was shot with a Canon 5D handheld camera; Pinball, n.d.). Although notes on the visual language are hidden behind the PR campaign (perhaps to prevent questions about its value as a professionally made feature film or delimiting the range of possible screening platforms), it is clear that the scenes selected for the movie trailer and other marketing material foreground medium shots and close-ups (see Figure 2). These shots and the fairly long takes showing one or two characters at a time with little in the way of changes in perspective or camera movements seem to fit to the mobile-screen experience and the potentially distracting viewing space: they present only a small amount of visual information organized around an enlarged semantically meaningful object, such as a face or upper body. This observation yields two preliminary claims. First, framing in the trailer for *The Silver Goat* compensates for eventual distractions (e.g., looking away from the screen) and provides necessary information viewers can access through the visual and auditory channels. Second, these takes

augment the visual information of the characters' bodies and the presence of the spectator holding the screen. This dominance of body- and face-centered shots is not representative of the entire movie, which suggests that the filmmakers deliberately used scenes in the trailer that embellish mobile photography with the aim of enhancing the ontological proximity of the filmmaker, the viewer, and the characters.

By exploring the remarkable intersection of past and future (well-illustrated by *The Silver Goat*'s black and white imagery streamed on mobile platforms), I intend to direct attention to the mechanisms that distinguish and associate cinema and portable screens and identify the levels of viewing experience at which the two platforms differ and where they share common ground. In this pursuit, I analyze the technological and social dimensions which have shaped consumption and which keep cinema in strong competition with the newly emerging personal screening platforms. Comparing the spectator's physical, phenomenological, and social relations to the respective screening platforms presents the role of cinematic spectatorship in new media consumption. In the following, additionally, I explore whether viewing experience differs significantly in terms of attention, physical presence, and pre- and post-filmic interactions, and whether the true peculiarity of smartphones lies in the new behavioral frameworks that their mobility, versatility, and interactivity require.

Relocating the Spectator: Where Is Cinema, Where Is the Spectator?

Inspired by historical and cultural domains, analyses of screens through the lens of cinema are not unprecedented. Around the turn of the new millennium, film scholarship underwent a performative shift with a rising inquiry about where cinema would be amid new media; and whether, if it could be found everywhere among domestic and commercial screens, it really was cinema. This focus marks the continuance of a spectator-centric view, which gained importance in the discipline during the last few decades. The paradigm that bestows an observable social and mental role upon the spectator is the result of technological and economic catalyzers of consumption, which defined on-demand movie culture with extended spatiotemporal spheres. These processes established a lone spectator with an unyielding attitude toward the private exploitation of movie spectacles.

I argue that an audience is a dynamic system of individuals with individual experiences instead of a homogenous body. This implies that an examination of

the collective reception of cinema spectators is somewhat flawed when pursuing a comparative framework of cinema and smartphones. Such an approach is validated by a sociological understanding of audience behavior[5] which highlights the fact that individuals facilitate their behavior according to the presence of others. But this phenomenon functions less like an elemental behavioral framework of physical copresence; rather it points to a chain reaction in which social presence affects behavior simply because of an increased quantity of environmental input. This also means that if we observe spectatorship as a province of perception, cognitive processing, and social presence, the key differences between film screening platforms cannot be condensed into a framework of collective versus individual. It is equally crucial to address physical and mental presence in public and private spheres. Hence, I approach the spectator as an individual whose behavior and knowledge are embedded in and affected by collective social dynamics—in the cinema's case, by accommodating audience behavior and, in the smartphone's case, by accommodating virtual and physical presences in various spheres. Consumption and spectatorship, which are the common denominator between cinema and smartphones, are the point of departure for understanding the individual spectator who consumes, senses, and perceives the movie and who interacts with others online and offline.

I hold that both cinema and smartphones occupy significant roles in present-day movie spectatorship. *The Silver Goat* reminds us that mobile communication and portable screening platforms may keep having an impact on cinematic premieres and audience numbers, while cinematic references proliferate in various social contexts, such as a screening aboard a bus. This concurs what Lev Manovich (2001) formulated at a fairly early moment of modern commercial computation. He claimed that cinema, just as much as computers and other media, belongs to a system of information transmitters, which means that its form and content are in close interplay. One implication of Manovich's observation is that because of the merging of form and content, there is no clear divide between screening platforms. This notion is confirmed by Francesco Casetti's conclusion that, with the end of its monopoly, cinema is everywhere and nowhere at the same time. Casetti (2008b, 2011b, 2012) explores the *relocation* of cinema onto various screening platforms, and holds that the cinematic medium persists even on the small screen of an early version of Apple's iPhone (Casetti &

[5] *Audience effect,* or *social facilitation,* explains the tendency to perform tasks differently under peer pressure due to a different level of mental activation and task complexity. See Strauss (2002) and for recent work on cinematic audiences, consult Hanich (2018) and the discussion below.

Sampietro, 2012). In a corresponding manner, Odin (1995, 2012) and Fowler and Voci (2011) hold previous cinematic experience accountable for the mental engagement with moving-image narration, drawing a similar conclusion to mine earlier in this chapter regarding the evolution of moving-image storytelling.

Although screening apparatuses and environments often accommodate transmission of stimuli, the dynamics of relocation implies the presence of another component: when watching audiovisual content in a different (private or public) sphere, the desire or urge to recreate cinematic space is driven by cultural, rather than purely technological factors. This means that upon encountering a cinematic object (i.e., a movie), a spectator tends to create an accommodating environment by, for instance, darkening the room or adjusting the image and sound so that they resemble a cinematic encounter.[6] Relocation enables exploration of new spaces and practices while recreating the conditions of cinematic experience.

Despite the apparent differences between the technical, social, and perceptual means of smartphone and cinematic spectatorship, there are similarities that reveal the cinematic roots of mobile spectatorship. In *Brief Encounters: Theorizing Screen Attachments Outside the Movie Theatre*, Fowler and Voci (2011) state that the visual and sonic associations indicating cinematic references mean conformity and predictability. Fowler and Voci's explanation concerning the "other cinematic screens" (para. 1) ignores the capacities of screening apparatuses and correlated social norms. It instead emphasizes the role of actual perceptual activity, that is, the act of focusing one's consistent attention on a stimulus for a given period of time. They hold that these processes are responsible for defining and transmitting the cinematic experience and that it is a temporal framework that distances or bridges mobile spectatorship and cinema. The timeframe in cinema is usually described as an extended, uninterrupted period, while on the highly attraction-based smartphone, it is momentary. Such "brief encounters" with audiovisual representation and the account of cinema's relocation indicate that, apart from the material dissimilarities, cinematic references reappear on new media screens. This also suggests that a methodological link between cinema and smartphones is beneficial for understanding how portable screens refer to and challenge both the tangible and social features of cinema.

Cinema's relocation, thus, is not a linear, one-way drift, but rather a dynamic synergy that involves technology and audience behavior. Quoting the opening

[6] A current example of the tendency that the concept of relocation outlines is what virtual reality screening platforms, such as the Netflix VR or Cmoar VR Cinema applications facilitate: movies or video content are visualized on a virtual screen in a virtual screening environment set for a cinematic or near-cinematic experience.

words of Friedrich Kittler's book, *Gramophone, Film, Typewriter*, "media determine our situation" (trans. 1999, p. xxxix); additionally, spectators determine the ways media is situated into social systems. As new environments and physiological processes define the social status of movie watching, media formats tend to conserve references to earlier practices. Cinema possesses an established institutional toolbox that ranges from dimmed lights to fixed viewing positions, but the act of "relocation 'constructs' its *own* [emphasis added] original, in order to be able to reestablish it as such in a new situation" (Casetti, 2008b, p. 6). Casetti suggests that, during the process of relocation, some features of the cinematic medium are transferred onto other media platforms. By this, he underlines that the factors that affect relocation concern media convergence, private spheres, and the spectator as an active agent in possession of experience and conventions. His insights reveal three components that require closer observation: technology, environment, and the spectator.

The cinematic apparatus is nothing less than a system of appliances, mechanisms, practices, and the spatial arrangement thereof with a tight interplay of strategic and systematic functions. Therefore, I argue for the necessity of a systematized view of the alloy of technological and cultural factors. Such a view offers the potential for understanding the relation between cinema and new screening media to serve as a base for defining the cognitive and technical means of extra-cinematic spectatorship. This methodological path is the one that motivated both the initial definition and the analysis of extra-cinematic screening devices and the need for both a comparative inquiry of smartphones and the big screen and an empirical validation of smartphone viewing experience.

Movies and Mediation

After relocation, a material change of viewing environments and circumstances, an additional approach is required to outline a more complex, contextual definition of media references. This approach, unlike relocation, is based on the temporal dimension and omits remarking upon the hegemony of one screening platform over another. In their influential work, *Remediation: Understanding New Media*, Jay David Bolter and Richard Grusin (1999) introduce the concept of *remediation* and the double logic of remediation referring to the duality of hypermediacy and immediacy. One of Bolter and Grusin's most fundamental claims in the question asserts that during the process of remediation, a given medium

integrates the attributes and functions of one or more other media (hypermedia component), but it endeavors to eliminate these references behind operational mechanisms (immediacy component). By featuring a high-resolution screen, smartphones remediate other screening media, such as television or cinema, just as media player applications remediate some features of video players. Also, by their portability (due to the small size), they remediate other handheld devices. However, as moving-image content is played on the device, these references and the user's awareness of them become masked by the content (narrative elements) and usability practices (e.g., knowledge of the ways to start or pause the video or hold the screen).

The complexity of media references and remediative techniques generate hypermedia qualities, indicating that each medium may be observed through the contextual lens of several other media. In the case of digital media tools, such as smartphones, hypermediacy is inherently observable as the visual, sonic, and haptic representation of a myriad of cultural and media references. I provide a more detailed account of the technical features and related sensory reactions of hypermediacy in Chapter Two. This section serves as a preliminary view in which these technical and sensory elements are introduced through the two strategies Bolter and Grusin (1999) list as necessary to achieve remediation. According to the first strategy, digital applications (e.g., smartphone applications) appear as objects with perceivable and operable features which afford user intervention by revoking usability scenarios related to similar, physical, or computerized tools. These methods of use appear as sensory references on the smartphone: for instance, a video player application is operated through the toolbox of visual references (e.g., the designated icons for playing, pausing, or searching the footage) and mechanical references (e.g., the pause icon is operated by touching it) to earlier video players. The second strategy Bolter and Grusin propose is the multiplication of mediation, which generates a sensation of modal completeness, again in terms of appearance and usability. They claim completeness or fullness as a direct result of the wide range of media references involved in the process and explain that, in line with the double logic of remediation, mediated items are sensed as immediate, specifically because of the complexity of a given medium. In more tangible terms, this means that the designated icons and buttons are operated (touched, pressed) without consideration of their predecessors or reflecting upon previous encounters with similar mechanisms.

Movie watching features a number of referential systems that stem from storytelling instruments, visual and sonic representation, and screen- and

screening-related practices (more details are in Chapter Three). The evolution of film narration and the development of screening technologies are interwoven—the repertoire of structural, visual, and sonic tools to mediate narratives is allocated within the physical, physiological, social, cultural, and behavioral framework of cinema. The hypermediacy of spectatorship lies exactly in this triangular system of audience, content, and medium, which controls the technological, economic, and artistic instruments that entwine filmmaking, distribution, and consumption. Paradoxically, it is the same complex system that facilitates immediacy, the direct perceptual relationship between content and consumer. To illustrate the correspondence between hypermediacy and immediacy, I return to the aforementioned example of *The Silver Goat* and iPad-screenings: hypermediacy is perceivable in the material presence of the mobile device as a result of the economical limitations of independent film production and distribution. Immediacy signifies the direct perceptual link between the viewer and the narration about the London art world that eliminates the constraints of the screen and the screening circumstances. Similarly, the visual language of the movie trailer (featuring medium shots and close-ups) appears to be a marketing stunt to build on the intimacy of private spectatorship and to enhance the spectator's sensation of being in close bodily contact with the characters. In other words, as Bolter and Grusin (1999) suggest, hypermediacy refers to technological and to social and monetary elements (e.g., distribution on mobile devices) and immediacy serves to connect the viewer to more rudimentary narratives, such as the actions and characters depicted in the movie.

The interactive physical adjacency of the smartphone reminds the viewer of the fictional and artificial nature of the stimulus, as does cinema's environment, the protocols of admission, and the presence of fellow viewers. But fictionality is not exclusively indicated by technological and cultural elements: movies reflect on their respective mediative status by direct indication of non-diegetic elements, such as temporal and spatial discontinuities, texts, or music. Thus, cinema achieves the sensation of immediacy by prompting mental processes and evoking substantial physiological responses in spectators in relation to the narration, masking the mediative strategies of the technical apparatus, the institution, and the movie itself. In the case of smartphones, these layers are complemented by bodily correspondence and the conformity of physical and mental immersion that are results of the spectator's involvement in holding and positioning the device.

Culturally familiar concepts or objects, such as the image and sound of movies or digital video player applications, generate the sensation of immediacy. Through the process of remediation, the materiality of the film reel morphs into the ideological object of cinema that is reenacted when watching movies on other platforms. This reenactment motivates the arguments of the following pages along the intersection of the spectator-centered (relocation) and the medium-centered (remediation) theoretical framing of audiovisual experience.

Just as much as the tie between cinema and extra-cinematic screens is the consequence of an indirect progression, the current form of cinema has not developed along a direct course: visual illusions, pre-cinematic and post-cinematic media platforms have likewise contributed to shaping both its form and status. Cinema, therefore, has no unequivocal genealogy, which—for instance, through television, video, and video games—led to smartphone screens. Such a statement serves as a methodological reminder to the necessity for a comparative analysis between cinema and smartphones as contemporaries and as a directive to guide a technological, psychological, and cultural framework. Altogether, eschewing an exclusively media-historical and anthropological overview grants me a way to use the widely explored cinematic paradigm as a point of reference for the scrutiny of smartphones as movie screens in order to reach a suitable formula for defining the perceptual, cognitive, and physical processes involved.

On-Demand Double Bill, in Cinemas and on Smartphones

A great number of film screens and players developed in the realm of cinema. Presumably, the smartphone neither is one of the first apparatuses to question and reform cinema's spatial or temporal limitations and its peculiar collective privateness; nor are they the last to remediate cinema and extend its physical and social sphere. Television established co-viewing within the scope of domestic presence; video and home cinema effectuated on-demand timing; and portable TV sets and computers further extended the plasticity of the viewing environment. Yet, what makes portable multimedia devices, such as tablets, phablets, and phones worth analyzing as movie or video-screening instruments is their unprecedented mobility and simultaneous omnipresence in both online

and off-line spaces. Moreover, even though commercial feature films seldom conform merely to mobile media platforms (for some notable exceptions, review the previous subchapter and consult Chapter Two), the increasing amount and quality of professional and amateur content constantly pushes for reform of the global film industry (see also, Jenkins, 2008).

From the 1980s onward, and specifically from the introduction of the European, then global, mobile communication system, the GSM, portable phones have occupied an increasingly significant role in both business-related and leisure activities. Later, by means of cellular data standards and other wireless computer networks, an infinite number of extra features became available for personal assistance, information access, and entertainment (see Agar, 2003). As a consequence, by the end of the second decade of the twenty-first century, billions of video-capable mobile phones, tablets, and players have become connected to the World Wide Web globally (see Richter's, 2013, prediction and ITU statistics, 2018). Besides their capability to integrate tools, such as calendars, maps, cameras (and telephones), contemporary smart devices provide admission to a plethora of films, TV shows, and user-generated video content—all *to go*.

The various media and screening instruments that smartphones' technical apparatus accommodates advance them to a hypermedia system of spectatorship (see Snickars & Vonderau, 2012). Video players and other multimedia applications are substantialized in two ways: on the one hand, as physical objects that the user controls through mechanisms afforded by the device. On the other, these applications remediate other instruments and become *referential objects*. Inspired by "the pioneering work" of C. K. Ogden and I. A. Richards (1923), I recognize a referential object through the analogy to its referent. Unlike Nanna Verhoeff's (2009, 2012) *theoretical console* that entails a conceptual and contextual definition of digital tools,[7] referential objects (here, mobile applications, for instance) provide modal and representational links between earlier media remediated by a graphic computer interface and the pragmatic and material specifications of the container device. Thus, I argue that smartphone applications are ambiguous media, incorporating digitally visualized remediated objects and media practices. When a video player is running on the smartphone's screen, the visualized representation and command mechanisms remediate the frameless design of the cinema screen as well

[7] Digital tools, according to Verhoeff, are analyzed and used through the specific functions they enable. The concept is applied in Chapter Two to discuss user involvement and interactive spectatorship.

as the icon kit and button interaction of analogue and digital projectors and players. Such semiotic and haptic references function as temporal and spatial signals to trigger associations and inform usability, in other words, to synchronize operative mechanisms in the physical and virtual spheres.

As introduced earlier, it is not only technical tools that play a part in remediation processes; filmic instruments, such as audiovisual composition, editing, or other non-diegetic features, such as music, also appear as references in smartphone spectatorship. But whereas buttons, icons, screens, and the like inform mostly operative aspects (e.g., how to start a movie or adjust the screen's position), storytelling methods inform cognitive processing in the pursuit of answers to, for instance, whether the protagonist will succeed in his goal at the end of the story. In terms of accessing content, the depicted story, this means that diegetic and non-diegetic elements (for instance, the protagonist's success; musical themes and length of takes) serve as referential objects in the same manner as buttons or other visual, sonic, or haptic indicators: they remediate conceptual formulas that help comprehension, but, according to the double logic of remediation, become transparent in favor of narrative content.

The infinite loop of remediation within a referential object, for instance, the remediation of physical and virtual objects on a smartphone, generates a theoretical *mise en abyme*, or Droste effect of recursive references. These references are repeated and recreated in an infinite number of reproductions at each and every occurrence of user intervention—similar to the effect of two facing mirrors. Successive remediations between the narrative and the spectator's mind through narration, non-diegetic elements, the screening apparatus, stimulation, and mental processes generate new functions and sensations each time. Casetti (2011a) embarks from a similar mechanical structure when he studies cinema's relocation and the broad range of spectatorial economies in Atom Egoyan's short film, *Artaud Double Bill* (part of the anthology, *Chacun son cinéma: Une déclaration d'amour au grand écran* [*To Each His Own Cinema*], 2007). Casetti's account chiefly concerns illustrating the multiple faces of cinematic spectatorship, which only partially applies to this analysis. By further exploring Egoyan's film, I intend to draw attention to the layers of remediation that connect cinema and smartphones, scrutinize the processes involved in spectatorial behavior and narrative comprehension.

In *Artaud Double Bill* (Figure 3), two friends—or alter-egos, if you like—accidentally end up in two different screening rooms watching two different films: Anna is watching Jean-Luc Godard's *Vivre sa vie*, while Nicole is sitting in another theater viewing *The Adjuster* by Egoyan. *Vivre sa vie* opens with the main

character, Nana, watching Carl Dreyer's silent classic, *La passion de Jeanne d'Arc*. Thus, Anna watches Nana watching *Jeanne d'Arc* on one screen. Simultaneously, she is recording the film to involve Nicole, who live streams it on her mobile phone: Nicole watches *La passion de Jeanne d'Arc* and *The Adjuster* on two separate screens at once.

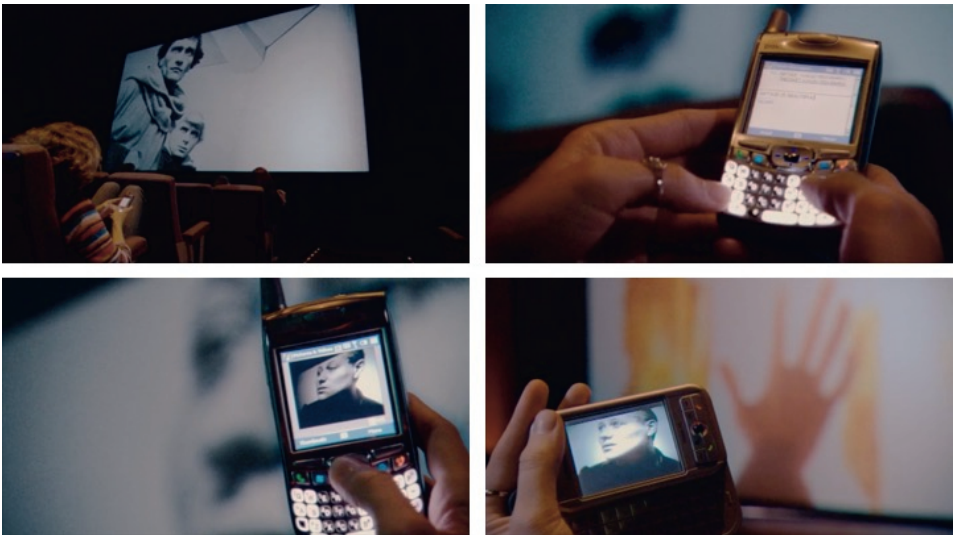


Figure 3. Stills from *Artaud Double Bill*. Screenshots from film.

Casetti's inquiry epitomizes the ways screening media fuse, discussing the effect of this fusion on viewing as a social phenomenon. However, Egoyan's film holds some additional references for my analysis, namely, the role of media, mediation, and embodiment. Anna, by recording and streaming *Vivre sa vie* through her mobile phone, produces another, definitely irreproducible version of both Godard's and Dreyer's films using her own body, intentions, and technical tools. She records her personal, momentary impression of the film: by involving her physical connection with the mobile device, she augments both the dynamics of her hand and arms and the camera's image-capturing quality to the footage. On the other end, Nicole perceives Jeanne d'Arc (or, to be precise, Antonin Artaud, the actor playing Father Massieu) remediated through the layers of Dreyer, Godard, her friend's body and intentions, the recorder device, and finally her own posture and the technical specifications of her phone, which is captured through Atom Egoyan's camera for spectators.

Besides the film's visual scope that merges multiple layers of representation, another angle of concern here are those of action, perception, and the spheres of

social connectedness. Anna's and Nicole's plots signify the archetypical consumer scenario of being simultaneously exposed to a series of mediated and unmediated sensory stimuli and accommodating pertinent behavioral codes. In Anna's scenario, the act of recording contrasts proactive use of a multimedia tool with the stationariness of cinema spectatorship. She involves her knowledge of the phone's recording mechanisms and perspectives. Nicole's spectatorial behavior indicates the conjunction of the two screens, signifying divided attention between the predetermined viewing mode (the cinema screen) and an embodied, and therefore yielding, relationship with the phone's screen. In each case, the two sets of mediated stimuli complement, but at the same time separate, each other, primarily due to the variance of consumption patterns and attention. Consumption extends the physiology of viewing (i.e., the phenomenological experience) and spectatorship (watching a certain type of content using a certain type of apparatus). In my reading, it attains the assemblage of psychological, cultural, technological, and economic variables, which conclusively define the spectatorial manners, behavior, and individual interpretations of a film narration. This means that the different layers of narratives that Anna, Nicole, or Egoyan's spectators perceive are aggregated into distinct practical usability formulas and cognitive processes. Anna is engaged in processing the intentions and motivations of Godard's character, and through her, the momenta of Jeanne d'Arc's life while she completes another action, recording the film for her friend who is sitting in another screening room. For pursuing this goal, Anna makes use of her knowledge of mobile phones, recording, camera positioning, and so on. In the other room, Nicole pays attention to two different movies and narrative lines, one about the sufferings of the Maid of Orléans and the other about the unorthodox methods of an insurance adjuster, through her knowledge of cinema behavior and mobile phones.

During movie viewing illustrated by Anna's and Nicole's cases, two separate cognitive processes are in play, which can be divided into two semiotic units: movie narration and comprehension, on the one hand, and mediative and spectatorial formulas, such as understanding operational mechanisms, on the other. Torben Grodal (1997) schematizes these dimensions in his theory of the *metaframe* and describes how extra-diegetic and even extra-filmic elements have significant impact on emotional and cognitive connections to a movie or movie sequence. This suggests that spectatorial experience is highly influenced by a wide range of mental and material filters from genre expectations to screen position. Correspondingly, we can establish that the two dimensions of spectatorship are the presentation and comprehension of a narration and the

corresponding technological and contextual instruments, which latter components are superimposed during commercial circulation.

Each separate act of film viewing typifies an individual metaframe: the narrative and its audiovisual composition, the size and position of the screen or projector, the image quality, the space and its participants, as well as the spectator; her cultural background, knowledge, associations, attention, behavior, and reactions are all factors that contribute to creating a unique item. In *Artaud Double Bill*, the protagonists' dual scenarios illustrate just such a complexity. Anna and Nicole are not merely involved in the cinematic and mobile spheres of spectatorship; they engage in processes of mobile filmmaking, distribution, redistribution, and online streaming. Moreover, they encounter alterations of film format, sensory modalities, and quality, while experiencing a variety of viewing modes.

Mental associations attach mediated and unmediated information (or multiple sets of mediated information), but attention disconnects them. Nicole fails to concentrate all her attention either on the cinema screen or on the mobile phone; her focus travels back and forth between the two simultaneously present screens. In *Visible Fictions: Cinema, Television, Video*, a work that intended to strengthen the link between film and media studies, John Ellis (1982) asserts that the key difference between cinema and "new" media (he uses television as the counterpoint in his comparison) lies in the length of attentive states. While cinematic spectatorship is defined by the prolonged gaze, domestic (personal) media platforms induce shorter-lasting glances. Ellis analyzes both cinema and television as institutionalized media and applies a systematic framework to link viewers and technology. His comparison of spectatorship and consumption patterns enucleate, first, the disparities in immersive qualities and screen-centeredness between the two groups of media and, second, the factors that prioritize the spectator's freedom in what she consumes.

Casetti (2011a), on the contrary, carefully omits defining attention-shifting as something inherently distracting. Instead, he borrows Mariagrazia Fanchi's (2005) approach and argues that partitioning gaze between simultaneously presented visual stimuli marks a multi-centered, attentive, but dividing mode of watching. The logics behind Casetti's theory nevertheless correspond with the main inquiry of this dissertation, namely, that narrative comprehension is based on the viewer forming context-specific and context-unspecific information clusters which connect or disconnect stimuli from different sources. Yet, his train of thought is somewhat incomplete, as it disregards the peculiarities and priorities across different varieties of stimulation, which I will explore later.

Divided attention and its analogy with any everyday stimuli are still fittingly illustrated in *Artaud Double Bill* through Anna's emotional reactions, as Casetti pertinently notes: her attention makes momentary shifts, which hinders her from achieving an immersive state. Anna is obsessed with Antonin Artaud's aesthetically pleasing image in Dreyer's film, but she fails to perceive Jeanne d'Arc's emotional drama and Nana's sorrow. To rephrase it, Anna responds to the artifact elements, the visual representation and composition of the film without immersing herself in the fictional world.

When developing the characteristics of artifact and fictional-world emotions, Ed Tan (1996) outlined a model of films' emotive strategies that assigns different emotional reactions to different forms (levels) of mediation and immersion. Fictional world emotions, he claims, are either shorter reactions catalyzed by the spectator's expectations regarding the outcome of narrative events (phasic emotions) or cover a longer, more complex segment of the story (tonic emotions). These reactions require an increased feeling of presence within the boundaries of the diegesis. Tan also suggests that, precisely for this reason, when the spectator's attention is directed toward a narrative and she is immersed into it, emotions primarily originate from empathy, surprise, fulfillment, and similar connections. This likely suppresses emotions built upon the artistic presentation; a fact that he credits to the narrative style of fiction films and the diegetic effect.[8] Bordwell, Thompson, and Staiger (1985) claim that identifying with the events on screen—given that the mode of observation (e.g., camera angle) resembles natural perspectives—advances the development of presence feeling and disguises the viewer's surrounding physical space. Therefore, if the level of immersion decreases, for instance, because the narration is not engaging or one's attention is divided, film as an artifact or its visual composition, music, representation, or other non-fictional features are more likely to occupy the center of focus. Casetti's observation holds true with regard to attributing shorter periods of focused attention to the sensory presence of a broader range of stimulation, as in Anna's case: instead of being immersed in Nana's or Jeanne d'Arc's storylines, she observes narration from a distant and temporary position. Egoyan highlights this distracted state with Anna's text-message remarks on the actor's look and her lack of empathy for or emotional connection to the characters in *La Passion de Jeanne d'Arc* and *Vivre sa vie*. Egoyan denies his characters' and viewers' immersion; he constantly reminds them that they are part of a hyper-

[8] The term diegetic effect refers to the spectator's illusion of being present in the film's fictional space as if it were her own environment. For a broader discussion, see, among others, Burch (1979), and Bordwell (1985), as well as Chapter Three.

mediated system of narratives. In addition, he maneuvers between divergent visual and auditory stimuli, surround effects, vividness, and multimodality, a question that will be addressed in greater detail below.

As the logic behind the theory of remediation (Bolter & Grusin, 1999) implies, cinema has played an extensive role in shaping the cultural domains of movie consumption. The audience, hand in hand with technological improvements, has characterized the content and the spatiotemporal spheres of spectatorship. Demands for more realistic imagery developed during the emergence of digital technology as did the desire for an increasingly immersive experience, when movie watching became less typically defined by the physical properties and behavioral protocols of institutions. On-demand modes for watching and the domestication of screens that connected cinema to personal devices have effectuated an increasing spectrum for both content and behavior. The bidirectional nature of cinema's remediation on new media platforms has elicited medium-specific visual representation and strategies for content making and distribution.

Remediation and Premediation

The increasing number, quality, and availability of screening platforms that remediate cinema and amalgamate it with newly developed technologies reveal a persistent trend in reinterpreting cinematic practices. Richard Grusin (2016; see also Grusin, 2004) refers to this tendency as the *cinema of interactions*, and regards film consumption and experience as part of "a hybrid network of media forms and practices" (Grusin, 2016, p. 66). Whereas remediation indicates the reinterpretation of objects, practices, and their mediated appearance, cinema of interactions implies another ontology, namely, premediation. Premediation "imagines an interactive spectator in a domestic or other social space rather than an immobilized spectator in the darkened dream-space of apparatus or gaze theory."

Grusin's understanding implies that mediation can be interpreted as occurring along two trajectories: backward, as a reference—in the context of remediation, and forward, as an opportunity—in premediation. The forward dynamics of mediation denotes no specific mediating agent (media tool). Thus, the mechanisms of premediation lack a direct connection between old and new media, which are deputized by *potential* paths of development. What Grusin

endorses is the prospect for technological, social, and behavioral evolution, which, in this case, places a direct link between cinema and newer movie screens such as smartphones. In other words, premediation implies a role in which existing media provide frameworks for understanding, preparing for, and contextualizing future media and practices. This means that usability schemes and protocols related to cinema premediate the circumstances in which flat surfaces connected to screening apparatuses become film screens, which, subsequently, remediate cinema's behavioral and physiological conventions. It is precisely this manner of expanding the means of practice by both medium and consumer that facilitates the birth of the cinema of interactions.

The analogy of Grusin's cinema of interactions to Tom Gunning's (1986) *cinema of attractions* is no coincidence. Gunning argues that the reason early films were game changers lies in the aesthetics of wonder and the contradiction between sensation and expectation—mechanisms that play important roles in new media too. Also, the astonishment of early cinema that arose from the screen and the sensory illusions the spatial arrangement fostered was a first step in the evolution of both cinematic perception and behavior. Cinema's cultural role requires some clarification here, however. According to Grusin (2016), a reading of cinema as something that directly mediates the physical world to provide a peculiar frame for sensory experience is incomplete. It must be supplemented with the remediation of various other media: photography, theater, sound recording, and synchronized presentations of sound and image. However, Grusin's theory also implies that new media screens not only remediate cinematic practices and protocols, but in fact premediate spectatorial behavior on other screens and in other spaces as well. Although Casetti's point of departure is different, his above-discussed idea of cinema's relocation (Casetti, 2008b, 2011b, 2012) reaches a similar conclusion: the post-cinematic spectator endeavors to recreate the cinematic spectacle by using its behavioral references.

On a smartphone's interface, the premediative role of cinema manifests itself in virtual hypermediacy and a system of fluctuating artistic, social, and technological practices: smartphones incorporate and synchronize different media and allow for great freedom in combining different modes of usage. Similar to cinema or other screening platforms, movie-watching on portable devices paves a bidirectional logical path between the audiovisual stimulus and the spectator through different media references which define the modes of spectatorship. By presenting audiovisual footage with the help of video player applications, smartphones exhibit the convergence of several distinct layers:

(1) movies (storytelling practices, cultural references, linguistic, and semantic elements),

(2) the cinematic apparatus (the parallel presentation of image and sound through one or more devices),

(3) a wide range of video players (and their options, as well as audiovisual representations),

(4) smartphones (platforms for various functions),

(5) mobile devices (that are portable and mobile), and finally

(6) the space and time of spectatorship.

For sharing viewing experiences, a smartphone also remediates and visualizes applications for communicating, surfing the web, and distributing content in various forms. Moreover, video players represent a complex system of options that enable interacting with the screening (e.g., pausing or rewinding the video) and also provide and visually indicate modes of interaction by remediating projecting mechanisms from digital and analogue screening apparatuses.

I argue that the smartphone as a medium for movie spectatorship emerged from cinematic conventions just as much as it did from the physical and technological specifications of handheld devices. Such conditions, technical and cultural alike, operate in terms of mnemonic and projective logics. If smartphones are regarded as movie screens, spectatorial conventions are informed by those from cinema, telecommunication, and portable devices, meaning that the roots of moving-image consumption lie in technology, user behavior, and the environment. Again, this notion concludes in a twofold reading. First, the material and social framework of cinema and mobile media envisages the modes of smartphone spectatorship (premediation), and second, these modes feed upon the practices of the two (remediation). This projective–mnemonic parallelism that Grusin and Bolter (Bolter & Grusin, 1999; Grusin, 2016) also exhibit entails intercommunication between new and old media, by which both technological advancements and forms of consumption tend to reach a transparent framework for perceiving content.

In psychological terms, I regard perception as a direct, immediate link between stimulus and spectator without any technical mediation that instead relies on sensory, linguistic, semantic, and narrative schemata. The immediate connection is analogous to the sensation of sensory immediacy, as explained above, in which visual and sonic indices contribute to the spectator's understanding of objects' and events' temporal and spatial system. Yet, while movies are often referred to as the closest medium to physical reality, film as an art

form (dependent on formal, temporal, and spatial discontinuities) has followed tracks too ambiguous to be described in terms of an indexical connection between representation and referent (see Kracauer, 1968). This approach, which opposes the indexical form of representation to mediation, punctuates a significant methodological system: that of sensory connection, emotional engagement, media practices, and media-specific modes of consumption.

Following a similar argument, Michael Grabowski (2015) begins his work on neuromediation by proposing a comparative analysis of mediation and user behavior as a potential method for investigating communication and communicating actors. His solution stands in accordance with the modal logic of immediacy explained above. By choosing to scrutinize mediated communication through presupposed roles and behavior, Grabowski pinpoints the importance of a conditional framework and sketches the link between technological and cultural advancements. Whereas Grusin and Bolter approach mediation through the social means of hypermediacy (i.e., the logical connection between social processes and representation that presumes an interplay between consumers, media, and content), Grabowski sets out to theoretically connect the human body to communication devices. Such an ecological model, however, still cannot omit the question of remediation and its role in user behaviors. Therefore, I will revisit a compensative, well-balanced union of the two in Chapter Two in my analysis of the spectators' role in embodied watching. But at this point a framework must be defined that aids the comparison of cinema and smartphones in the upcoming section: one fueled by a cognitive, or neuromediative, premise, but still explores the realms of media-specific consumption and behavior. A feasible method proceeds in the direction of premediation and remediation reaching their further vanishing point in the form of behavioral and consumption references. Thus, I hypothesize that a collection of previous media practices manifests itself in the social, cultural, and psychological processes of smartphone spectatorship. The smartphone as a movie screen can only be understood as the extension of the human body and nervous system, if we ground it within a hypermediated system of media references and assigned practices.

The above deduction by no means attempts to question the codependency of perception and apparatus. I resolve such a seeming controversy between the comparison of smartphones (or any post-cinematic media) and cinema by returning to the dual question of reality or realness. This specifically implies an examination of whether film is "real" in the sense that it is a medium for representing physical reality or whether it is an artifact and a mediated version of

reality that exists only in the context of a screening apparatus and related technological and social conventions. The "realness" of filmic representation lies in its direct phenomenological connection with its viewer, on the one hand, and in the acknowledgment of the act of mediation, on the other. When describing the illusion of reality a medium produces, E. H. Gombrich (2000) distinctly denies fidelity a role and instead highlights the way the observer (spectator) uses her own experience of actual objects or concepts. Immediacy, according to this argument, implies a perceptual connection rather than a physical or technological one—similar to Grabowski's theory. However, Gombrich elaborates, immediacy must acknowledge physical space, since an immersive state can only be achieved by either closing the distance between the observer and the represented objects or by making the space idle. Distancing, therefore, grants a material and cognitive understanding of the physical interspace between the spectator and the screen, as well as selective attention to the respective stimuli of space and apparatus. This augmentation of the human factor also connects to the strategies Bolter and Grusin (1999) suggest for reaching an embodied state as the object of mediative experience. The Alberti's window metaphor they use to reflect on televised, cinematic, and virtual reality experiences similarly points to eliminating the distance between stimulus and perceiver, where different types of represented objects and their pervasiveness occupy the senses.

Correspondence between the spectator's bodily presence and sensory environment implies that the represented narrative information (i.e., the film plot) is incorporated into the space of spectatorship. Therefore, immersion originates from the illusion of nonmediation; in other words, the sensation that the depicted diegetic events unfold in the environment surrounding the viewer. This concurs with Gombrich's idea of spatial composition, as well as with the spectator's position as a point of reference for which Bolter, Grusin, and Grabowski all advocate. Grabowski's ecological—more precisely, media-ecological—view provides an even more concrete tool for connecting the dots between the relative components of remediation and the viewers' position: his reading of the ecological grounds for mediation draws attention to neural processes as a point of departure for describing how individual media engage perception. Based on a complex overview of communication methods, this model highlights the importance of premediation and remediation by providing a framework for observing the balance of the technological, social, and perceptual mechanics of spectatorship.

Introducing the concept of media ecology in education, Neil Postman (1970, 1979) inserted the term into the then ongoing discourse of media, technology,

and their effects on human communication. Postman proposes a holistic perspective that includes technology and consumption into his model of mediated communication; moreover, he cites information transmission as the responsible factor in perception and behavior. Using Postman's ecological perspective as supplementation (or background) to Grabowski's spectator- and cognition-centric view leads me to the following interpretation: regarding media as self-standing cultural products, the reinforcement of perceptual and behavioral paradigms configures neural pathways through which mediated stimuli are sensed irrespective of the mediative context. Grabowski elaborates, "the associations between signifiers and various referents can be understood as interconnected neural networks that are bounded by evolutionary genetics but are formed as new experiences create new connections between already established patterns" (p. 10). This idea produces two equally important results. First, it affirms the theoretical foundations of how new media spectatorship revisits cinematic practices. This highlights the link between stimulation and behavior, namely, that mediated stimuli generate associations to implement social and behavioral patterns. Second, reversing the equation, behavior is connected to perceptual and cognitive strategies, which serve as basis for testing whether practices and attitudes toward specific media would interfere with narrative comprehension. Building on Grodal's (1997) theory concerning the functional and neural processes involved in movie spectatorship,[9] Grabowski's theory takes the same direction. He explains, "a Kindle is like a book but also like a computer. An iPad can be like a magazine, but only if one has experienced a magazine" (2015, p. 10). Similarly, a smartphone can become a movie screen if the user has corresponding knowledge of movie screens and video players. Grabowski adds, "as experiences are repeated and reinforced within a cultural structure, networks are strengthened and begin to bypass the prefrontal cortex,[10] building habits and emotions of which we become unaware" (2015, p. 10). In effect, this means that the pragmatic argument behind the logics of remediation and premediation is tightly connected to an inspection of neurological habituation or adaptation processes, as I will explain further in Chapter Three. Such an approach also implies that, by establishing a specific type of sensory and behavioral reference network, each medium contributes to a

[9] Grabowski refers to Grodal's PECMA flow model, that is the system of perception, emotion, cognition and motor activity. For a comprehensive definition, see Chapter Three and Grodal (1997).

[10] The prefrontal cortex is the brain area responsible for decision-making. It orchestrates expectations, hypotheses, executive functions, and goal-oriented behavior (Goldman-Rakic, 1996; Shimamura, 2000).

consumer's sensation of immediacy. Here, immediacy is understood as the product of the cultural references a medium applies and the user's contextual understanding of the sensory information related to those references. This means that media references can only result in the sensation of nonmediation, if the user is aware of both referent and reference object and if she regards both of them in terms of functionality. For instance, an e-book read on a Kindle can only be perceived as a literary text if the reader knows the methods books or other corresponding media use for mediating a story in writing and perceives the sensory information as a text instead of as a visual representation on the screen of an electronic tool. Similarly, in order for a spectator to contextualize audiovisual information as a movie narration, it is necessary to be familiar with storytelling as well as the mediating methods of storytelling and screening apparatuses.

Since this approach treats a single medium as an integral system of stimuli and behavioral patterns (as do remediation and premediation), it produces a somewhat generalized account of the sensory and cultural aspects of media consumption. Grabowski admits this, noting the need for disaggregating factors such as environment, social background, protocols, and emotional engagement from the actual moving-image stimuli. Grodal's (1997) previously mentioned metaframe thesis establishes a methodological direction for a similarly systematic definition of the cinematic (filmic) phenomenon. The most significant element in Grodal's theory is the functional presence of the screen, which localizes, frames, and defines the spectator's relation to (namely, proximity to and dissociation from) the movie. Moreover, the central role of the screen establishes the framework for attention, which, in this regard, creates a unified, schematic feature that unites the audience. Using the screen's fundamental omnipresence (i.e., its palpable presence in the space of spectatorship) as a point of departure, let me reframe Grodal's theory and highlight elements of his explanation in order to disaggregate the variables of spectatorship. These variables are the following: space as a "pointer-mechanism" (Grodal, 1997, p. 210), the temporal dimensions of attention, emotional and cognitive filters (which apply to both individual and social behaviors), and the spectator's intervention and enactment. Thus, the elements associated with spectatorship can be rendered into four main categories to serve as a theoretical skeleton for locating and establishing the links between cinema and smartphones and for identifying the remediative role of smartphones in cinematic and the premediative role of cinema in smartphone film consumption. These four categories concern the *screen*, viewing *space* and *time*, as well as the *behavioral protocols*, including

copresence and social interaction. The spatial and temporal frameworks for screening media define the features that disembody the spectator's diegetic engagement from physical presence and, thus, reveals inner dynamics from the broad realm of spectatorship toward the filmic content. The screen and enactment (i.e., the spectator's embodied relationship with the specific screen and her surroundings) follow a trajectory from the movie toward medium-specific attributes, meaning that mediated stimuli initiate the framework of consumption. Hence, I propose that a bidirectional link defines the mediative relationship between cinema and smartphones. This link must be observed through a lens of screen alignment, through the social and sensory environments, audience dynamics and, finally, through the historical, material, and cognitive discourse of the embedded, but selective attention to the movie stimulus.

From Cinema to Smartphones

Following the dialectical argument of remediation, movie screens refer to earlier media practices for establishing their own respective reference systems and to achieve immediacy. Single- and multimedia appliances from television sets to computer interfaces hold references to the cinema screen, yet embed an array of other sensory references to support usability. But as smartphones' hypermediacy and multimodality—that condense the act of viewing, creating, and interacting—foreshadow, the notion of immediacy, or media transparency, requires a twofold approach. This approach combines perceptual and referential domains. The perception-based understanding of immediacy (that, for instance, an e-book is perceived as a literary text instead of pixels visualized on a screen) correlates in many ways with what Bolter and Grusin (1999) describe as *transparent* immediacy in the theoretical section of *Remediation*. However, transparency neglects physiology, in other words, the medium of sensations. Even though they acknowledge the perceptual connotation, Bolter and Grusin's statement maneuvers along another path: they base immediacy on the immersive quality of a medium, which demands its continuous, reflexive presence. This notion fails to hold true in some cases, for example, virtual reality platforms, which can provide immersive visual and sonic stimulation, but often involve bulky equipment or unrealistic sensory representations. I intend to resolve this paradoxical connection between the sensory and mediative compo-

nents of immediacy by replacing Bolter and Grusin's focal point on the medium and its attempts at achieving immediacy with the spectator's perception. A perceptual approach recognizes the *sensation* of being present in a diegetic (mediated) environment and focuses on meaning-making processes, which is inevitable when discussing the cognitive processes involved in movie watching.

From another angle, immediacy can also be mapped in epistemological terms to describe the mechanisms that play roles in recognizing and reacting to objects. Such passage is inspired by J. J. Gibson's (1979) ecological approach to sensory perception, more specifically, his proposal for observing potential modes of interaction between objects and users.[11] Usability, as defined in Gibson's affordance theory, is based on an object's capacities for interaction, which signifies why explaining immediacy in terms of transparency is problematic, or at least incomplete: a conceptual definition of transparency applied to a medium's immersive quality places remediation into a media-archaeological context without crediting a cultural reading, according to which representation originates from a medium's functions and capacities for interaction. At this point, it is important to highlight and dissociate the material specifications of the smartphone as a container device (its size, mobility, touchscreen, etc.) from the pragmatic contextualization of the interface (e.g., sensory representation of applications). This is crucial because the two signify different trajectories of remediation, and while both contribute to spectatorship (discussed further in Chapter Two), they achieve immediacy using disparate strategies. The smartphone in the form of a screen and audio-transmitting device feels immediate due to the viewer's conceptualization of sensory modalities, and thus gains a status similar to the cinematic apparatus, while as a graphical interface, it gains immediacy by familiar depictions and usability formulas, such as clickable icons.

Fusing the concepts of relocating cinema and remediation and premediation into an operative method, that is, defining media references by means of material conditions and aesthetic and cultural practices, allows for an overview of what makes cinema cinematic and what cinematic qualities smartphones have when they are used for watching movies or video content. More specifically, such a method supports exploring which exact factors pertain to cinema's efficiency to transmit narrative information in a social setting, and how smartphones attempt to incorporate such competences. It seems essential to contrast the sensory qualities and the collective and individual settings of spectatorship across cinema and smartphones, which attributes echo with the phrases of spontaneity, freedom, integrity, embodiment, quality, and isolation;

[11] For a more detailed discussion of Gibson's theory of affordances in usability, see Chapter Two.

none of which has positive or negative connotations. Yet, traditions of film culture show a seemingly conservative and essentialist promotion of cinema[12] prompted by architecture and behavioral protocols, and this stand for cinema's superiority disobeys the venture of multimedia in gaining status in film spectatorship.

When introducing cinema's abilities to execute collective experiences, Julian Hanich (2018) elucidates that, even though collectivity was constructed around a static form of presence (derived from the sturdiness of the screening room design), the most substantial factor that has contributed to the cinematic experience is, in fact, the perpetual correspondence of facilitating and inhibiting individual reflections in a space of communal copresence. The cinema audience, thus, is a composite of individuals, but also a collective mass propelled by social dynamics; Hanich's argument presupposes a physically present collective.

The basic theses of cognitive film theory are founded upon the notion that the spectator constructs the narrative and contextualizes it using mental representations of sensory modalities (Branigan, 1992). Hanich, too, acknowledges this and treats it as a cue for mapping and sketching social impulses under the term *audience effect*. Borrowing the concept from social psychology, he allows for a parallel reading, according to which a cinema audience is a community that forms less because of physical copresence, but rather because a behavioral framework in the screening room happens to confine these people into one space.

Quite obviously, the smartphone "audience" is less likely to form the same spatially and temporally circumscribed collective of spectators and thus, the sensation of collectiveness is more attached to mediated channels than phenomenological, sensory presence. The "audience" in the case of smartphones is a plastic body of individuals that operates through online presence and interacts on mechanically constructed platforms, such as social media, blogs, and text messages, with an extended capacity for movie consumption and a flexible spatial and temporal frame. Thus, besides the question of sensory richness and fidelity, which are mainly consequences of screen size and audio quality, I treat these factors as the key components for defining the resemblance and differences between cinema and smartphones. The spectator's freedom in choosing, streaming, playing the film or video footage, and interacting with it individualizes each encounter with moving images to a greater extent than in cinema, while online channels (such as social

[12] See the discussion of the new trends of content production and smartphone viewing in the last section of Chapter Two.

media, messengers, or streaming sites) frame and converge individual consumption patterns. An example that captures these mechanisms features a smartphone spectator who accesses a movie or video through a website or mobile application and watches it while in motion (e.g., commuting) and stretches the temporal frame of spectatorship in accordance to the physical activity of traveling and the chosen pace and time interval. Before, during, and/or after the movie, she may engage in online communication to rate or discuss her experience and share the clip itself or her personal reflections of it. The smartphone spectator in this scenario connects with other (potential or actual) viewers to share her experience, thereby forming a loosely coherent community, with social dynamics similar to Hanich's cinematic collective.

However, the dynamics of a physically copresent collective cannot be solely understood in social psychological terms, as Hanich's theory suggests. Besides the synchrony of behavioral schemes (such as sitting still or expressing emotions relevant to the movie), audiovisual stimulation also generates real-time psychological and emotional integrity across the present audience. Psychological and emotional integrity provides the sensation of cinema: these are the elements that can provoke public laughter even if a scene is only funny for some, that impede loud weeping for fear of shame in front of others, or provide comfort during a scary moment. Comfort and discomfort are equally part of the effect of collective spectatorship (see Harris & Cook, 2010; Hanich, 2018). In smartphone spectatorship, these processes differ, first, in their spheres of manifestation, that is, in the form of off-line versus online interaction platforms, and second, in their spheres of stimulation—the private or public presence of the screen.

The fundamental function of copresence has been widely praised as the essence of enjoyable movie watching (even in scholarly circles, see Bordwell & Thompson, 1979/2001), and cinemas indeed maintain the loop of collective viewing experience not only in how they screen, but also in how they promote movies. Collectivity, however, sacrifices the comfort of personalized spectatorship, where post-cinematic and mobile media prevail. What propels trends in movie and audiovisual culture at present is marked less by commercial interest than by technological advancements, or, as Bernard Stiegler (2009) states in *The Carnival of the New Screen*, the role of cultural industries is replaced by that of technology. Accordingly, participatory film culture (to which I devote a section in Chapter Two) operates by the very presence of ubiquitous devices like smartphones and attempts to demolish the rigidity and one-sidedness of cinematic spectatorship. While applauding online video's admission to the film

industry, Stiegler acknowledges the spectator's increasing freedom and compares it to how the French New Wave reformed cinema.

I have no intention of engaging deeply with the possible moral or political consequences of a tailor-made film culture. Yet, in line with the logic behind remediating film screens and Stiegler's description of the plethora of audiovisual content, two summarizing points must precede scrutiny of smartphones' cinematic roots. By integrating moving-image content, culturally accepted indicators, the social dynamics of audiences, and the viewer's subjective pursuit of a pleasurable and effective viewing experience, smartphones operate with proliferating media references, which are actively involved in connecting users with familiar communication methods (see also Grusin, 2009). This means that icons and other sensory indices of areas of interaction on the smartphone's screen, as well as online and off-line interactions in a social sphere support immediacy by applying practices borrowed from other media. In agreement with the preceding one, the second statement reflects on immediacy as a link to connect technological development (allowing for such a large and complex blend of functions enclosed into a portable device) to the user's embodied presence. Here, the sensory impressions stemming from the spatial configuration of the body, the screen, and other social actors enable, even advocate for, real-time adjustments.

The union of the human and cultural components of media and the endeavor to achieve immediacy signifies an instrument that aids observation of the eventual differences and similarities between screens, spatial and temporal structures, and social constructs. As proposed earlier while sketching the different directions of mediation, drawing a parallel between cinema and smartphone spectatorship while avoiding the verification of a particular cultural status is crucial for two reasons: first, to determine the commodities of movie consumption, based both on pre-programmed schemes and consumer freedom; and second, to reflect upon the role of mediated references in forming and developing the cultural directions of new media. Cinematic references, be they behavioral, narrative, or otherwise, create a theoretical thread toward post-cinematic film screens. Cinema's presence in smartphone spectatorship is likely the responsible party for framing the technological, economic, social, psychological, and bodily components of film consumption and film cognition, and hence, necessary for defining the legacy of mobile film screens.

In the following pages, I apply a close view of cinematic references that appear in smartphone spectatorship and the factors that induce immediacy. My aim is to address the question of technology and social practices across the two

platforms and to describe which parameters are adopted from cinema and which typify mobile screens in particular. A bidirectional flux between the two platforms is revealed in connection to material attributes (such as screening space, the screen itself, and sound system); behavioral protocols and the spheres and features of social interactions; and the general nature of the physical, cultural, and perceptual relationships that form between the spectator and each of the two screening platforms. Firstly, the comparison covers the respective screen's size, surface, and location; then I define the extent of user involvement to schematize patterns of sensory engagement with regard to both vision and sound. Secondly, I turn from a technological approach to a cultural discourse to examine the persistence and legacy of the elements created by cinematic viewing protocols: screening time, social space, behavioral norms, and the sensation of a diegetic presence.

Screen

The most evident, and perhaps the most substantial difference between cinema and mobile devices, is the size of the screening surface. This includes not only image size, but also its position, the screen's relative mobility or stationariness in the space of spectatorship, as well as the proprioceptive dimensions that define the spectator's body in relation to the screen. Even though an average cinema room houses a screen 15–20 meters wide and an average smartphone screen is around a hundred times smaller, the angular size of the cinema screen (the visual angle subtending the viewer's eyes) is barely more than twice of that of the smartphone's if the device is held close to one's eyes.[13]

A bigger screen does not necessarily enhance the feeling of presence: although its surround effect is guaranteed, it can also make distracting image flaws increasingly apparent. Moreover, such a blunt conclusion would neglect variety across the other mentioned factors, which will be discussed further along this section. What, however, is pivotal is the physical position of the screen in relation to the viewer's body. Proprioception is the sense that orients the body in relation to other bodies or objects using proprioceptors that are located in muscles, joints, and tendons (Colman, 2015). Whereas the relative distance

[13] As measured during the behavioral experiment presented in Chapter Four, with an average viewing distance at around 45 centimeters, the horizontal visual angle approximates 20 degrees, while a cinemagoer in one of the most favorable middle seats watches the movie from an angle of around 50 degrees (Talwalkar, 2011).

between the viewer's sense organs and the screen can equalize the visual angle, the proprioceptive element projects distinct rates of bodily involvement: in cinema, a viewer's body is small relative to the screen, while feels large compared to a smartphone held in one's hands. This dimension implies two aspects, a biological and a social one. Biologically, proprioception is responsible for adjusting the body's position in order to avoid collision, for instance, when a vehicle is seen approaching or when an object is flying toward the observer. When seen on a screen with a larger relative size than the observer's body, these objects might evoke a flight response manifested as ducking or cocking one's head. Being aware of kinesthetic power over the mobile screen (i.e., that its position can be adjusted by muscle movements), these reactions are less likely in the smartphone spectator's case. There, a more feasible reaction would be moving the screen farther away from the body. Awareness of control over the relative image size also affects immersion (see in Chapter Three) and changes in visual angle affect gaze strategies too.

Screen size—in terms of the extent to which the visual field is covered—can affect attention to visual cues. On a large cinema screen, the viewer performs more and longer saccades (rapid eye movements between two fixation points) in order to effectively explore the details on screen. Since the brain only provides a sharp picture at the points where the eye fixates, saccadic movements are crucial for mapping the visual elements that attract attention. An average smartphone screen from the late 2010s, however, is barely more than 10 centimeters in length diagonally; ergo when focusing on the center, the spectator sees much more of the screen in peripheral vision than she does a cinema screen and, as a result, the amplitude of saccadic eye movements decreases.[14] Following an eye-tracking demonstration, Tim Smith (2014) drew similar conclusions. As part of a BBC Radio show, Smith compared the eye movements of the two hosts watching the same animated movie on screens with different angular size: one of them on an IMAX screen, the other on a laptop. Although the amount of research on the subject is still too limited to draw definite conclusions, Smith's results show that the IMAX spectator's gaze shifted on a wider scale, while the laptop viewer mostly moved his eyes around the center of the screen. Since, depending on the distance of the screen, a smartphone can cover a similar amount of the visual field as the thirteen-inch laptop Smith used, it is fair to suspect that the viewers' gaze more likely concentrates at the center of the

[14] The larger the image the more details of it fall outside of the fovea, the area that provides sharp vision. See, for instance, Troscianko et al. (2012) and Chapter Four.

smartphone screen and the attention paid to certain visual elements varies between cinema and smartphones.

The indisputable difference between screen size and the ensuing physiological effects is the core ingredient in the factors that define smartphone spectatorship and the initial point of departure for a methodological combination of cinema and smartphones. However, besides the difference in the physical sizes of the screens, another crucial path must be paved before addressing the questions of respective spatial, cultural, and temporal harmonies and dissonances. Each screen functions as a separate physical entity, circumscribed by its ontological presence within different spaces. As functional objects, both a cinema screen and a smartphone materialize in the form of a frame separating the diegetic space from physical surroundings. The material facet of this approach entails the embeddedness of the screen in the environment, while the functional side denotes the screen's pseudo-substantial position. This latter only exists in a system of temporal, social, and cultural correlation, in other words, at the merging point of the film's screening, spectatorial attention, and a social-psychological framework. Reflecting on the Bazanian idea of the frame, Anne Friedberg (2006) proposes a similar idea, in which she compares the illusion of reality mediated by a screen and the physical world framed by a window. She pertinently claims that cinema screens serve as frames "between the material surface of the wall and the view contained within its aperture" (2006, p. 5). In a material sense, the screen separates the "mobility of images" (2006, p. 6) from the static, immaterial surrounding space—the visible virtual space of the movie from the idle space of the spectator. According to this division, the cinema screen obscures the physical world in favor of the diegetic reality, and demarcates the scope of the image inside the architectural construct of the space.

Whereas the cinema screen only marks the location of the image projected by another apparatus, the smartphone screen combines the functions of a demarcator with that of a transparent window over the computerized image. The built-in screen in this regard resembles some historical forms of entertainment such as the eidophusikon (1780s), the phantasmagoria (from the 1790s), or the diorama (introduced by Bouton and Daguerre in the 1820s), which share the presence and framework of a transparent surface that defines the area of representation. In these three apparatuses, the mechanical representation is a painted or crafted scene which the audience sees through a glass wall that ensures the angle of mediation. Given the two-dimensionality of the image, movie screens no longer demand the same definition of viewing angle to create the illusion of a three-dimensional space: each viewer facing the

screen sees the same angle captured by the camera. At the same time, the screen's immediacy does appear in the capacity of a virtual window, which remediates the archaic modes of visual storytelling precisely by maintaining the illusion of the simultaneous presence of diegetic and non-diegetic spaces.

The presence of the screen is more palpable in mobile spectatorship. Similar to cinematic screens, smartphones represent the same demarcator—as the screen fills the entire surface of the device. Yet, instead of separating the motion of the images from the space of controlled behavior, the functional and embodied (principally haptic) presence of the screen synthesizes the two spheres: just as in Nicole's situation in *Artaud Double Bill*, when her visual scope includes both the live-streamed image of *Vivre sa vie* on her mobile phone mediated through her friend's recording and a live presentation of *The Adjuster* on another screen. More specifically, the continuous rendering of the preferred spatial composition forms a synthesis between the mediated (filmic) and unmediated (physically present) stimuli. Nicole's is a typical case in which the spectator has the freedom to adjust the viewing angle and distance through bodily connection and through the device's specifications enabling mobility. The scenario presented in *Artaud Double Bill* accurately illustrates the plasticity of visual information and spatial composition: the two films which Nicole watches cover different segments of her visual field and belong to two different spatial domains. When watching *Vivre sa vie,* her bodily (hand and arm) movements become involved in her phenomenological experience, and by focusing her momentary attention on her phone, the screening room and the big screen shift to her peripheral vision. The cinema screen, thus, loses its window status. However, while concentrating on *The Adjuster*, the mobile screen becomes effaced as part of the physical, yet hollow environment, eliminated by the cinema screen's affective presence. Screen placement of smartphones, hence, challenges the reconstruction of the cinematic experience in two ways: first, by the spectator's active participation in defining the proportion of the screen and therefore, second, in the lack of a clear demarcation of diegesis and the cognitive and social framework of the physical space, which will be the subject of the next section. The epistemological grounds for these challenges call for a holistic rationale of viewer enactment that presumes an interdependent system that includes the viewer's senses, the screen, and the projected sensory modalities, as well as the physical and social attributes of her surroundings.

The enactment theory proposed by Varela, Thompson, and Rosch (1991) postulates the observer as a mediator between the mental and physical dimensions of perception. This suggests that viewing factors, here, the movie

stimulus and the environment, rely on a dynamic interaction between the viewer's perceptual and social position. Following this definition (which will be revisited in Chapter Two), I claim that the screen's position and the spectator's body (that defines this position) mediate, or even manipulate, cognitive patterns in favor of a coherent understanding of filmic narration. In practice, this means that by holding and moving the screen, the smartphone spectator accommodates the perceptual relationship with the screen (including the visual angle, and therefore gaze) to balance sensory information projected by the screening apparatus and her social presence in the physical space. The attributes of spectatorship, such as immersion, engagement, and comprehension, closely depend upon the screen's position in the space and, therefore, the proportion of the filmic to real-world stimuli. As per a succinctly put attempt to determine the way movie screens separate film from reality, and to strategically highlight the relevant elements of narration, I pursue a triangular model with the spectator, screening apparatus, and viewing space located at the respective vertices. Cinema, according to an idealistically motivated approach, maintains a balance between each vertex of the triangle, while, in the case of the smartphone spectator's substantial embodied presence, the hypothetical angle formed by the screen and the sense organs is in constant flux in response to the demands of the surrounding space. Consequently, as will be further examined in the next few pages, the functional reciprocation of what is virtual and what is physical defines spectatorial attitudes toward cinema and smartphones, especially as a result of physical proximity and manipulability.

Viewing Environment

Beyond the indisputable fact of smartphone's ubiquity, the personal and embodied nature of mobile viewing experience stems chiefly from the smartphone's one-man cinema status, engendered by the matter of subjectivity and the fact that the screen caters typically only one individual. Spectatorship materializes as the irreproducible hybrid of the movie footage, the viewer, and the surrounding environment, while other people inhibiting the physical space remain perceptually independent. To the contrary, the walls of a cinematic screening room demarcate a space of ephemeral collective presence, which ceases to exist when attention is shifted toward the diegesis and which is conditional upon sensory neutrality, bodily immobility, and cognitive attraction. Neither the surrounding unenclosed (or undefined) space, nor the smartphone

screen's allocation is constructed to attain a cinema-like concentrated focus: sensory modalities transmitted through the device mash into the surrounding space and stimulate physical and mental conjunctions between the two parallel fields of vision and sound.

The collective environment demarcated by social, cultural, and architectural benchmarks surpasses the question of copresence. As noted earlier, it also applies to the modal construction of visual and auditory stimuli and the collective dynamics of an integral system of actors—an approach of transience hallmarked by Marc Augé's (1995) *non-places* and Michel Foucault's (trans. 1986) *heterotopia*. But while Augé's premise can be applied to establishing the social characteristics of spectatorship, what Foucault proposes operates more typically in the spatial composition of copresence. In the prologue to *Non-Places*, a book on the anthropological non-existence of people at common places, Augé tells the story of a Frenchman named Pierre Dupont who encounters textbook examples of non-places while taking a trip from Paris to Bangkok. Early in the story at Roissy (Paris's Charles de Gaulle Airport), Dupont confronts the customs and crowds of a passage hub, a place that only exist by convention and because people happen to crisscross each other's ways in this single spot. Augé recounts:

> 'Roissy, just the two of us!': these days, surely, it was in these crowded places where thousands of individual itineraries converged for a moment, unaware of one another, that there survived something of the uncertain charm of the waste lands, the yards and building sites, the station platforms and waiting rooms where travellers break step, of all the chance meeting places where fugitive feelings occur of the possibility of continuing adventure, the feeling that all there is to do is to 'see what happens.' (Augé, 1995, p. 3)

Cars, aircrafts, airports, and auditoriums are all versatile spaces that connect physically existing or mediatized locations. Equivalently, a screening room functions beyond the matter of a constructed stationary space: it fuses collective presence and the collective's mutual interest with functionality, and functionality with the multiple temporal and spatial spheres. The metaphysical and anthropological connotations of spaces in my comparison, however, ensures an approach that accesses any spheres of spectatorship through temporary functions and temporary presences. These spaces operate as frameworks for physical or mental participation and the viewers' embodied presence in the midst of sensory stimulation, which brings theatrical screenings and smartphone spectatorship to common ground.

In line with Augé's theory, the common identity in forming a functional sphere (which could also describe a physical space for spectatorship or a virtual

environment for exchanging movie-related content) secures and shapes a dynamic relationship with the screen and platforms for participation. Yet, as opposed to the concept of *non-places*, spaces understood through heterotopic dimensions omit the aforementioned social dynamics as a referential point, and instead shift the functionality of the screening space to a cultural realm, in which these spaces gain (often inviolable) social capacities as a result of physical or even phantasmic qualities. Such heterogeneity implies the flexibility of spaces and their aptness for transforming the realm of social concepts and sensory stimulation, although not necessarily and certainly not axiomatically. Foucault claims, "we do not live inside a void that could be colored with diverse shades of light, we live inside a set of relations that delineates sites which are irreducible to one another and absolutely not superimposable on one another" (1986, p. 23). At this point, Foucault's meaning approximates Augé's *non-places* and his notion of temporality and cultural function, but Foucault uses a twofold classification: on one side stand spaces that are utopias (social constructs without real sites), and on the other, there are heterotopias ("counter-sites," or "effectively enacted utopia[s]," p. 24).

Heterotopias are spheres that merge embodied presence with an imagined one in a state of absence; some sort of spaces with cultural markers. Foucault allows heterotopias for juxtaposing more than one spatial realm—even unsuited ones—in one physically present environment to connect different temporal spheres, events, protocols, and characters, which can be accessed through cultural rituals (e.g., the formalities of admission). Such spaces are surrounded by physical or cultural barriers that make them isolated yet penetrable, and which frame and secure participation and galvanize exclusion. In a strictly material sense, cinema, for instance, as Foucault describes it, "is a very odd rectangular room, at the end of which, on a two-dimensional screen, one sees the projection of a three-dimensional space" (1986, p. 25).

Space

The approach inspired by the common point in Foucault's post-structuralist and Augé's anthropological view of spaces and emplacement aids a departure from the theatrical screening room's cultural and the smartphone's anthropological position. Besides the apparent differences in screen size, location, and capacity to connect and at the same time divide spheres of spectatorial presence, another fundamental and evident detail in this comparison is the actual environment that the film's sensory scope inhabits. In other words, this is the spectator's

surroundings within which image and sound are exhibited. Such a spatial domain, accessed from material as well as social and behavioral points of view, progresses along the line of the contiguous or even overlapping roles of spectatorship.

En route to the emancipation of cinema, Hugo Münsterberg (1916/2014) drew attention to the seeming dichotomy between the optical dominance of the large screen and the imaginariness of the enlarged, superimposed figures. However, his conclusion nevertheless highlights that it has been specifically this dichotomy that defines the cinematic illusion by securing the sensory dominance of film. The dark and neutral space of a cinema room delimits motion and verbal interactions and sets the spectator's body in a static state. Although screening rooms are manifold, the psychophysiological processes involved in attentive behavior toward motion and luminance foster the prevalent tendency of the diegetic space to oppress the physical space, as well as both the social and sensory copresence of viewers. The fusion of bodily and mental presence is defined by the screening apparatus. Contrarily, the smartphone spectator's embodied presence and physical proximity to the screening apparatus prompt an eminently plastic spatial and sensory organization: by bodily movements, for instance, moving the screen or unplugging earphones, one can change or even abort the perceptual connection with the device and intermit the screening.

As noted earlier, if we consider a movie screen to be embedded in the surrounding space, its position in relation to the spectator will govern the balance between mediated and unmediated sensory information, as well as the social provinces of spectatorship. Thus, as a result of an imaginary dislocation of the spatial dimension in the aforementioned triangular system of viewer, screen, and surrounding environment, the perceptual link between viewer and screen would also diverge. This can be induced, for instance, by the presence of distracting elements in the spectator's surroundings. More evidently, in case of sensory distractions, such as sudden noises or changes in luminance the spectator registers, attention is likely drawn to the assumed source of the distracting stimulus. Due to changes in viewing position, in that case the screen will no longer dominate the visual spectrum to the same extent as before.

The spectatorial viewpoint and social dynamics of copresence, in other words, the synergy of a cinema audience or a smartphone spectator's virtual or physical interrelation with others yield a spatial arrangement of sensory information. Thus, three factors must be closely inspected besides the trajectories of stimulation. First, the material construction of the viewing space and the extent to which it is an isolated or penetrable viewing environment highlight the

architectural measures and the screening room's capacity for prompting effective viewership. The cinema space has evolved in relation to spectatorial position, but also has a great impact on forming the behavioral and temporal dimensions of movie consumption. Second, entry rituals signify a fundamental point in isolating the plateau of mental and bodily presence: this reflects the separation between spectatorship and other social roles, which has remarkably shaped cinematic experience, but which also configures the mental focus of the smartphone spectator. In other words, cinema transfers social roles and functions in relation to spectatorship. Lastly, I note the ways the distinct cultural and anthropological functions of viewing environments (the *place of transition*) are juxtaposed. Such an approach yields initial theses concerning how viewership shapes the viewer's social presence and how social presence affects viewing strategies.

The architecture of movie theaters frames the movie's images and sound and designates the area of stimulation. A screening room is a darkened and often sound proof room, an isolated container for audio and image, which augments a virtual space, but obstructs (or at least attempts to obstruct) physical-world stimuli and non-spectatorial behavioral norms. On account of the spatial arrangement, a screening room functions as a mediator. It syncs the spectators' presence by freeing them from socially defined roles and providing a context for immersing themselves into the film diegesis: the legitimate behaviors assigned to social roles (e.g., adult, son, teacher, etc.) become unified into the behavioral framework of spectatorship.

The long-disputed question regarding the efficacy of cinema's isolated space is not resolved with a stroke of the pen. Yet, it seems clear that both the cultural and material foundations of what cinema is in contrast of its migratory tendency and its frequent emergence on extra-cinematic screens in extra-cinematic spaces propose an entry from the *walls* of cinema. The material boundaries and the admission to screenings—the rituals performed between purchasing a ticket and taking a seat—demarcate a clear frontier that isolates spectatorship from other activities and attitudes. Moreover, these establish a temporary link between the individual members of the audience and redistribute forms of behavior. Within the four walls, the intensity of sensory information is without notable sensory differences that would otherwise divide spatial presence; thus, no matter where they are seated, spectators have similar perspectives of the movie stimulus. The amplitude of stimuli arriving from a smartphone, however, is not distributed equally to all who occupy the space of spectatorship. Certainly, the smaller the apparent distance from the mobile screen and its single viewer, the less publicly

available the screened information becomes. Although this statement overlooks the smartphone's obvious feature as an intimate, personal tool and the initial thesis of personal spectatorship as opposed to cinema's public nature, stimulus amplitude proves to be an important point of comparison. First, on account of the audience effect (Hanich, 2018, see above), the sphere of communal copresence is formed due to the mutual sensory experience. Second, it stands in close correlation with the smartphone's immersive quality. More precisely, in case the amplitude of the movie stimulus decreases (e.g., the screen covers less of one's visual field), the intensifying sensory presence of the surrounding environment (and its sensory distractors) can affect viewing experience.

Theatrical screening rooms are specifically designed to promote intensive immersive spectatorship by masking the physical space and by providing intense visual and sonic stimulation (M. Slater & Wilbur, 1997; Wirth et al., 2007). In the case of smartphones, it is bodily presence (i.e., the viewer's physical contact with the mobile screen) that effectuates involvement (see Hirose & Nishio, 2001). But with the lack of a formal spatial arrangement, there is a chance that the proportion of mediated and non-mediated stimuli would fluctuate, which could negatively affect the sensation of narrative immediacy and therefore, immersion and comprehension. A more thorough analysis of user behaviors and screening features to enhance immersion lies in Chapter Two, but in relation to the remediation of cinema, it is nevertheless important to summarize the question by addressing one key notion. Whereas cinematic screening rooms and the related physical constraints, cultural functions, and behavioral patterns have developed an environment for movie spectatorship, a handheld video player synchronizes and thus remediates not the single combination of spectatorship and moving-image storytelling, but also several other cultural and phenomenological concepts. As screenings take place in undetermined spaces, that is, on an (online) computational platform at any of a variety of locations, spectatorship must suit behavioral and mental references from online participation to physical attendance to various social scenarios the mobile spectator confronts while watching.

It is exactly the plasticity of cultural frameworks that defines cinema's viewership-centered and smartphones' participation-centered interpretation—remediation—of moving images. The transience of disparate functional and cultural elements that construct viewing experience out of sensory information (image, sound, and even haptic references) and a contextual framework (narration) are present in the case of both media platforms. According to Foucault's (1986) statement recalled above, the presence of visual and sonic

indicators of the diegetic space transforms the material composition of the screen and the room; and the functional role of cinema lies in this very transformation. Moreover, from another angle, the versatile neutral space integrates the audiovisual narrative in a way that facilitates the physiological processes leading to immersion. The theatrical screening space is thus a transitive sphere, which, by housing the screen and the sound system (or, in certain cases, live performers), exists in relation to spectatorial presence and the sensory modalities originating from or related to the screened movie. Cinema's profound synthesis of sensory information creates what Kracauer (1987) endorsed as "*the total artwork (Gesamtkunstwerk) of effects*" (p. 92); a compilation of artistic intentions, psychological and physiological processes, as well as cultural meanings and subjective cognitive processing.

The composite sensation of the cinematic is thus established upon the illusion of nonmediation (immediacy, a direct sensory connection to the diegetic space) within an environment framed by social, rather than purely material copresence. In addition, cinema's functionality is based on behavioral protocols and the customs of moving-image consumption, which at each and every screening transform the space into one specific component of storytelling. The smartphone marks an even more complex spatial domain. In *Art and Illusion*, E. H. Gombrich (1960/2000) recounts his past experience monitoring radio transmissions for the British Broadcasting Corporation, where he had to collect and interpret information from barely audible recorded speech. The description of his work monitoring highlights the process of comprehending symbols and linguistic patterns as well as the "extent our knowledge and expectations influence our hearing" (1960, p. 204). Watching movies (or consuming any other kind of media, for instance, reading a book or playing a game) in an unenclosed space inherently entails the simultaneous presence of unrelated or irrelevant sensory information, from which one is compelled to select the relevant and cohesive segments to make meaning. In cinema, most sensory modalities carry semantic or some other sort of logical link, for instance, sound effects confirmed by visual representation or containing information that stands in reference to images, and vice versa. In an unenclosed environment, however, it becomes the spectator's mental task to select what is relevant and what is not—akin to Gombrich at the BBC. Accordingly, this extra task in the cognitive processing of incoming stimuli decontextualizes the mobile spectator's participation within the diegetic space, which can affect cognitive workload, attention, and therefore the level of engagement with the movie or video content.

Soundscape

In his acclaimed anthropological work, *The Hidden Dimension*, Edward T. Hall (1966/1990) states that sensory (vocal) shifts in a spatial arrangement differentiate the spheres of interaction between people as *intimate*, *personal*, *social*, and *public* zones. These shifts are the geographic points where, in Western culture, human voice changes its volume from whispering to talking, from talking to shouting, and so on. The intimate zone is the space of physical contact up to half a meter's distance; the personal zone can stretch to approximately 120 centimeters, while the social zone to three and a half meters. Hall appoints the boundary between the two latter zones as the "limit of domination" (1966/1990, p. 121). In viewing spaces, sensory shifts and the amplitude of sonic and visual stimuli carry important semantic cues in accordance to the sociological and physiological dimensions of spectatorship. Using Hall's anthropological classification to analyze the soundscape of mobile spectatorship fosters two premises, one pertaining to the relevance of sonic information and the other concerning behavioral norms (on the latter, see the upcoming subchapter).

I analyze a movie's sound predominantly in terms of its juxtaposition to imagery, as sound projection (the sound system and speakers) is both ideologically and technically intertwined with image projection (the screen). Such a logical link would assume that the auditory element is examined in accordance to the screen. However, sound has an equally (if not more) peculiar physical presence in the viewing space. Instead of directly following the course of screening technology or the complex process of sensory perception as points of departure, it seems more profitable to treat sound as a spatial feature. Michel Chion (1994) echoes that sound waves propagate similarly to gases filling the viewing space, unlike light rays which spread in straight lines. This signifies not only an inevitable (and obvious) feature of sonic experience, but also defines the zones or limits of cultural pertinence, as Hall's theory suggests.

In a screening room, the spectator is seated in a spot with a linear connection to the screen and in the crossfire of surround sound. Unlike radio or music recordings, a movie's sound exists in relation to the imagery: noise and speech have direct, indexical connection, and music has indirect, synesthetic connection to what is presented on the screen. The source of sound (the speakers of the sound system in modern cinemas), however, remain independent of the screen, in which sense sound is *acousmatic* as spectators perceive sonic information via mental spatialization. As a matter of cognition, the sound producing agent (for instance, a speaking character) in most cases is indeed visible on the screen. Sound effects are interpreted as originating from the

segment of diegetic space where the relevant visual cue is located or, in the case of off-screen sound, where it is *presumably* located. To explain this phenomenon, Chion (1994) uses the term *spatial magnetization*, highlighting that, in case of monaural screenings, sound effects are perceived as if they arrive from the diegetic location of the visually depicted source, even if the speaker is located elsewhere. Spatial magnetization serves to recondition perceptual coherence that originates from the physical characteristics of cross-modal (audiovisual) stimulation.

Spatial magnetization certainly did not emerge with Chion's theories in the 1990s—well into the era of multiplex movie theaters and at the dawn of digital cinema. Attempts to advance the spatial composition of cinematic sound were already made in the late 1920s, even before the rise of sound film. In his contribution to the history of cinematic sound in *Sound Theory, Sound Practice*, Rick Altman (1992b) reports that sound pioneer Lee De Forest demanded the placement of speakers playing recorded music in screening rooms in a way to simulate the presence of an orchestra. Later, when audible dialogues arrived, separate speakers were used for music and speech. As if switching between musical and diegetic sound was not enough of a burden for projectionists, another vision would have required them to switch the sound to emanate from one of a number of speakers every time the source location changed on the screen, thereby achieving a spatial effect. The fantastic idea of transferring sound to the speaker closest to the source location was technologically infeasible: not only because operating such a switchboard would have required unattainable accuracy, but also because correlating multiple speakers spread across a rectangular screening room with a space projected on a flat surface simply cannot produce surround effect. Altman recounts, the Society of Motion Picture Engineers presented another solution, that of multiple soundtracks containing different sonic information connected to separate speakers, but that became achievable only at a later stage.

The term, spatial magnetization as well as the attempts to produce stereo sound in screening rooms, resonate with the neuroscientific concept of the *ventriloquist illusion*, which explains the role of visual capture, in other words, the dominance of vision in perception over sound. According to the ventriloquist illusion or ventriloquism effect, the speech of the ventriloquist is perceived as if it were coming from the dummy: the observer connects the speech with the dummy's moving mouth and involuntarily assigns it as the origin of the sound. A similar effect is observable when two sound sources alternate, resulting in the sensation of sound moving from left to right or right to left, but when a visual

cue moving in the opposite direction is added, the sound effects are perceived to be shifting in the same direction as the visual stimulus (Bonath et al., 2007). I hold this point fundamental when discussing cinema's soundscape, since it draws attention to two crucial cognitive features: spatial immersion and contextual meaning. The first feature implies that spectators map the diegetic space through the location and motion of objects and actors, which enhances immersion into a virtually created space, even if it is visualized on a two-dimensional surface. According to the second one, acoustic and visual information extend each other's range by endowing each other with extra dimensions of narrative meaning.

Cinematic sound—whether it is played or performed live—permeates the physically enclosed space, yet closely relies on the phenomenological illusion of synchronized, spatially coherent sensory information (see Kane, 2014). A similar mental processing is applicable to smartphone spectatorship, with one crucial difference. In that case, the soundscape is most often defined by a pair of earphones or headphones that decreases the impact of external auditory stimuli and isolates film sound from the surrounding environment. This auditory feature of spectatorship is based, however, on a generalized scenario; not to mention that the quality and noise-filtering capacity of different earphone models vary (a topic outside of the scope of this text). Yet, it can be safely stated that in this scenario, external sound is more or less blocked; but earphones designed for leisure use are unlikely to keep all noise out.

Beyond extending narrative information, sonic cues also orient viewers and keep them focused, thereby compensating for the deficits the smartphone screen's size. These effects are bound to one important factor, namely, the natural (physical) proportion of sound volume and visual representation (Altman, 1992b). As Altman suggests, this question arose at the same time as the speaker switchboards described above and manifested itself in different technical concepts during the history of moving-images. Such concepts incorporated the spectator's body and sensory presence, instead of focusing all attention to sound production. In movie theaters, not only is a spectator sitting in one specific location throughout the screening; it is not negligible that her sensory organs for seeing and hearing maintain a constant distance from each other. This means, naturally, that the design of cinematic soundscapes must entertain the correct angle between the eyes (directed toward the source of visual information) and the ears (of which we possess one pair for the sake of binaural hearing).[15] Sound

[15] See Altman's (1992b) contemplation of the "normality of a many-eared spectator" (p. 49–50) as a reference to Cass (1930).

in natural environments provides information concerning the perceiver's position in relation to surrounding objects by its intensity (that is, loudness) and reverberation effect (see also, Altman, 1992a). Although, as Langkjær (2000) adds, the perceiver's position (point of listening) in spectatorship holds a conflicting duality between the location of the microphone and that of the audience. Thus, the reference point of the perceiving character within the diegetic space and the spectator who hears the sound through a set of speakers differs. This is amended by equalizing sound intensity and echo.

Nevertheless, a movie's soundscape serves to reconstruct the diegetic space and to involve the spectator and ensure her feeling present within it. This, again, is dependent on the coherence of the visual and sonic information, and the point of reference is exclusively the spectator's body and sensory organs. The nature, intensity, and reverberation characteristics of sound effects that the spectator perceives must be on a scale with the visually represented diegetic environment and the sound source's distance from the camera. In this way, sound and vision are able to provide corresponding spatial information. My argument here stands also in unison with Hall's aforementioned cultural theory, in a sense that sonic cues contain essential semantic information regarding the spatial position of the sound source and its relation to the listener. This point, although very much bound to the specificities of independent cultural systems, reflects on the following, discussed in detail below: the acoustic soundscape and its private and public relevance; the proportion of diegetic, non-diegetic, and extra-filmic sound; and the perception of context-relevant and irrelevant information and its immersive effects (see also Chapters Three and Four).

In the absence of "earlids," obstructing the flow of sound as we can block vision is naturally impossible. Such an abstract idea comes from Chion's (1994) description, in which he suggests that the lack of framed sound in natural environments, either in a metaphorical or in a physiological sense (as, for instance, moving images are framed by the screen), sonic data becomes spatial. When a sound is played through a pair of earphones, the process is reversed: sonic stimulation becomes dependent upon the mechanical channel between the apparatus and the ear and filmic sound simply disappears in cases of interference or removing the earphone. By adjusting volume and thereby including less or more of the physical soundscape in the viewing experience, the spectator strengthens or weakens the mental borders of her personal viewing space.

Observing sonic information seamlessly embedded into space orients the comparison of sound presentation on smartphones and in cinemas toward a

pragmatic rationale, which delineates the differences based on the size and privateness of the acoustic environment. When discussing the collective phenomenology of a cinema audience, Hanich (2018) compares the dynamic formation of a collective experience to the acoustic experience, emphasizing sound as a fundamental source of the collectivity and privacy of spectatorship. Thus, it seems an indisputable conclusion that copresence in a screening room entails a collective sonic experience, whereas prevalent headphone or earphone use during mobile spectatorship both delimits and privatizes the filmic soundscape. My approach to separating cinema and smartphones according to their respective attributes indicates publicness and privacy for describing an array of media- and consumption-specific features. This approach also serves the methodological domain of this text, even if it disregards out-of-the-ordinary circumstances, such as silent cinema screenings.[16]

As touched upon earlier, movie audio mixing allows for the sensation of surround sound, even if auditory channels are restricted to the two speakers with which a pair of earphones is equipped. Yet, these sonic soundscapes differ from natural soundscapes where different kinds of noises, speeches, or music fill the space and provide precise information not only about volume and pitch, but also the location of the source.[17] In accordance with Chion (as well as Pierre Schaeffer's acoustic discoveries—see Chion, 1983), musicologist and composer Denis Smalley (2007) asserts that acoustic space expands according to the direction of the perceived sound. Smalley's argument is based on surveillance of sonic spaces and concludes with the statement that the trajectory of sound provides valuable contextual information about the proximity of a sound-producing object, but not its exact location. This means that one only perceives the *panoramic space* (the acoustic space extended into one's peripheral view) through stereo or surround sound. Consequently, by connecting earphones to a mobile device, the viewer separates the sound and the visual channel, defamiliarizing the natural, parallel perception of visual and sonic stimuli. Thus, listening to a film's sound through a pair of earphones limits access to spatial

[16] In silent cinema screenings, audience members watch movies wearing wireless, often noise-cancelling, earphones, which promotes a fully immersed spectatorship, often with several dubbed versions to choose from, without having to deal with distraction. "The crowd can laugh, sing, eat as loud as they wish without the worry of ruining the experience for others." Quote from the promotional site of Silent Disco Direct, a silent-cinema provider. See also Silent Summer Screenings, Silent Cinema Hire, and others, as well as Hanich (2018).

[17] Binaural (in-ear) microphones and audio post-production enable sound mixing specifically for earphone users that provides similar surround effect as in screening rooms or natural environments. However, this practice is still rather uncommon.

information, and sonic cues unrelated[18] to the movie often interfere with narrative information. Beyond this, sensory modalities lack other sensory references, such as sound to vision or vision to sound, more often than occurs in a cinematic screening room (see also Alexander, 2017). The detachment of sound from the image facilitates cognitive processes to compensate for the missing sensory information, namely, associations. If the spectator becomes distracted in watching or listening, the remaining sensation guides her through the film scene: sound and vision enhance and extend each other by generating associations with preceding encounters with filmic events, narrative themes, and real-life experiences.

When examining the question of soundscape in spectatorship further (and here, as a result of generalized behavioral and consumption patterns, I would use mobile spectatorship as a model), discussing sonic distraction as part of environmental stimulation is unavoidable. When it comes to sound perception and cognitive processing, sensory links between vision and sound, *audiovisual metaphors*, "synaesthetically fuse the appearance of figures, objects, and spaces with cognitive and emotional meanings on the level of narration [to] elucidate affective and physical experiences that go far beyond the communication of cultural symbols" (Fahlenbrach, 2008, p. 86).[19] Applying the theoretical base of the neural and cognitive processes involved in emotions and emotional reactions, Kathrin Fahlenbrach (2005) explains the processes involved in integrating and interpreting multimodal narrative information and assigns the various modal stimulants of a film or other stimulus source to clusters of emotional prototypes or schemata. Such a process, Fahlenbrach explains using a developmental psychological approach, is ascribed to a neural "level of amodal qualities that are processed by all senses, like duration, intensity, position" (2005, p. 87).

In reference to Vittorio Gallese's (2005) work on mirror neurons (further explained in Chapter Three), Fahlenbrach (2008) also presents a socio-neural approach according to which emotional reactions, for example, mimicking or changes in voice timbre, also define emotional schemata. The interpretation of emotional cues (sonic or other narrative information) in Fahlenbrach's reading, then, implies that narrative scenarios can be classified according to cultural symbols and response modalities. When linked to spectatorial behavior, collective and personal alike, I argue that classification of sonic cues and responses is based on emotional control and semantic meaning. The latter

[18] Various auditory data that are connected to activities and actors within an unenclosed space.
[19] See also Fahlenbrach (2005, 2006, 2007).

suggests that auditory and visual cues in spectatorship are indicators of narrative schemata. The former points to a prototypical theory for classifying relevant and irrelevant (distracting) stimuli in screening spaces where different sensory information is assigned to different social functions. As an example, let me take the smartphone spectator who watches a film on a portable device while in transit. She is simultaneously involved in multiple activities and behavioral frameworks (watching and traveling) with different extents of mental presence, emotional devotion, and relevance according to which her responses are organized. In such a scenario, the variety of sonic information that occupies the space may have diverse levels of ecological relevance to the diegetic and physical environments and effectuate reactions of different urgency.[20] For instance, an emergency alarm or closing-doors signal is more likely to evoke an immediate reaction or attempt to investigate than any kind of atmospheric sound or a pair of softly chatting fellow travelers. However, even an alarm sound could go unnoticed if it does not stand in dissonance with the film plot.

Protocols of Spectatorship

Unlike in a screening room, where the movie's sound and any apparent noise deriving from copresence occupy the same space, the use of earphones separates sonic cues in terms of their narrative relevance. This means that sound arriving from the earphones is semantically related to the movie, whereas external sonic information is related to the viewer's physical and social surroundings. The mental and physiological responses to sonic and visual stimuli nested in any viewing spaces are not only dependent upon the alignment of the screen and speakers, but also upon the behavioral and social protocols employed in spectatorship. Cinema experience ranges from film art to the related cultural practices, so both the aesthetic and the ideological components hold roles in the contextual meaning-making process.

After the centuries-long era of theatrical art, the appearance of cinema created a sphere where high culture fused with more mundane entertainment, and which addressed the self instead of the community. Yet, despite the focus on individuals, the simultaneous presence of various communities and interests exposed to the same stimuli reformed the collective experience. As opposed to a variety of customs across the globe, Hollywood and the Western art house

[20] For an extensive description of my paradigm of stimulus urgency, intensity, and ecological relevance, see the discussions of environmental distraction in Chapters Three and Four.

admonish spectators to refrain from interacting, thereby confining the social sphere to movie-watching's spatial and temporal frameworks. The socially "passive" Western audience evolved from the cultural and technological transformations of theaters and cinemas at the turn of the twentieth century. Introducing lighting systems, which shifted the center of attention exclusively to the stage or screen removed the relevance of social status and granted audience members the pleasure of a fairly immersive entertainment. This practice also aimed to eliminate public political discourse inside theaters or auditoriums, thus reserving opinions to the private sphere.

Cinema's functional spatial outline compels me to address another path for mapping cinematic behavior: the evolution of narration. Before the first years of the twentieth century, cinemas were more similar to other types of theatrical entertainment than to modern screenings. Just like in a revue show, screens (stages) served as background spectacles to social get-togethers. Following Kristin Thompson's distinction between primitive and classical cinema, Lev Manovich (2001) reminds us of the early tumbling spectators who were "free to interact, come and go, and maintain a psychological distance from the virtual world of the cinematic narrative" (p. 107). Then, cinema established the link between film duration, space, and spectatorial behavior, and reformed the norms accommodating storytelling.

The spatial plan of screening rooms not only defines the center of attention, it also demarcates the protocols of copresence. In *No Sense of Space*, Joshua Meyrowitz (1985) explains that practices and physical objects, such as buildings, characterize social and physical spheres. To Meyrowitz, such spheres are circumscribed by physical and social frontiers that one must pass through in order to attend a given activity. As briefly addressed earlier, the rituals of participation, that is, purchasing a ticket, taking one's seat, and staying for a fixed amount of time define the phenomenological domain of spectatorship, the social domain of copresence, and the cultural domain of cinema screenings (as well as the functionality of the screening room as per Augé's *non-places* and Foucault's *heterotopias*). As Meyrowitz argues, by partaking in the modalities of a social event, participants share and apply relevant social schemata in the form of knowledge and expectations. Applying Erving Goffman's (1986) theory, individuals identify and interpret social and biological frameworks using cultural knowledge and apply behavioral protocols in consonance with the given situation. In his thesis on primary frames, Goffman notes that each frame holds its own set of behavioral principles that are only applicable to particular cases. Yet, multiple frames—and consequently different personal and social roles—can

function within a single institution or space. Cinema spectators fulfill multiple roles simultaneously; for instance, the comedy fan, someone's viewing companion, the spectator that stays silent and maintains a constant focus on the screen, and so forth.

Here, I must return to heterotopias for a brief reminder: whereas the spatial and institutional frameworks of cinema define the basic attributes of spectatorial behavior, screens of domestic consumption, such as televisions, computers, and portable devices must reform various spaces in different ways. These screens are short of established institutional settings, therefore spaces, such as living rooms, trains, shopping malls, or offices become spheres that accommodate these screens using two strategies. The viewer either sets the stage for movie-watching and adopts screening routines (e.g., a darkened room) or balances the related activities so the spheres of viewing and other tasks divide the space (e.g., watching while traveling).

My statement also recalls the dynamics of remediation and premediation in that consumer practices, here extra-cinematic spectatorship, frame and are framed by behavioral references borrowed from cinematic protocols. The Goffmanian roles of an extra-cinematic viewer, and in particular, the smartphone viewer, appear to comprise an even broader scope of behavioral norms than those of the cinema spectator. This predominantly emanates from the fact that the viewing space fails to cohere to the actual act of viewing. The spectator, thus, is obliged to employ the behavioral frameworks established for the physical location and the online space—she is a traveler aboard a train, a shopper strolling through boutiques, an employee on a break, a blogger, a social media user, and the list goes on. In any of these situations, the viewing activity is complemented by other social frames, norms of interaction, and copresence.

Unlike cinematic settings that leave little room for distractions or parallel activities, extra-cinematic spectatorship is largely defined by the urge to recreate uninterrupted viewing. Although, the above discussion highlights that this is only one part of the picture. An equally crucial component is managing the apparent (incidental) clash of social or functional behavioral schemes which may "break the frame" and temporarily or permanently disengage the viewer from any of her roles (Goffman, 1986). Both Anna's and Nicole's behaviors in Egoyan's film signify this sort of disintegration of the cinematic spectator and spectatorial behavior. Anna, instead of engaging with *Vivre sa vie* and maintaining the coherent focus as the movie's temporal and the screening room's spatial frameworks dictate, divides her attention between the act of watching and the act of transmitting the film. This not only bestows dual social

roles upon her (that of the movie spectator and the filmmaker and distributor) but the two opposing roles constantly attempt to overwrite each other. Being a filmmaker or distributor of her own version of *Vivre sa vie* distracts her from following Godard's interpretation of *La passion de Jeanne d'Arc*: she fails to comply with cinematic protocols, which Egoyan highlights by her obsession with the actor Antonin Artaud in the role of Father Massieu. Anna's presence in the screening room and in the temporal unit the film circumscribes is compromised and her detachment from mental involvement foregrounds her participation in an online communication sphere where she connects with Nicole, who is unwittingly sitting in a different screening room.

Nicole's scenario is similar, although less about breaking social roles in Goffman's sense. Her scenario instead puts greater emphasis on the parallels between the behavioral and consumption norms of cinema and mobile spectatorship. As she simultaneously watches two films, one on the cinema screen, another on her mobile phone, she is present in two different spheres of spectatorship, involving two different modes of watching and two different narrative and temporal schemes. While paying attention to *The Adjuster* on the big screen, Nicole is temporarily immersed in the cinematic spectacle, concentrating all her attention on the screen and the movie's diegesis. This process, however, is frequently interrupted, either by Anna's live-streamed video or the text messages they exchange. Both of these are independent of Nicole's engagement with the cinema screening, but both are closely tied to the subjective viewing scheme related to mobile personal devices. Her experience becomes characterized by her active participation in maneuvering within multiple spatial and temporal frameworks.

The other crucial factor that defines spectatorial behavior is the screening's temporal frame. In *Concerning the Nature of Time*, Henri Bergson (2002) suggests that duration is a referential unit that indicates a consciously chosen segment of time, which exists simultaneously with other, different durations. Thus, cinematic time (i.e., the duration of a movie) indicates a unit defined by cinematic space. As a result of the aforementioned spectatorial conventions, cinema involves a culturally conditioned, fixed temporal frame, which stands in accordance with the movie's duration, and in which the time preserved in the movie fluxes with the screening time. I argue that during its runtime, a movie replaces the spectator's perception of the physical time with diegetic time. Thus, the temporal frame marked by the material and social framework of screenings becomes an inactive, idle period spent, instead, in the time of the narrative events (see also Doane, 2002). Mary Ann Doane provides an additional, struc-

tural explanation of cinematic time and movies' durations, which is not subject to assimilation or manipulation of any kind and which necessarily structures spectatorship. She claims that the fixed temporal frame for a cinematic screening is the result of the indexicality of the photographic image: each frame on the film reel is one segment of the depicted event in real time and the length of the film strip defines duration. The temporal rigidity of analogue film is maintained in digital cinema too. Smartphones, contrarily, more often operate with easily manipulable digital footage designed for interactive usage. This suggests that, other than material properties, the structure of temporality must be regarded as the result of the viewer's phenomenological connection to the narration as well as the social protocols framing screenings. An interpretation, which differentiates between the filmic spectacle and narrative structure (see also McGowan, 2015), enables a methodological detachment from treating a movie's runtime as a temporal unit and, instead, signals the role of spectatorial behavior in setting the temporal frame of viewership.

Besides the clear evidence of duration, the key indicator for analyzing the temporal framework of spectatorship pertains to appointed viewing times: while duration circumscribes the timeframe of cinematic viewing as a consensual social act, it is the mode of consumption which defines the subjective and flexible time for the smartphone viewer. This mode reforms what Anne Friedberg (1993) calls the "inherent temporal flânerie" (p. 125) of the cinema spectator and manifests in interactive viewership that has an impact on screening tempo and the occurrence of sequences. Due to behavioral clashes that often occur in extra-cinematic spaces, the smartphone viewer defines her own temporal frame, instead of merely witnessing the "display window" of a screen behind which the film unfolds irrespective to her attention. By performing manipulations of the footage, Friedberg claims, the post-cinematic viewer becomes "a ready-made *montagiste*, cutting and pasting images from a wide repertoire of sources at the push of a button," which also effectuates a more subjective *durée* (1993, p. 142).

Spectatorship, according to the same line of thought, is timed by "the mechanism of consumption (i.e. constant unsatisfied desire, the constant hope of a forthcoming but never realized plentitude)" (Kaplan, 1987, p. 28). What E. Ann Kaplan means in her investigation of consumer culture and television is that broadcasting around-the-clock enables a timeless exposure to televised content, and as such, the timeframe for spectatorship is defined simply by the period when the TV set is on. It naturally follows that, instead of content duration, television watching is defined by other activities, or, more specifically,

a lack thereof. Television broadcasting—fairly unequivocally—functions as a loop of content; it goes on regardless of whether someone is watching or not, and thus, its temporal framework approximates that of cinema in the sense that the scheduled start of a program and its duration are equally independent of spectatorial presence. Yet, when it comes to consumption patterns, Kaplan's model can also be applied to smartphone spectatorship. Movies screened on personal, portable screens can appear as temporally displaced *heterochronies*, where the time window for film watching is set according to social situations and the viewer's role in them. Even though the term heterochrony is most often used to describe deviant temporal shifts of development in biological entities, Foucault (1986) adopts it as an indication of *temporal heterotopia*, a "break with…traditional time" (p. 26). Foucault's interpretation appertains to institutions, such as museums and libraries that exhibit multiple temporal dimensions at the same time or to vacation homes and locations hosting reoccurring festivals that reserve a segment of time for certain events. Akin to heterotopic spheres, in which provisional functions take effect through modal changes, heterochronies accumulate and displace the Bergsonian referential time units in relation to social capacities. As a space becomes a sphere of spectatorship with the consonance of the viewer and the screen, screening time is framed in relation to the act of viewing. For instance, smartphone users often choose a movie or video that fits the available temporal frame or watch longer footage in segments or adjust its length by fast forwarding. This, according to Neta Alexander's (2017) thesis of speed watching, often serves as "a compromise between the viewer's schedule and the length of the cinematic work" (p. 105).

The development of screening technologies, the changing cultural habits, and the emergence of portable screens transform the basic act of viewing, as well as the related social interactions. Because of the personal nature of viewing and the flexible space and time, planning may lead to a more carefully scrutinized choice of content—driven by online reviews, descriptions, and social media entries. As discussions shift into online space, after-screening evaluations extend a movie's social and temporal frame: written or audiovisual reflections are less likely to pursue a common viewing experience among parties, but rather function to collect and connect opinions and interests. This virtual community of viewers links previously independent viewing times and spaces. My approach to the (a)temporal framework of mobile spectatorship through behavior and the social and cultural protocols of spectatorship is utile (or even unavoidable) when assuming that smartphone spectatorship frequently takes place in between other activities. Not to mention that the timeframe (the aforementioned subjective

durée) is settled by parallel activities as well as behavior and attention, which, in certain cases, divide spectatorship both spatially and temporally. In this manner, diegetic time exists simultaneously with the act of traveling, waiting, or other activities, instead of dissolving into the material framework of time as it does in cinemas.

Post-cinematic viewing, or post-spectatorship in *Artaud Double Bill* is hallmarked by such a cultural framework and by an abundance of meaning and narratives. The hypermediated sphere that surrounds the two protagonists, Anna and Nicole, in other words, the complexity of movies (and movies within movies), subtitles, visualized text messages, and the online domain as a whole define distinct and even conflicting levels of commitment. The cinematic settings question the legacy of the spatially, temporally, and socially coherent conformity, whereas the augmentation of new media platforms reflect on the way audiences are individualized in a virtual space that is always within reach. Social media platforms, such as YouTube and Vimeo, and even the streaming and producing giants like Netflix or Hulu, further compress the virtual social space on television sets, computers, or portable devices offering high-quality, video-on-demand content. Moreover, by including modules for comments and ratings, as well as direct links to other (social) media sites, these platforms generate a sphere of visible participation and subjective access. Due to the altering dynamics of screens, video sharing sites and applications enable users to experiment with production, ready-made content, and viewing conventions to form the producer's or observer's self—the *you* in YouTube and the *me* in Vimeo.

Chapter II.

Pocket Cinemas

Mobile New Wave: Devices and Spectatorship

A series of monologues are filmed with a mobile phone against bright-colored backgrounds: this is the underlying situation behind the story of Michelangelo, whose invisible and inaudible character interviews the key figures of a New York fashion show parallel to a police investigation of a murder on the catwalk. Sally Potter's movie, *Rage* (2009), centers on Michelangelo's mobile phone camera to strip down the distancing power of the film apparatus, of directing, and of acting and to create what she calls naked cinema (Potter, 2014). The portrayal of the characters—a designer, a fashion critic, a photographer, some models, and other fashion-industry personalities—slowly sets them free from mannerisms and poses. As the movie translates racial and social stereotypes and the issues of fashion industry into humorous, simple, but vivid representations, it also experiments with a formal and, above all, visual language for bringing the characters' faces and bodies within the reach of the spectator (see Figure 4). Potter tells the story exclusively through the characters' viewpoints—showing nothing more than them talking one at a time into the camera, expecting the spectator to connect the dots and solve the mystery. *Rage* introduces two crucial departure points for the following analysis of smartphones and spectatorship. First, the movie reflects on the fact that, by shrinking the spheres of interaction and bringing viewers and content together, smartphones and mobile devices create sensory addiction and bodily immediacy. Second, it highlights the ways in which

participatory culture and mobile-device usage universalize participation and anonymize users and creators.[21]

Figure 4. Stills from *Rage*. Images reprinted with permission from Adventure Pictures.

In the last decade, an increasing number of users have made use of the novelties that smart devices bring to entertainment. Television and video, and later computers and portable video players, reformed the heritage of cinema, and now the smartphone, too, is rushing to ensure its legacy in the moving-image industry. As a tool that functions in close connection to the user's body, the smartphone is capable of connecting users with a large variety of visual and auditory content at any time. But the question arises: what exactly does this connectivity mean in terms of moving-image experience? Inquiry in the emerging research field, which deals with mobile media's application, advantages and drawbacks, as well as its social and economic effects is often limited to the analysis of content (such as forms of communication or consumption: Ling, 2004; Hjorth & Goggin, 2014), politics (digital rights, global and local policies: Steinbock, 2005), or mobile filmmaking (filmmaking and distribution: Odin, 2012; M. Berry & Schleser, 2014). Little attention is paid to user–device interactions and the ways (audiovisual) media references shape and recreate relationships between spectators and smartphones. Informed by interdisciplinary approaches, this chapter aims to reflect on these relationships and their effects on film and video experience and viewing strategies while building upon the remediating role of mobile devices. It is not only of fundamental importance to expand the limited insight into mobile media and moving images, but also to elevate smartphone spectatorship out of the utterly chaotic web of a great variety of screens.

[21] Whereas cinema is associated with a more formal declaration of authorship, content on online platforms often circulates without clear traces.

While cinema gained cultural autonomy by shaping behavioral patterns, audiovisual stimulation, and storytelling, smartphones have been developed in accordance to society's demands for mobilizing telecommunications, business, and entertainment—including moving images (Ling, 2004; Jenkins, 2008; Goggin & Hjorth, 2009). On account of increasingly smart mobile devices, the evolving palette of film and video players, and the ever more experienced spectators, the social significance of pocket-sized devices is definitely growing. Moreover, the changes in consumption schemes affect spectatorship and content, too.

The discussion in Chapter One of Bolter and Grusin's (1999) remediation theory demonstrated it to be valuable to access spectatorial behavior through the smartphone's status as a movie screen and to seek the ways in which it relates to cinema or other platforms of domestic consumption. In addition, insights into the public–private, spectatorial–interactive, and even exploratory–performatory dualities contribute to the upcoming analysis of the factors that influence consumption patterns and the spectator's engagement.

In analyzing one of the mid-generation editions of the Nintendo DS (dual screen) game console, Nanna Verhoeff (2009, 2012) introduces a practice-based, dynamic approach to portable devices. In Verhoeff's reading, a *theoretical console*—as she calls the DS, reframing Hubert Damisch's concept of the theoretical object—is a tool of social and historical status, which is defined by its modes and context of usage instead of its physical properties. A theoretical console communicates with the user through its input and output devices, namely, its touch sensors, buttons, and the screen, respectively. In light of Verhoeff's theory and its applicability to physical devices and interfaces alike, I claim that smartphones must first be accessed here through two of their most fundamental capacities: the contextual understanding of usability and the modal and representational references which are remediated and implemented into the technological affordances of the container medium, the screening apparatus.[22]

In the following, I revisit the notions of remediation and premediation and the theoretical console and apply them to the tasks and opportunities that smartphones as movie players present in order to define the roles of mobility, portability, size, and interactivity. I also demonstrate the ways in which smartphone spectators can control and shape their encounters with moving images in accordance with behavioral frameworks adopted from previous screening media. The analysis presents possible modes of user–device interaction and their effects upon the overall viewing experience. While my

[22] See the explanation of media references and referential objects in Chapter One.

description builds upon what I calculate to be the primary potential ways in which a spectator can engage in interactive viewership, I also acknowledge the existence of diversity in habits, viewing spaces, and stimuli.

One of the attributes that define a smartphone is that it is an inherently private tool, a personal property, however, it mostly functions in relation to public spheres: it connects users with content and fellow users through the internet and is frequently used in public environments. Moreover, smartphones are often regarded as tools for secondary activities, a fact that has a significant impact on the modes of consumption and the amount of consumed content. While cinematic and other types of screens are situated in isolated spaces that promote information transmission, some of the environments in which smartphones are used require users to split their attention between multiple sets of tasks and stimuli. In addition, as I addressed earlier, the viewer often needs to actively filter context-relevant information from the content and external stimuli. A smartphone, therefore, is a media tool that in many cases demand the user's intervention to adjust the spatial, temporal, behavioral, and contextual frames of spectatorship—altering the screen's position, volume, viewing times and duration, as well as self-curating the audiovisual content in response to social and environmental domains.

All these factors considered, smartphone spectatorship, which merges the roles of spectator and user, is characterized by bodily presence that induces an interactive mode of moving-image consumption based on the constraints of the device, user preferences, and environmental demands. This means that bodily intervention defines the personalized characteristics of spectatorship that range from the position of the screen, the firmness and stability of that position, to subsequent manipulations of the image and sound presentation. The smartphone affords a threefold mode of interactivity. First, the viewer has the opportunity to define the proportion of moving-image stimulation in relation to the physical world by fixing the distance between the screen and her sensory organs and performing adjustments to a desired position. Second, she has the option to determine stimulus intensity using the device's built-in functions, such as volume, luminance, and image size.[23] Third, the smartphone allows for interactions that change the flow of the narrative presentation; so the spectator can, for instance, jump between scenes, freeze or pause the footage, or alter the

[23] Although zoom function is not a feature of all of them, most smartphone media players offer some sort of image-calibration functions for setting image ratio or choosing full-screen display.

speed at which it plays.[24] These functions are the ones that not only enable personalized screenings, but also offer the chance to refine the content of spectatorship both spatially and temporally.

Interface and Usability

Although most mobile operating systems offer multitasking solutions (for instance, split screen visualization), the smartphone's interface is optimized for presenting one set of functions at a time. Each of these functions appears as an individual medium that mimics the look and operational mechanisms of physical tools or computer devices while inducing novel spatial, temporal, and behavioral conventions. Thus, a running smartphone application remediates the aesthetics and related practices of other media to grant the user immediate access to the content. For instance, a running video player fills the screen with moving images, thereby remediating other screening media, such as cinema. But while the user is in direct phenomenological and perceptual connection with the film or video, the screening is easily intermitted when switching to another application. The presentation of applications and content fits the properties of the device, such as its size or screen resolution, which are the very attributes that define the smartphone's ability to play and accommodate audiovisual content.

Different mobile operating systems have different interface designs. What they share in common is that smartphone users navigate entirely graphic surfaces, where available commands appear in the form of stylized icons or texts. The representation and range of gestures for the most part differ only across applications and application windows. Each window possesses a particular graphic layout, individual functions, and specific commands, which indicate navigation.

Another significant characteristic of nearly all available smartphones is touchscreen interaction. Most operations are performed by simple moves like touches and swipes, which are easily integrated into the navigation of the graphical user interface. These gestures do not require motor capabilities other than placing one or more fingers on the screen; still, their roots are of great significance: gestures arise from cultural knowledge and usability formulas, and are responses to visual, and in many cases also haptic and sonic stimuli. The

[24] It is important to note that most media players allow for changes in speed, and that this affects (increases or decreases) both the duration of a screening and the sequence. Increasing playback speed categorically excludes a number of frames.

primary move, the touch refers to pressing a button and is used for activating applications and performing various commands. A longer touch (touch and hold), often used for moving elements within the interface, employs mechanisms from computer interactions and resembles mouse navigation. Other gestures, such as pinches or swipes imply physical actions that are used for relocating physical objects in three dimensions.

Integrating multiple functions, references, and protocols, a smartphone interface is best described through its immediacy and intermediality, in other words, sensations of direct sensory connection with a variety of media tools, operated by the above-named gestures. A calculator application, for example, is perceived without the conscious awareness of it being only a representation on a smartphone's screen. It is, instead, perceived as a tool capable of performing calculations after typing digits and choosing functions. The same applies to video players and spectatorship: the device and screening applications are transparent in favor of accessing the narrative content, the characters, their actions, and their surrounding diegetic space.

Reframing Bolter and Grusin's (1999) remediation theory, smartphone applications are perceived as immediate and practical instruments due to the array of sensory, cultural, and technological references among which users maneuver. These references are interface metaphors that provide visual, sonic, or tactile indices regarding the modes and areas of interaction and implement cultural and biomechanical operation methods. The metaphors used must be pervasive enough to cause users to focus on the specific function and orient themselves within the virtual platform, which, in the end, minimizes awareness and automatizes interaction with the device. By applying haptic gestures, such as "pressing" (touching) a "button" (a specific area on the touchscreen), the smartphone gives sensory feedback in the form of animating inward pushing, a sound effect, or vibration (for some examples, see Figure 5). As the range of output devices are limited, the sensory feedback does not necessary correlate with that of the referenced physical tool. Yet, these sensory cues nevertheless strengthen the user's experience of an immediate action and link the virtual representation to actual mechanical feedback. The usability of graphic communication platforms rests on a lawful, causal, and subjective connection between media representation and the potentials for interaction, where both sensory representation and modes of interaction presuppose reference to earlier media and practices.

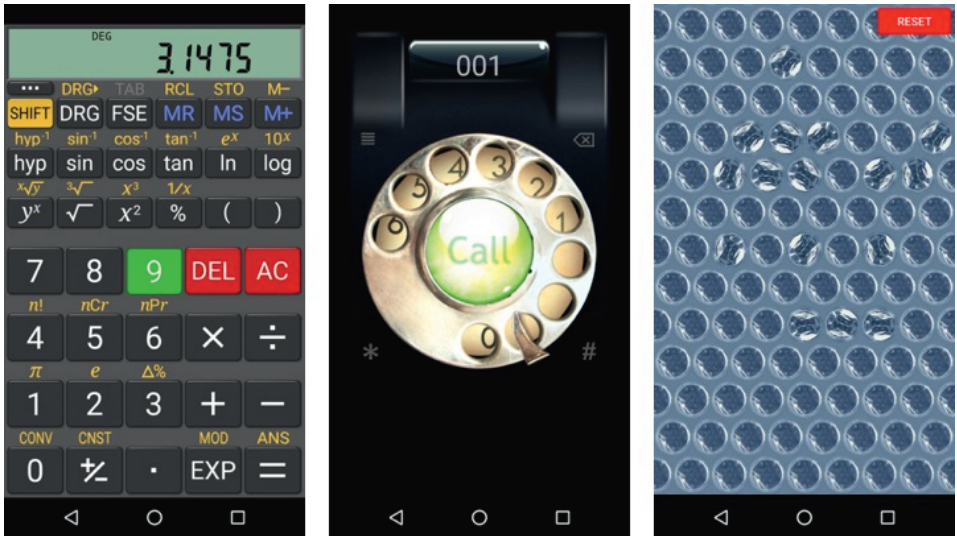


Figure 5. Smartphone applications that provide tactile, sonic, and visual feedback to touch. Screenshots of RealCalc Scientific Calculator; Rotary Phone; Bubble Wrap, Realistic Zen App.

The role of subjectivity requires an explanation here. In a pragmatic sense, functional indicators are references to modes of usage; practically, they provide some sort of sensory (in this case, typically graphic or textual) information about the possible actions. Yet, as cognitive scientist Mark Bickhard (2009, 2017) notes, such indicators must account for the human factor and, therefore, the possibility of trial and error in order to anticipate effective means for communication. Bickhard's position contributes to the discussion of smartphone interfaces because it allows for a wide margin for individual preferences and capabilities. This corresponds to some aspects of the Gibsonian definition of affordances as well. The ecological approach J. J. Gibson (1977, 1979) proposes places affordances in the authority of the perceiver or user. But while it concentrates on a tool's usability, it neglects cognition: memory, experience, and hypotheses testing as part of any forms of interaction. A cognitive aspect challenges the independence of an object's properties, to which Gibson is so committed in most of his theoretical discussions. Physical properties, such as size, surface, or weight can limit modes of usage, but the notion of these limitations can hardly do justice in the case of smartphones. The reason is that the

graphical user interface determines a smartphone's usability just as much as, if not more, the material constraints of the device.[25]

Interactive Spectatorship and a Framework of Games

As stated above, smartphones afford a wide range of interactions, which, in the case of spectatorship, include adjustments to the screen's position, stimulus intensity, and playback modes. Before investigating these interactions and their effects on narrative engagement and audiovisual consumption, I feel obliged to briefly introduce two areas contiguous to this chapter's central discussion: interactive movie screenings and decision-making processes in video-game play. A short introduction will certainly not do justice to these two fields and my intentions lie elsewhere. Just as addressing smartphones' technical specifications and interface design equips this analysis with tools for assessing modes of interaction, interactive screenings and games provide insights for studying viewers' motivations and the involved cognitive processes.

Interactive movies in general frame and direct spectators' choices both materially and contextually. First of all, moments of interaction happen at clearly pronounced times following a specific, pre-determined design, which means that these moments are set to narrative tipping points and announced by moderators and/or intertitles. Decisions are made in a forced-choice manner, where two or more alternatives are offered to spectators to decide upon a character's next action or a situation's outcome. Second, in connection to the forced-choice method, it is the arranger or director of the screening or movie who determines the range of possible outcomes during the process of constructing a "network" of narrative events. This implies that spectators' choice is often just an illusion: while they are given the opportunity to determine certain narrative episodes, these decisions seldom impact the movie's conclusion.

One of the first-ever interactive movie screenings was part of the *Kinoautomat* project presented in the Czechoslovak pavilion at the Montreal Expo in 1967. During the screening of the movie, *One Man and His World* (*One Man and His*

[25] The widely chaotic discussion of design, usability, and affordances that is often embroiled in misinterpretations, however, offers some aid to the upcoming discussion of the role of interaction in movie watching. Extending Gibson's theory, Gaver (1991) takes a stand for active exploration for potential interactions; Vera and Simon (1993) include cultural references as indicators; Mateas (2001) distinguishes material and formal affordances to define opportunities as separate from intentions; and Young and Cardona-Rivera (2011) introduce the notion of narrative affordances to define moments in a narrative (game play) that impact understanding of a broader storyline.

House in other versions, directed by Radúz Činčera), 124 viewers voted for one of two possible outcomes at different points in the story, which defined how the main character, Petr Novák, encountered the next key event in his adventures. At each of these tipping points—nine altogether—a moderator (in fact, Miroslav Horníček, the lead actor) appeared on stage to present the two possible continuations of the given scene from which audience members could choose by pushing a button attached to their seats. Then, the ensuing scene that had received more votes was played from one of two synchronized projectors. The outcome of the narration in the *Kinoautomat* was not something the audience could influence, however: Činčera's trick was that, although the screenplay contained two different plot lines at each moment when a choice could be made, they both led to the same climax and the film ended in the same way anyway (Hales, 2005; Naimark, 1998). Satirically, this deceit was not even hidden; the film made in the Communist-era Czechoslovakia opened with the final event, a scene depicting Mr. Novák's burning house that revealed the absolute illusion of audience intervention. A show that proclaimed people's choice in a plot with a pre-defined outcome ridiculed the illusory democracy in Communist states.[26]

Similar to interactive screenings, user involvement in narration is likewise an essential element in video games. But unlike interactive screenings, interventions during game play are motivated by real-time decision-making processes and include either forced or free choices. The contribution of a game framework to this analysis includes the cognitive processes, decision-making, and behavioral responses involved in momentary actions that comprise two trajectories, the temporal and the affective dimensions of interactions. In terms of time, these can be direct or off-line involvement in a video game play, where the former signifies momentary decision-making based on cognitive mechanisms and the latter reflects on long-term engagement, primarily characterized by affective involvement.[27] These processes are manifested in mechanical (motor) control used to orient, move, and operate avatars and objects within the game space, in interactions with other agents and players, and in emotional reactions, which are all based on the game's rules and goals. Thus, video games provide a flexible window for the time and modes of interactions that either reward or punish the

[26] Years after the Montreal Expo, the movie was aired on Czech television, simulcast on channels 1 and 2. Viewers were free to decide which outcome they preferred by changing the channel. It, however, failed to be a success: people complained of feeling tricked as the truth behind the *Kinoautomat* became apparent (Naimark, 1998).

[27] See also Calleja (2011), who distinguishes interactions in response to micro or macro involvement.

player, and such emotional feedback is rarely present in the case of forced-choice interactive movies. The reason, I argue, is that due to the structure of interactive screenings, interactive episodes interrupt the viewer's emotional engagement with the narrative and the freedom of choice becomes less dominant. Emotional rewards in this case are negligible. In contrast, decision-making while playing video games accommodates the player's affective involvement, suggesting that emotional reactions to the results of actions are closely connected to the player's immersion in the game's diegesis.

A prerequisite for interaction is attention. As unambiguous as it sounds, involvement is an embodied and emotion-driven experience and it is as impossible without directed attention to a stimulus as it is without the awareness of the mechanisms of interaction. Selectively paying attention to a game would bias immersion; too much attention to physical control would result in a shallow level of immersion in the narrative or the diegetic space. This identifies the most important difference between video games and interactive movies in the emotional or affective factor. Game players express emotions in response to the game narrative, which are translated into reactions by the player's avatar and, then, manifested as outcomes in the narrative according to the game's context, rules, and protocols.[28] For instance, a gesture of fear or anger exhibited through the controller, mouse, touchscreen, or other input device appears as the avatar's reaction in the virtual sphere: the player performs a mechanical intervention according to the rules of physics and the avatar represents it in relation to its functions and capabilities in the diegetic space. In order to apply this reasoning to interactions with smartphones while watching movies, it is important to view interactions and interactive smartphone spectatorship as decision-making rather than purely mechanical processes in which the viewer or smartphone user applies her abilities to connect physical gestures and virtual representation. Here, viewers use their knowledge of interface metaphors, experiences from other media, as well as their motor skills.

The proposed potential modes of interactive spectatorship on smartphones correspond to interactive screenings in that spectators can seldom influence the outcomes of the narration, even though they can change the process in which they access narrative information. But while during interactive screenings interactions are typically bound to specific moments, smartphone spectators can apply changes to the content's sensory outline and the narrative presentation at any time. These occasions are not bound to specific actions either, but are instead typically defined by the temporal frame or the social or physical sphere

[28] See the affective loop model (Sundström, 2005).

at the spectator's disposal. In this regard, smartphone spectatorship resembles video game play in that interacting with a graphical interface through a touchscreen generates affective involvement, which integrates cognitive and physical participation. Analysis of interactive media experiences thus contributes an especially crucial insight into analyzing viewers' agency in spectatorship and understanding the mechanisms behind tailoring viewing activities.

Multisensory Viewing

Personalized viewing experience is not unique to smartphones. In fact, it is perhaps safe to say that all screening media possess aspects that in some way alter the presentation or perception of content. These are typically bound to external instruments, for instance, the properties of a shared place or the technical specifications of a screening apparatus. Yet, handheld smart devices stand out from other screens: the relationship that is defined by the dynamic, exploratory, and reflexive modes of interaction between a device and its user catalyzes hitherto-unseen freedom in movie and video watching.

Interaction with smartphones and their content emerges from sensory input arriving from both mediated and unmediated sources and is defined by two sets of material properties. The first is portability and mobility, the device's capacity to be transported and moved. In other words, portability signifies that the smartphone can be taken to and used in various spaces due to its light weight and small size, whereas mobility implies that it is handheld and thus the visual angle covered by the screen is adjustable to the user's sensory abilities and preferences at any time. The second set of key properties springs from the interface's design and the touchscreen and entails haptic operation.[29]

The peculiarity of interacting with virtual objects through a graphical user interface lies in the multiple dimensions in which the user's body meets affording surfaces. Such dimensions, which Verhoeff (2012) calls *mobile spheres*, define the modes of physical interference between the body and the device (here, the touch-sensitive surface that turns physical gestures into virtual

[29] In an essay on the aesthetics of interfaces, Manovich (2007) delivers a similar approach when he identifies the duality of the virtual and physical dimensions as a point of departure for studying post-millennial multimedia devices. Manovich draws equal attention to the mechanical and aesthetic domains of media representation which is undoubtedly inevitable when considering spheres of interaction, more specifically the physical and bodily encounter with mediated stimuli.

responses). My inquiry, however, extends this toward the metaphysical limits of utility and cognitive processing. Thus, in examining the potential ways to interact with a smartphone while watching audiovisual content, one must do more than simply observe the outcome, that is, the operative result—a completed action. I therefore dedicate space to remarks on the related sensations, decision-making processes, and ecological practices, necessary to explain the entire communicational process between content, user, and the smartphone's image- and sound-producing instruments. Implementing these notions in the domain of the potentially interactive modes of smartphone film consumption is not only effective for recognizing the modal range of physical gestures and mediated responses, but its phenomenological overtone also helps to define the principles of haptic interaction.[30]

It is the user's or spectator's bodily connection to the smartphone which determines the boundaries of the sphere of interaction. Portable screens are not only objects established by the surrounding environment and system of protocols; haptic communication plays a significant role in defining the appearance and quality of the sensory space to, in the end, augment physical reality with mediated reality (see Verhoeff, 2012). For instance, when one moves the screen, it changes its proportion to the surrounding environment. Such an "intimate" connection fuses materiality with meta-reality, and accordingly, the user's procedural connection to the physical device (smartphone) and the movie screen (touch-sensitive surface) influence the perceptual and cognitive processes of watching and interacting.

The hardware design of smartphones fuses haptic, visual, and sonic elements, which may cause dissonance between sensory input and output. When it comes to touch control, tactile information gathered from the screen (size, surface, weight, etc.) fails to correspond to other sensory modalities communicated by the interface. That is, it is impossible to gain tactile information about the visual representation of a face; the touch will not correlate with the texture, temperature, or other attributes of human faces. Instead, it will collect information about the hardware that visualizes it. It follows that the user's body in this equation serves as both an input and an output device: information is gathered by touching, holding, hearing, and seeing, but responses are mostly executed using haptic control, for instance, moving the device or interacting with the touchscreen. The correspondences and dissonances between the smartphone's

[30] Vivian Sobchack (1992, 2014) experiments with a corresponding idea, only on the position of screens and their alignment in space: her argument rejects the constant balance between screen and viewer and instead emphasizes the capacities of both for dynamic interaction.

design and human perception denote that comprehension and interaction are based on media references and previous encounters with different sensory sensations. For example, the appearance of a virtual calculator application reminds the user of interacting with an electronic calculator. Likewise, the visual (and other sensory) representation and the sensation of touch is cognitively integrated in the action of "pushing" "buttons."

Defining spectatorship on mobile screens through the spheres of bodily connection and intervention is pertinent because of two main factors. First, it reinforces the account of mechanical variables that link the physical and diegetic spaces as during video game play; and second, it aids the discussion of multisensory viewing experience on handheld devices. The addition of haptic control to my analysis of the sensory modalities integrated into the act of smartphone spectatorship produces a coherent understanding of the ways in which viewers augment or divide diegetic, filmic, and environmental stimuli. In the long run, this combination supplies a thread for inspecting embodiment and how context-relevant and context-irrelevant information is filtered during movie or video watching.

Touch Interaction: The Skin of the Smartphone

By establishing a bridge between different sense organs, the touch-sensitive screen adds a new factor, hapticity, to the framework of spectatorship. Through touch, smartphones enable bidirectional and dynamic interactions so users may personalize the sensory and even the structural outline of contents. This turns spectatorship from the phenomenological experience of witnessing a screen into physical contact, into tactile visuality. My argument regarding the relocation of cinema in Chapter One implies that spectators tend to recreate a focused viewing environment and strive for an uninterrupted experience. However, unenclosed environments and undefined (or at least plastic) behavioral formulas, as well as the need to redefine temporal frames or sensory intensity (e.g., the screen's distance) often emerge irrespective of one's preferences or desire to commit to the cinematic. Although it is an essential part of communicating with smartphones, touchscreen interaction is hardly an integral part of moving-image spectatorship. Yet, the most popular mobile video player applications on the market[31] all have user interfaces designed for touch interaction during viewing.

[31] The top-rated streaming and video player applications with most downloads from Google Play Store and Apple App Store in February 2019.

The marketing for these applications openly claims superiority to cinema's static screen, celebrating functions such as changing screen orientations, image-size settings and zoom function, split-screen or popup replay windows, position seeking, as well as playing in fast and slow motion. These functions are available by touch and are designed to accommodate the mobile spectator's emerging needs in the late 2010s. As an essential feature of smartphone usage, touch interaction demands a detailed description which I attempt to complete by first addressing the biological and technical mechanisms involved and then analyzing the opportunities it brings to spectatorship.

Touch interaction necessarily induces the need to define haptic perception, the sense of gathering information through somatosensation (by skin receptors) and kinesthetic sensing (by muscles, tendons, and joints). Somatosensation, also referred to as tactition, registers tactile data through the cutaneous nerves in the skin, such as touch, pressure, temperature, or texture. Kinesthetic sensing happens through sensors in muscles, tendons, and joints and provides information about an object's size and weight. According to my earlier claim, the smartphone provides limited haptic information as it is a principally (audio)visual instrument. This means that users gain little haptic information about tools, objects, or figures mediated by the interface, as they lack weight, texture, or even size in a physical sense explorable by touch. As physically palpable objects, smartphones engage in kinesthetic interactions but are, again, incapable of providing independent tactile information that correlates to visual or auditory representation. In this sense, objects that appear on the touch-sensitive screen possess metaphysical haptic qualities.

Instead of being touchable in reality, the visual and auditory (even audio-visual) representation of bodies, objects, and surfaces on a touchscreen evoke haptic memories, which complete the missing information about texture, temperature, and the like. This is the process that guides the options, needs, and even desire, to interact with objects in the virtual sphere where the authority of muscles and receptors is minor. Interaction with a touchscreen evokes a particular alliance of the screen and the user's body, involving physical touch that elicits virtual actions manifested in sensory responses, like a push-button animation, sound effects, or vibrations. Because these sensations fail to correlate, the user must rely on references and associations.

The protocols that guide touchscreen interactions with smartphones are complex; inevitably more complex than command-specific physical instruments that allow for one-way interaction or a gesture that yields a single assigned result. For instance, turning a door handle effects opening a latch; pressing a key

on a keyboard activates a specific command. On the smartphone's screen, visual indicators (e.g., panels, bars, links) represent the affording areas and possible gestures, which I explain using a media-ecological approach before turning to the more specific case of haptic control in smartphone spectatorship.

When analyzing interaction mechanisms, and more specifically the physical, modal, and intellectual components of user involvement, it is key to return to how smart devices reference other media. In practice, this points to observations that the interface remediates the appearance and mechanisms of other instruments in order to optimize usability and gestures (i.e., the aforementioned touching or pinching gestures) that potentially lead to a sensory output. Hence, interaction is conditional upon the mimetic qualities of the device and the user's recognition thereof. Gibson (1977, 1979), by contrast, suggests that the question of usability must focus on material qualities, as I briefly introduced above. Gibson claims that affordances are the potential ways of interacting with an object, proposing an object-centered view even though he acknowledges that interaction is impossible if affordances are not accessible to the potential users. But this fails to account for experience, memory, and hypothesis testing (see above, and Bickhard, 2009, 2017).

One of the most prominent theoreticians of affordances and usability, Donald Norman (2008, 2013), offers another approach: he proposes that the prerequisites for interaction lie in signifying its location(s) and mode(s). Norman's approach clearly illustrates that his position stems from design rather than from the products of nature that motivate Gibson's arguments, which explains the way he debates Gibson's phrasing. In another work, Norman (1999) claims that it is crucial to distinguish affordances from *perceived* affordances, which only seemingly solves the dispute: the question of whether affordances are chiefly based on an object's qualities (Gibson, 1977, 1979), signals (Norman, 1999), or the perceivers themselves (Vera & Simon, 1993) is a perplexing one. But when it comes to smartphone screens, the objective is not to insert another reading of design and usability, but rather to provide a cogent link between the semiotics of multisensory representation on the screen and the potential modes of interaction from the user's point of view.

Due to the smartphone's constantly changing virtual interface, interaction is only partly based on the elemental functionalities of the touch-sensitive screen. Another significant factor is what the interface affords: what gestures applied to which specific regions of the screen complete a desired task. The plasticity of the visualized image recalls Norman's (1999) distinction between real and perceived affordances, as well as the definition of *feedback*. While real and perceived

affordances refer to the physically possible actions and actions perceived to be possible, respectively, feedback indicates the presence (or absence) of perceivable information regarding these actions. Feedback is not only important as part of the mechanisms for interaction with an object, but also because without such a concept, the study of cognition and touchscreens would be precarious.[32]

Reviewing Gibson, Norman, and McGrenere and Ho, Rex Hartson (2003) defines a similar concept, coined as *sensory affordance*. But while this term demonstrates the importance of sensing (i.e., the user's perspective), it is more suitable here to regard sensory information and sensory affordances in the context of interactions, for two reasons. First, because Hartson, like Gibson, excludes cognitive factors and regards sensory affordances as predetermined (i.e., designed, built, or programmed) features. Second, because the term does not include the duality of physical motion in relation to a mediated (virtual) referent.[33] These points are significant because the smartphone's affordances are constant and are determined by its specifications (for instance, its size and the touch-sensitive surface), whereas the modes of touch interaction are highly dependent on the ongoing task and its sensory representation.

When embedded into the domain of human–computer interaction and the mechanisms of video game play, a sensory referent achieves two main goals. Besides realizing actions, on a smartphone screen, sensory referents present information regarding the areas touched and the gestures used (e.g., touching, swiping, etc.), which are end results of the user encoding sensory stimuli and decoding them into muscle signals. The smartphone user is consequently required to have sufficient knowledge and capabilities to interact with the physical device and with the interface. This means technical knowledge, for instance, of the sensitivity of the screen or gesture types and the ways to click or drag icons.[34]

---

[32] The concept of sensory feedback is also acknowledged by Gaver (1991), Vera and Simon (1993), McGrenere and Ho (2000), and Mateas (2001).

[33] I use referent here as a term for affordances that are momentary and which appear in relation to user intervention.

[34] Rambusch and Susi (2008) provide solid ground regarding the alignment of technical and virtual knowledge by discussing affording surfaces and areas in game play from Tetris to adventure games. Their approach focuses on the integration of physical control and player strategies in game environments and how "players constantly escape their virtual confines and mingle with the physical and social environment" (Rambusch & Susi, 2008, p. 85). As Rambusch and Susi claim, the challenge, both for players and researchers analyzing their behavior, is translating responses to virtual stimuli into biological and social reactions in a real-world environment. This approach extends the exclusive duality of the physical device and virtual environment and makes cognitive processing part of human–computer interaction, which, conclusively, acknowledges embodiment.

While actual bodily connections to objects supply tactile and kinesthetic information, in the case of smartphones the process is rather abstract. Similar to playing a video game, neither the touch-sensitive surface nor the mobile device itself provides clues as to what is visually represented on the screen. What the smartphone does is mediate: it integrates the sensations of physical touch with virtual image and sound. However, unlike arguments that attribute tactition solely to a function of remembering (as in Marks, 2000; Huhtamo, 2007; Barker, 2009; Verhoeff, 2012), smartphone spectatorship entails a more complex effect. Touchscreen sensors only register the presence or absence of physical conjunction, that is, whether or not there is an operational gesture, whereas the user's bodily sensation is connected to the deformation of the fingertip, the mechanical contact between the flexible skin and the rigid surface. The device subsequently generates responses, which appear as sensory stimuli. Many studies discussing usability and interactions with computer interfaces (some mentioned above) begin by correlating physical touch with virtual responses to make epistemological deductions and draw conclusions about user behavior. Heidi Rae Cooley's (2004) approach follows this method, but actually focuses on the *physique*, more precisely the structure of the user's hands, to examine what she calls the *mobile screenic device* (MSD). Her insights are especially valuable to my discussion of screen positioning, as well as the biomechanical aspects of touching and holding a mobile device, and produce useful inferences for tactile exploration. Cooley labels the intersection of image and touch as *screenic screening* and claims it to be the result of an automatic process: the user memorizes the area of touch and is thus able to focus on the result, which is perhaps the most notable in the case of gaming.

Although her focus is on palpable instruments such as buttons, knobs, and slides, Cooley's reasoning can serve as a basis for comprehending tactile spectatorship. However, full comprehension requires yet another aspect of memory and cognitive processing. "Touching" moving images as part of video player commands introduces similar dissonances to those in game play that divide physical input and virtual output in space and sensory experience. Physical gestures and tactile sensing are detached from the object sensed, which exists in a different spatial construction and holds only visual or sonic qualities. To resolve this ambiguity, I require a cultural understanding, similar to what Laura Marks (2000; see also Marks, 2002) uses in her book, *The Skin of the Film*, to define tactile sensations in connection to the visual language of movies. This solution foregrounds the correspondence of (re)mediation and sensory modalities, notably the way textures and other palpable, haptic qualities repre-

sented on screen stand in reference to haptic memories. Visual (or sonic) representation, accordingly, evokes bodily experience, which, in the case of smartphone, is extended by the physical sensation of touch.

While being exposed to (audio)visual content, the smartphone user maintains constant physical and perceptual contact with the device, which enables touching and interacting with the virtual platform. The screen serves as a medium between the user's skin and virtual objects. Marks's multisensory understanding of cinema is applicable to this analysis considering the modalities of embodiment and the role of haptic memories. Marks (2000, 2002) suggests that narratives (even film narratives), however abstract their presentation is, are understood through associations and references to earlier physical encounters with corresponding objects. Marks illustrates this with the haptic qualities that artworks presented on film evoke. Functional correlations between objects and mental images (i.e., memories) refer to earlier haptic inputs[35] and memories of objects' physical properties. Marks (2000) claims that the sensory understanding of narratives goes well beyond the technical capacities of any screening apparatus; the process can evoke memories of touch, smell, or even taste. "The eyes themselves function like organs of touch" (2000, p. 162), she argues.[36]

Naked cinema that Sally Potter's (2009) *Rage* promotes attempts to recall a similar phenomenological experience (Potter, 2014). By enclosing the characters in their respective colorful boxes, Potter (or the interviewer, Michelangelo's cellphone camera) brings them and their bodies into the spectator's immediate vicinity, a space from which the characters cannot escape, where they are exposed to the gaze of the observer. The bright background colors, the variety of textiles, accessories, hair and skin tones, makeup, wrinkles, and small objects all materialize vividly on screen. Such an organic representation erases the actors behind the characters and creates inherently present, palpable figures.

---

[35] On modes of haptic sensing, see also Lederman and Klatzky (2009).

[36] Marks bases her argument about the memory of touch in spectatorship on Alois Riegl's distinction between optics (an image gained through vision, in which the image is separated from the observer) and haptics (that provokes touching). Marks includes Deleuze's theory of the optical image or the *opsign* which refers to indices that evoke subjective associations. She excludes concrete definitions of screening platforms and the complexity of audiovisual media, instead, focusing mainly on the correlation of vision and touch on any screens.
In Marks's work, the eyes (or in certain cases, the ears) mediate haptic sensations, which Erkki Huhtamo (2007) and Nanna Verhoeff (2012) later label the haptic quality of gaze. In another notable analogy, Jennifer M. Barker (2009) proposes a correspondingly synesthesia-inspired paradigm grounded in phenomenology, according to which intersensorial interaction (that is, the correspondence between vision and tactile memories) guides perception along a contiguity of different sensory modalities. Viewing experience, in this sense, is born of precognitive organic mechanisms.

What the portraiture style of *Rage* achieves is to present the characters through overly haptic figures. The background, their bodies, and attire dominate the viewer's senses: their looks, movements, and how the camera and microphone register even their smallest gestures in the absence of background music or nondiegetic sound effects. To enhance the sense of natural representation, some characters even comment on the settings in which they are filmed—for example, Merlin, the fashion designer, ridicules the choice of background color and Minx, the "celebrity," instructs Michelangelo to shoot close-ups of her face. The role of images in evoking haptic memories necessarily suggests increasing sensations of bodily presence. This is what the backdrop of *Rage* (the flaws in fashion industry) prompt in particular: the tactility of bodies, textiles, jewelry, even names, such as Tiny Diamonds, Lettuce Leaf, or Dwight Angel. Although there is little physical action in the movie, Potter's direction translates vision, sound, and other narrative information into bodily sensations, sensations perceived as if they had originated from direct physical contact.

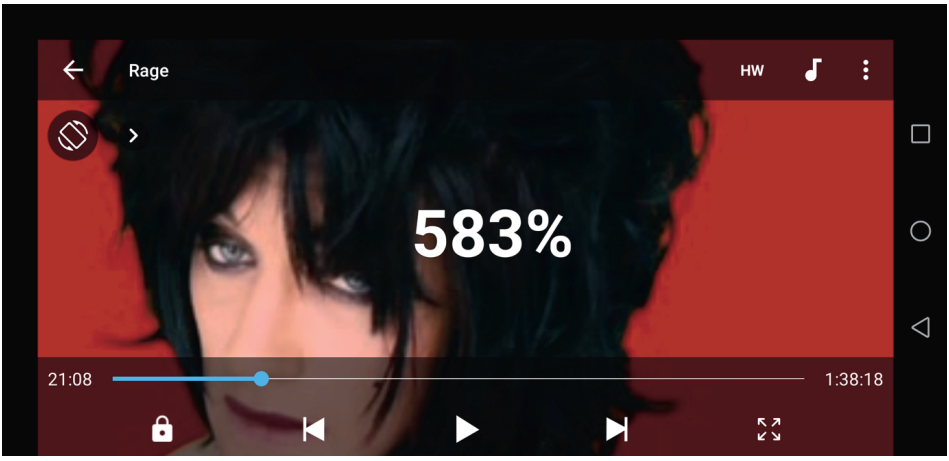


Figure 6. Screenshot of zooming in on a still from *Rage* using MX Player video player application for Android.

The phenomenological link that connects touch with vision and sound not only assigns the optical and sonic representation of bodies and objects to haptic (and mainly tactile) memories. It also involves the real physical touch of the screen which effectuates functions, such as freezing or enlarging the image, and thus elevates touchscreen interaction to the level of physical and spatial interaction. Including such a spatial dimension reflects on the way in which the brain anticipates possible (or necessary) physical interventions and activates the

relevant haptic senses, thereby connecting the sensation of touch with diegetic information that the screen and the screened space contain. It follows that, for instance, enlarging still or moving images by touching the screen (as in Figure 6, the image of Minx in *Rage*) can evoke memories of visualized textures, such as hair or skin, and activate the sensation of tactile sensing in relation to what the viewer sees.

The analogy between physical (motor) gestures and virtual representation logically presumes a mediating agent, which must be examined from a technological perspective as well. The smartphone's touchscreen has a dual function: first, it is an organ of representation and, second, a mediator between the movie's diegetic space and the spectator's body. Therefore, I turn to the screen and its capacity as a touchable surface and a transparent window, an instrument that frames the diegetic space. Providing a thorough overview of the parallels between screens' materiality and virtuality, Giuliana Bruno (2014) defines surfaces as spheres of mediation. While scrutinizing the relations between aesthetics and technology, Bruno emphasizes that materiality is based on *material relations*, that is, on the spatial and temporal coincidence of agents. The screen in general, she claims, "is an active site of exchange between subject and object," "a mobile place of dwelling, a transitional space that activates cultural transits," "a plane that makes possible forms of connectivity, relatedness, and exchange," and "a form of intersubjective transfer that engages the material world and the forms of transformation that operate within its space" (p. 8–9). The emphasis on such correlations foreshadows the mnemonic, or referential, quality of the screen, which, then, points to the transience of the conjunction of a material surface and a virtual image. The smartphone's haptic controls incorporate these elements into the human sensory system, the correspondence of which is the direct consequence of the device's size and mobility. Bruno mentions one additional factor: the spatial construct of spectatorship, which, in the case of smartphones means the planes of manual touch control and screen placement, the subject of the next section. Thus, the screen surface signifies a sphere of physical and virtual contact, which lies on the frontier of material (haptic) interactions and virtual projections.

What Bruno suggests prompts my inquiry into the screen's function as a physical object, a *thing*. One of the basic tenets of Bill Brown's (2004) relevant theory of things is borrowed from A. S. Byatt's *The Biographer's Tale*: by the dirt on the window the protagonist perceives that something is there. The materialism and material presence that encompass the thing concept raise an important question concerning screens and mediation, namely, whether the

material manifestation of a mediated object can mask the presence of the screen.[37] Thus, I propose, cognitive processing of a filmic narrative obscures the presence of the screen—understood both as a physical object with assigned specifications and modes of communication, and as a conceptual entity designed to incorporate and frame the movie's diegetic space. Brown turns Byatt's notion upside down and asks what makes an *object* a *thing*. Corresponding to Bruno's assumption, Brown's proposal suggests that the answer lies in the object's functionality. Brown explains, "the story of objects asserting themselves as things, then, is the story of a changed relation to the human subject and thus the story of how the thing really names less an object than a particular subject-object relation" (2004, p. 4). This implies that the smartphone screen only becomes an active surface when it maintains its projective and responsive function—the former in terms of spectator's perception, the latter in terms of the motivations for interacting.

The presence and potential of smartphone screens' transparent veneer becomes apparent and utile through touch and the user's prosthetic connection with the device. The human factor in producing materiality (or, in Brown's reading, a *thing*) evolves out of the idea of functionality, which brings us back to Verhoeff's (2009, 2012) theoretical console concept. Verhoeff assigns a contextual role to handheld devices, that here would be the context for spectatorship, interactivity, and hapticity. The surface of the screen, thus, acquires a mediative role, which, according to Bruno, is part of user intervention and manual control. The flat and rigid surface inducts the user's touch via visual representation and visual representation via haptic information and the affective quality of the objects depicted in the diegetic space, on the screen. Spatial configurations, Bruno explains, gain affective qualities through tangible sensory information, such as a sense of textures or sizes. Such a form of mediation connects the sensory with the sensational and sentimental to activate spectatorial intervention.

With the smartphone in hand, the spectator attains feedback, a sort of corporeal experience, which prevails in bidirectional communication.[38] Haptic viewing, in synesthetic terms and in tactile sensing, is adjacent to sensory bodily presence. The illusion of being able to touch objects in the diegetic space or even to touch the movie itself through the touchscreen relocates this discourse and its paradoxical overtone from substantial to epistemological grounds, with

[37] The term mediated object, here refers to a film sequence or even objects transmitted through the medium of moving images that appear on the smartphone's surface.
[38] For more on tactile media, see Huhtamo (2017).

an enhanced focus on cultural protocols. Bodily interaction between the smartphone and the spectator watching movies or videos is the result of the constant flux of visual and haptic input and output, which involves cognitive operations and narrative contextualization through seeing, hearing, and an additional sense, key to interactivity: touch.

Mobile Screen, Mobile Spectator

In terms of spectatorship, adjustments executed through altering the screen's position are more ubiquitous than touchscreen control and can be prompted by both conscious decisions and automatic reflexes. The potential involuntariness of kinesthetic interventions can be attributed to proprioception, as I explained in Chapter One: as a flight response, the spectator moves the screen away from her eyes in reaction to possibly unpleasant or dangerous events depicted in a movie or video (e.g., to avoid "collision" with an approaching object) or she moves it closer to gain more information. Moving the screen can change the visual angle and redefine the proportions between screen content and the surrounding space. Apart from its practical or convenience implications, the ability to adjust the visual angle has cognitive and social functions too: a screen held close can positively affect narrative presence and consequently comprehension, and can successfully mask environmental stimuli. Kinesthetic interaction involves two main trajectories, the spectator's motor interventions and the perspectival changes those induce.

The hardware design of modern smartphones diminished—and in many cases entirely eliminated—button interaction, and replaced it with a touch pad built into a high-resolution screen. Through touchscreen control, the user is continually involved in shaping the content on the platform upon which it is visually represented. But what prompted the popularity of watching films and videos on smartphones is their two most apparent specifications, namely, their size and weight, which enable the user's corporal involvement by portability (they can be taken anywhere and used in a wide range of possible environments) and mobility (they can be moved to adjust viewing positions and visual angles). Thus, unlike fixed screens, such as cinema screens or television sets, the viewer may revise the spatial dimensions and adjust the synthesis of film and physical stimuli.

The ontological and spatial characteristics of smartphones and spectatorship establish cultural implications as well: watching audiovisual content on a

smartphone is often considered a secondary task which is integrated into conditions in one's surrounding environment. This environment, as I have argued in Chapter One, is understood both as a social arena and as a physical space in which the spectator's body and the device are connected. This implies that the smartphone's portability and mobility project a novel kind of consumer and spectatorial behavior that affects the location, time, and pace of watching, as well as the placement of the screen and the sound source, and thus the intensity of the experience.

Bodily connection with a smartphone entails not only physical but also social presence, which leads me to copresence and the spectator's personal connection to her property. The clash between these is fundamentally pragmatic: unlike social watching (as in collective entertainment), the screen content on a smartphone is predominantly the concern of only one individual; however, both viewer and screen are exposed to people's gazes. This means that bodily intervention and the proximity of the surrounding environment impact spectatorial experiences and modes of watching. In order to understand the dynamics of kinesthetic control, I must scrutinize the factors that influence and accommodate screen positioning, as well as the cultural and physical spheres of smartphone spectatorship.

Similar to Verhoeff (2009, 2012), I consider the smartphone screen a tool that enables performative viewing in a sense that it provides the possibility to choose the content and shape its presentation at any time. Verhoeff's account confirms that the spatial outline of a portable screen (in her analysis, the Nintendo DS) can be explained either by physical presence (*taking* space) or by the way it forms the realm of usage (*making* space). Verhoeff considers a mobile device as a "mobile *dispositif*" by which she refers to its quality of being "a screening arrangement that encompasses both the perceptual positioning of the screen's beholder, and the physical set-up for the interactive interfacing of the screen's use" (2012, p. 77). While Verhoeff's identification of the specific relationship of the screen to space and soundscape is accurate, the most critical element enabling the fusion of viewing and interacting is not the device itself (that would make it a dispositif), but instead the user and her mental and motor involvement. This is why manual (hand-driven) controls are so significant, but even more important are the decision-making processes that potentially lead to changes in the sensory spectrum, such as the proportion of the movie or video stimulus and the surrounding environment.

The question regarding the sensory field that Verhoeff explores is not an exclusive feature of mobile devices. Even screens with a fixed position challenge

the two-dimensional exploration of a three-dimensional interface. Yet, the conclusion that can be drawn here is that flexible spheres of spectatorship (i.e., the smartphone's mobile quality) become means for interactive, environment-specific viewership. Recalling the comparative analysis of cinema and smartphones in Chapter One, personalized viewing time and space assembles on the axis of the user's body and the device and such control defines the balance between the diegetic, sensory (nondiegetic), and physical (extra-filmic) spheres. This process affects content interpretation and the identification of context-relevant and context-irrelevant narrative information.

To delineate the balance between filmic and non-filmic stimuli and access to information, Francesco Casetti proposes a corresponding solution: he approaches the question of simultaneously existing spaces and the spectator's role in adjusting the screen to environmental demands through the concept of *existential bubbles*. Within the boundaries of such "bubbles," the spectator classifies incoming information and invests her bodily and perceptual capabilities in an effective viewing (see, Casetti, 2008a, 2011a). While present in a collective space, the spectator seeks to exclude (or at least diminish) undesirable disturbances and irrelevant information in order to recreate the intimate experience of cinema. Casetti and Sampietro (2012) claim that "when using a medium in public situations, one often surrounds oneself with invisible barriers that offer refuge, even though one continues to feel open to the gazes of others" (p. 21). The interpretation of a collective environment in this sense goes beyond the matter of copresence—as in a movie theater. Such a line of thought inevitably points to the privateness of the mobile viewing sphere (the screen content and the spectator's engagement with it), as opposed to actual solitude. Casetti and Sampietro argue that an existential bubble is fragile, and has easily violable, but also easily amendable walls. The spectator interacts with the movie stimulus and the screening platform by holding and moving the device: when moving it closer to her eyes, the bubble's walls strengthen; moving it away from them incorporates more surrounding stimuli. By forming one's private sphere within the collective, the spectator synthesizes the narratives of the physical and diegetic spaces, thereby becoming physically immersed into the activity of watching.

The argument that Casetti and Sampietro offer is important for understanding how the spectator's bodily involvement guides her behavior, even if viewing seems to be challenged by the attributes of the screening apparatus or the viewing space. In addition, their theory supports the assumption that, although the spectator endeavors to recreate the most effective viewing

strategies, this attempt will likely fail because mobility necessarily foreshadows attention oscillation between two sets of stimuli and social presence. As a result, smartphone spectatorship is defined by noise, visual elements, social statuses, and behavioral norms generated by the surrounding space just as much as the smartphone and the movie sequence itself: visual and auditory stimuli from the device become embedded in the physical space.

Interactive Spectatorship

Bodily involvement and the options that a smartphone and its interface afford accommodate a broad scope for personalizing movie watching from screen placement, all the way to adjustments in image and sound. Some of these options have been integral parts of and customary to other screening platforms; pausing footage, browsing scenes, or setting sound volume are basic features of television sets and video players. Yet, haptic operations during smartphone spectatorship often exceed the frequency and even the range of these adjustments, and allow for tinkering with the flow of the narration by transposing story elements or changing image size, which concurrently promote involvement and dissociation. This calls for an approach of immersion, borrowing notions from game and virtual reality studies.

As with other interactive media platforms, interactions with smartphones epitomize the aftermath of mental processes and environmental stimulation (i.e., environmental distraction). This process, however, differs greatly from interactive movie screenings, as I argued earlier with respect to *Kinoautomat*. There, the alternation of viewers' physical and diegetic presence was moderated as part of the spectacle: at each voting session, the lead actor playing Mr. Novák appeared on stage and presented the audience's choices in broken English. The sessions created a formal framework, unlike the casual interactions that occur via smartphones. Presumably, the smartphone spectator's active role in the constant flux of stimulation analyzed in the previous sections enhances immersion: manipulations executed through haptic and kinesthetic intervention in favor of stimulus intensity (e.g., volume, image size, luminance, screen distance) and narrative presentation (e.g., playback speed, scene order) may trigger an increased level of cognitive involvement. Yet, self-curating screenings, that is, making impromptu decisions about the modes and content of spectatorship, enhance media awareness, too. Below, I engage in a more

intricate investigation of this paradox; but first, I reflect on the basic characteristics of user control and interaction.

In order to define the phenomenological and media-specific connection between spectator and screening apparatus, I approach immediacy through two modes that may help debunk the ambiguity described above: one mode reflects on perception, the other on media references. Together they imply two different strategies for creating the sensation of feeling direct sensory contact with the screened narrative, characters, and objects, without considering or actively paying attention to the attributes of the screening medium. While *perceptual immediacy* assumes a direct sensory and perceptual relationship to the content and proposes media transparency, *referential immediacy* is a cultural factor, according to which immediacy is achieved when the user discovers or identifies references to earlier encountered media or mediated interfaces—as the theory of remediation suggests (see Chapter One and Bolter & Grusin, 1999). Such a distinction girds the discussion with an adequate tool for recognizing information-projecting strategies related to smartphones and prompts the methodological course the upcoming section takes to examine the motivations and results of interactions.

As noted earlier, the smartphone's graphical user interface reveals clues about locations and modes of touch, which are familiar physical gestures for commanding the device. Most mobile media player applications employ clickable panels that appear on top of the video as a response to a touch of the screen and contain symbols remediated from previous digital or analogue video players. These familiar gestures accommodate usability and reduce media awareness, which makes it reasonable to assume that haptic interaction would not considerably influence the spectator's immersion while watching moving-image content. Although, according to an opposing hypothesis, interactions are often provoked by distraction, which can indisputably redirect the viewer's focus from the movie to the source of distraction. Yet another point links interaction and immersion: it concerns bodily (haptic) involvement and the extent to which such involvement induces or maintains the sensation of presence in the narration. But again, the option of fragmenting or adjusting the presentation of moving-image sequences results in increased awareness, which may remarkably affect immersion, I conclude that the device's specifications can both facilitate and hinder immersion. On the one hand, the small screen size, mobility, and portability increase the likelihood of distraction and wandering attention; on the other, the options for personalizing the screening though the haptic link that connects the apparatus to its viewer suggests increased presence. This duality

serves as a point of departure for the next two sections, which progress toward identifying the modes and possible directions of immersion, engagement, and interaction.

Presence and Engagement

Although immersion is mostly associated with particular mediated environments, such as virtual reality, it may also pertain to handheld devices, as, for instance, Verhoeff (2012) suggests. When discussing immersion and attraction—using Gunning's (1986) cinema of attractions as a starting point—Verhoeff distinguishes among different types of immersion based on spatial recognition, and claims that the sensation is either connected to physical presence in one's environment (*idiopathic immersion*) or to simulation in mediated, diegetic, spaces (*heteropathic immersion*). Applying such a distinction presupposes a perception-based approach, in contrast to stimulation-based approaches,[39] and also differentiates the *feeling* of presence (perception) from immersive *qualities* (medium). IJsselsteijn, de Ridder, Freeman, and Avons (2000) follow a similar direction in claiming that the fidelity of a stimulus source and the user's corresponding sensory-motor experience only partially stem from involvement and presence. They recall the significance of social and personal factors, which corresponds with my approach to examining the experience of smartphone spectatorship.

Observing the feeling of presence through engagement is distantly based on Witmer and Singer's (1998) work, in which they highlight the role of involvement, the conscious direction of attention to stimuli that are deemed meaningful. Witmer and Singer add that the feeling of presence in a mediated space (like a virtual environment or the diegetic space in a movie) increases with

[39] While the basic conditions for immersion depend upon mediated stimuli and their quality and intensity, the feeling of presence foregrounds sensory engagement—a parallel, which often becomes a source for methodological confusion and misconceptions.
The feeling of presence arises out of a cognitive and perceptual illusion of immediacy: the experience of being exposed to a coherent set of stimuli masks the existence of a mediator (e.g., screening instrument) between the spectator and the narrative (see Prothero, Parker, Furness, & Wells, 1995; Lombard & Ditton, 1997; Witmer & Singer, 1998; Mestre, 2005). Immediacy comprises multimodality and involvement (influence, interaction) as well as spatial immersion (Wirth et al., 2007), inclusion (i.e., the exclusion of physical space in favor of virtual space), and the quality and intensity of mediated stimuli (M. Slater & Wilbur, 1997). The lack of or at least a decrease in media awareness is most typically attributed to the user's capabilities and knowledge (e.g., in Suchman, 1987) or embodiment (e.g., in Hirose & Nishio, 2001), which automatize interaction with a medium.

the possibilities of interference. Following this, I classify smartphone spectators' engagement on a four-level scale. This scale demonstrates and compares the effects of media awareness and mental involvement on the user's feeling of presence, and it ranges from telepresence (feeling present in the diegetic space) to mediative presence (feeling present in the physical space and maintaining conscious awareness of the apparatus).

Telepresence, a term introduced by cognitive scientist Marvin Minsky (1980) to denote remote participation through telecommunication tools, is widely used in virtual reality and human–computer interaction studies to demonstrate the extent to which a virtual environment can induce a feeling of "being there." In smartphone usership and spectatorship, it implies a lack of media awareness during interactions. Telepresence includes navigating the device and video player using familiar haptic input methods. Its prerequisites are immediacy, that reflects the illusion of transparency in storytelling, that provides access to narrative engagement, but nevertheless suggests a division between physical (*here*) and virtual space (*there*). Thus, the feeling of being in a remote location contains more than stimulation and attraction; rather it concerns a broader contextual link between stimulus and observer.

The second level is identified as the sensation of engaging with a movie's events by being a first-hand observer or participant, which I call narrative presence. The third level signifies the viewer's parasocial connection with the protagonists involved in these events.[40] Both these levels are closely connected to empathy, and the feeling of presence here is induced by the spectator's emotional connection to characters and their actions. The term parasocial refers to the viewer's social connection to the characters, which, however, is never realized in actual social interactions. The viewer may relate to the characters and understand their roles and actions according to real-life social formulas, for instance, as a mother, doctor, man, or even by establishing emotional reactions to them, such as affection or hate.[41] Parasocial connections, as well as the interpretation of complex social scenarios, are grounded in cognitive and cultural narratives and consequently effectuate emotional reactions to manifestations within the fictional world.

[40] For theories on parasocial connections, see Horton and Wohl (1956) and Lombard and Ditton (2000).

[41] Lombard (1995) points out the indistinguishableness of virtual from natural stimulation, remarking that the involved sensory modalities are perceived in much the same way. Lombard demonstrated experimentally that viewers' reactions to televised figures (news presenters) and live characters were largely identical.

The fourth level, mediative presence, is substantially dependent on the viewer's environment: distracting sonic and visual stimuli decrease immersion in the virtual world and increase awareness of the mediated experience (and, of course, physical presence). The possible scope of interactions with smartphones and screened sequences, signifies spectatorship akin to a performance in which interactions between a viewer and digital content introduces a new, material involvement in the real-time development of the narrative. Within the subjective temporal dimension of the sequence, the viewer generates a self-reflexive private time and stimuli. This performatory mode of viewing aids learning about narrative actions (for instance, by replaying a sequence or watching it in slow motion). However, it hinders observation of diegetic elements in the way they were intended by the filmmaker or addresser.

Mental and Motor Involvement

My approach to analyzing the traits of smartphones as movie screens motivates the understanding of smartphone spectatorship as interactive due to the device's proximity. A viewer's involvement in stimulus construction and narrative production, as defined by Witmer and Singer (1998), signifies "a psychological state experienced as a consequence of focusing one's energy and attention on a coherent set of stimuli or meaningfully related activities and events." This is determined by "the degree of significance or meaning that the individual attaches to the stimuli, activities, or events" (p. 227). Involvement in smartphone spectatorship, as I argue above, is somewhat similar to interactive movie screenings. However, interactive screenings nonetheless limit the scope of narrative intervention to question-and-answer situations and detach interaction mechanisms from diegetic immersion. Involvement in interactive screenings outlines presence exclusively as an element of narration which mode of interaction neglects the spectator's immersion by psychological and motor involvement. In the case of smartphones, the focus is on telic spectatorship, which is not only influenced by the surrounding mediated and unmediated environment, but also by the user's internal motivations and consumption patterns.

In smartphone spectatorship, interactions are incidental rather than something that follows pre-determined timeframes or methods. They may be involuntary or voluntary and depend on the viewer's bodily and mental capabilities, as well on her psychological state in the moment. Interactions, in

addition, may serve to satisfy curiosity or be reactions to the demands of the environment or external, distracting, stimuli. Distraction is perhaps the most typical cause of intervention: the spectator's attention is redirected toward the source of a sound, a visual effect, or other sensory information in her surroundings or on the screen (e.g., a message notification), which can alter her physical, social, or diegetic presence. This means that she adjusts her social behavior and momentarily intermits the screening or strengthens her diegetic presence and performs adjustments to close out disturbances.

In regard to the question of involvement in relation to immersion and distraction, I argue that the palette of interactions manifests itself through two strategies: interaction either arises from external motivations, such as when external stimuli prompt interventions, or from internal factors, such as interest or curiosity. External control points to any pre-existing stimuli or social formulas that affect attention, whether on the screen or in the surrounding space. These distance the viewer from the movie stimulus and motivate her for adjusting the screening in order to restore an immersive state with the narration. Internal motivations, however, engage the viewer with objects, characters, and events represented in the diegetic space.

Changes in a movie's or video's structural outline or the playing thereof in connection to a chain of cause and effect relations create the same culturally induced illusion of control over the narrative as in the *Kinoautomat*. Yet, the mental and physical tools of interaction effectuate a holistic experience that involves the screen and its presence in the physical environment relative to the spectator's body and senses. This suggests that the most fundamental points of interactive spectatorship lie in diegetic engagement and in the viewer's identification with the fictional characters and events (narrative and diegetic presence), but it manifests itself in the execution of bodily interventions with the screen. In the absence of interaction indicators (anything similar to question panels or buttons that interactive screenings use), the viewer's subsequent activity is not well pronounced. However, the coincidence of output and input in one device (the touch*screen*) enhances the reflexivity of the technological, cultural, and cognitive aspects.

Sensory information (from the narrative and the physical world) functions to provide ground for associations. The triad of diegetic, narrative, and bodily dimensions links virtual and physical presences using cultural references, synesthetic and haptic qualities, and mental imagery. The spectator contextualizes the incoming information and builds a logically coherent meaning out of the semantic system of narrative formulas. Precisely this setting

provides a sphere of spectatorship with considerable plasticity, which allows for, motivates, and, thus, potentially leads to spectator interventions in the forms of touch control or kinesthetic operations. In other words, the (inter)action potential of moving images places narrative information into a subjective contextual frame.

The discrete entities involved in narration and interpretation are unified by means of perception and technology. The physical body, the object (the smartphone and its material presence in a physical environment), and the percept (the audiovisual stimuli) generate a formula, according to which off-line and online immersive states merge in simultaneous interaction with both the screen and the represented virtual objects. The physical and narrative presence of a medium manifests in the smartphone screen's capacity for mediating between the virtual image and physical touch. This also encompasses the spatial and temporal plasticity to recompose the quality, the order, and the intensity of images, and to create a tailor-made sequence with subjective immersive quality. The narrative content becomes metaphysical by a phenomenological connection to the viewer and physical via her prosthetic connection with the screen.

Smartphones and the Spectator's Body

The options for interaction, the direct bodily connection, and the smartphone's proximity to the sensory organs effectuate an intimate relationship between the viewer, the device, and its content. Affective qualities of this relationship point to a specific phenomenological experience that links the convenience and pleasure of both witnessing a story and influencing its presentation. The mobility and the direct operability of the smartphone, taken together, create a personal viewing experience that is influenced by an interplay between sensory input, the viewer's preferences, and the environment. Altering pace or sensory modalities affects cognitive processing, spectatorial behavior, and movie and video consumption. Connecting these factors, previously individually examined in this chapter, requires a new approach to theorizing spectatorial experience. The theoretical frame I propose addresses the potential to customize narrative experiences through haptic interactions and the effects thereof on viewing experience.

Embodied Spectatorship

Smartphone spectatorship is a malleable process that involves not only the affective presence of the viewer's body but also motor intervention. The source of this process is smartphones' hardware and software design that provide great latitude in tactile and kinesthetic interactions. The stimuli a smartphone produces during movie or video watching mirrors two processes: on the one hand, there are pre-composed sensory representations that originate from the moving-image content or from the device itself, and on the other, there are stimuli driven by the viewer's bodily configurations. This suggests that, besides narrative comprehension, cognitive processes are immensely dependent on impulses originating from bodily movement and posture, as well as the proprioceptive information that the body's relative position to the smartphone supplies.[42]

Introducing embodiment and embodied viewing into the study of smartphone spectatorship broadens the overview of the process for interpreting incoming stimuli. It also supports my inquiry into the mechanisms for narrative comprehension in unenclosed environments and the crossfire of filmic and extra-filmic stimulation. Altogether, the affective (and again, rather palpable) symbiosis between the spectator's body and the movie or video content leads me to interpret embodiment through the lens of a neo-cognitivist approach. For this, I make use of Hutto and McGivern's (2015) compatible collection of *E-approaches*, which interprets external stimulation as a bodily process based on, among others, embedding, engagement, emotions, and extension. The innovative power of E-approaches lies in appointing the observer's body as an agent that both perceives and contextualizes incoming information and executes responses in reaction to the surrounding physical environment. Hutto and McGivern define cognition as an integrated processing mechanism that stretches beyond the binary view of "perceptual inputs and behavioral outputs" (p. 1) and instead assumes the body's combined role in the two processes. What this approach proposes is that human cognition originates everywhere in the body. The collective performance of the perceptual and motor systems responds to the environment and produces mental, as well as physiological reactions,

[42] Sensory-motor impulses that register the body's configurations in space and inform limb movement are part of the neural model of body schema. For a comprehensive overview of the concept and its applicability to using external tools, see, among others, Head (1920), Holmes and Spence (2004), Maravita and Iriki (2004), and Lewis and Lloyd (2010).

which indicates that situatedness in the surrounding environment is a necessary component for cognition (M. Wilson, 2002; R. A. Wilson & Foglia, 2017).[43]

The ontological versus phenomenological understanding of smartphones sheds light on my discussion of format, interface, and screen. Grasping this duality, I recall Johanna Drucker's (2011) argumentation that draws attention to the mechanisms of interaction with interfaces. She claims that "interface is not a thing, but a zone of affordances organized to support and provoke activities and behaviors probabilistically, rather than mechanically" (p. 7–8). This resonates with my emphasis on the interactive capacity of smartphones, which allows for exploring both the user's perspective and the mechanics of the device. As an interactive platform, the smartphone affords a wide range of actions; but as a functional object, it also delimits the scope of stimulation and bodily control. For instance, the screen localizes and frames the visual composition of the diegetic space and the interface as the hardware localizes and frames various modes of haptic intervention. This proposition bridges the questions of embodiment and the smartphone's ontological and haptic presence while watching audiovisual content.

According to the above line of thought, an analysis of the phenomenological experience of smartphones must consider the complexity of incoming stimuli that arrive from the audiovisual content, the device, and the surrounding, unenclosed space, as well as the mechanisms that organize these stimuli into coherent information clusters.[44] Thus, the phenomenology of spectatorship focuses on the sensation of presence, as well as on access to sensory modalities, specifically, vision and sound.

The lack of a precise definition for phenomenology[45] allows for a certain extent of freedom in applying it to viewing experience without being hijacked toward an approach exclusively grounded in analyzing sensations. Treating perception as a symptom of affect restricts interpretation to perception of semantic units and misses the point of a structural view of the first-person

[43] Vivian Sobchack (2016) follows a similar line of thought, although approaching embodied experience from a phenomenological application to audiovisual media. She bases the phenomenological experience on kinetic perception (in addition to visual and sonic perception), which provokes a seemingly paradoxical reading in the case of handheld devices. On the one hand, I maintain that a smartphone is a physical object and its role in moving-image screening stems from its haptic connection to the user. On the other, the screen must be regarded in its functional capacity; as a compagination and an active entanglement of form, medium, and platform that produces sensory information.

[44] A detailed discussion of narrative comprehension follows in Chapter Three.

[45] The limitations of phenomenology and a lack of discussion of perception is what Maurice Merleau-Ponty (1962) explicitly criticizes in *Phenomenology of Perception*.

perspective in the relationship between perceiver and sensed object. Such argumentation appears in a great deal in Sobchack's (2016) interpretation of phenomenology and film experience as well (see also Sobchack, 1992, 2014). Her interest bypasses the affective quality of the screen, rather, she proposes a view of the *screen-scape* and *screen-sphere*. Sobchack argues, "spectators and screens are now primarily mobile and responsively 'smart' in relation to each other, their movements and interactions destabilizing the fixed position and physical passivity previously associated with watching cinema from a distance and in a theater seat" (2014, p. 88). This is particularly convenient as the current ubiquity and flexibility of screening devices hallmark the antiquatedness of studying screens as material entities attached to a specific space, or even as a primary stimulus source. Instead of iterating the static spatial position of the screen, this argument grants a definition of screening platforms that considers the fluidity of the spectator's phenomenological and bodily connection to the device and viewing environment. The notion of the phenomenological experience of movies questions the monocratic rule of the audiovisual stimulus, which resonates with the comprehension approach in cognitive film studies (see Bordwell, 1985, 1989). This underlines the process of integrating the totality of perceived stimuli into semantic knowledge.

In his seminal work on film and cognitivism, *Narration in the Fiction Film*, David Bordwell (1985) suggests several different paths for illuminating the spectator's role in interpreting information in movies. One of those is the formalist division between story, the chronology of events (fabula) and the fictional plot (syuzhet), that organizes these events. Fabula refers to the cause-and-effect chain of events that is constructed by the perceiver out of narrative cues and the continuous testing of hypotheses regarding how previous events relate to current ones. So, fabula is not only a formal, but also a mental representation of narrative events; subjective but still universal as it is informed by storytelling techniques of which spectators share knowledge (see Chapter Three). The fabula is extracted from the characters and their actions, while the syuzhet operates *with* the spectator, framing the fabula to manipulate, disguise, intensify, or efface events. However, when it comes to media and forms of representation that employ possible interactions between footage and spectator, the narrative gains another, more subjective layer. The story, in a cognitive sense, is filtered through not only a formal (or textual) layer, but also a subjective one—as a result of the continuous reinterpretations of the visual and auditory stimuli according to the spectator's preferences or environmental demands that extra-cinematic spaces make.

The connection between the narrative and the spectator creates a subjective metafiction. In his work concerning the *metaframe*, the "awareness of 'extra-diegetic' and 'metafictional' phenomena" (p. 209), Torben Grodal (1997) connects the filmic medium and narrative devices. This serves as a suitable direction of analysis for defining subjectivity and the cognitive impact of interaction with narratives on smartphones. The frame, Grodal explains, is a container of information—like a photographic frame that presents visual indices. Moreover, at the same time, it serves as a material and metaphorical boundary that demarcates the borders of content and surroundings.

The discussion regarding perspective in film has long been characterized by the exclusive "camera eye" (Spiegel, 1976) that secures the position of the filmmaker in creating a regime of perspective, implying that the screen conserves the viewer's position in relation to itself. The dominant view of the design of visual language—that, in fact, focuses on the composition of the mise-en-scène—sets the fundamental attributes for comprehending the depicted narration. However, and here I return to Grodal's thesis of the metaframe, the visual representation of a narrative event implies the spectator's role and the way she interprets what she sees (or *does not* see) on the screen. Grodal writes, "the cues in the focus are produced by the addresser; the viewer cannot actively control the signals from the screen, only decide whether he wants to provide attention and to reconstruct the emitted" (Grodal, 1997, pp. 210–211). This statement suggests the spectator's dual role in interpreting an audiovisual narrative that is primarily cognitive, but also physical: one can abort the reception of the visual or auditory stimulus (or both) anytime by simply not paying attention. A break in attention can be the outcome of a conscious decision based on cognitive processes, such as choosing not to look at an unpleasant scene, or a result of distraction, when the spectator selectively pays attention to the movie and other mediated or unmediated stimuli in her surroundings.

In his description of attention, Grodal further emphasizes the role of the frame in the process of generating and testing hypotheses about the outcomes of narrative actions. He claims that the frame serves as something that either draws attention to its content or reflects upon what is absent from it. A logical consequence of this statement is that, in the first case, hypotheses are created out of pieces of information that are directly visually perceived, in other words, one looks for patterns to connect these pieces and to explain their role in the fabula. In the second case, interpretation is more closely based on associations and narrative schemata that point toward missing information. The presented

information supports links between elements of the narrative, focuses the viewer's attention, creates a temporal and spatial frame, and serves as a filter for cognitive processes and emotional reactions.

Grodal's definition suggests that the primary function of the screen is to maintain attention to a coherent system of visual information. But this also means that the screen operates as the margin between relevant (on-screen) and irrelevant (off-screen) narrative information. In addition, the viewer's field of vision consists of mediated and unmediated fragments of information that are combined into a metanarrative, a subjective narrative interpretation based on associations that synchronize visual and sonic stimuli with knowledge, cultural background, and experience. The spectator assigns meaning and agency to objects and actors through the relevant contextual framework.

The above balancing of enactivism points to another conclusion, according to which context has a central role in the viewer's interpretation of stimuli. Paraphrasing Grodal's description, it is the subjective narrative context that creates and frames associations and establishes schematic links between elements and agents. But, I argue, not only diegetic information or experience influences narrative interpretation; depending upon the observer, schematic connections may be established among any sources of sensory information which the viewer accesses, even information with roots in the physical world. Taking a view that examines the interactive potential of movie, screen, and mobile device all at once, while also acknowledging the organic instrument's role in decision- and meaning-making, I examine the attributes and specifics of self-curated narratives.

Phenonarratology and the Mobile Mise-en-Scène

The upcoming discussion of the spatial, sensory, and cognitive scope of smartphone spectatorship follows the theoretical framework formulated above, and takes two notions as its point of departure. The first is embodiment, that is, the spectator's behavior in reacting to sensory input, as well as in customizing the screening with her bodily presence. The second is the multi-layered nature of the visual field that includes a systematic view of the screen, the smartphone's interface, and the viewer's physical surroundings. I have explored the question of embodiment through the lenses of mobility and haptic control and argued for its fundamental role in providing the means to shape the sensory and contextual presentation of movies and videos. Smartphones, with their user interfaces and

media player applications, enable a wide range of interventions that include aborting the screening (e.g., pausing or moving the screen), changing image size and aspect ratio, or defining the frequency or duration of visualized film frames by, for instance, changing the position of the seek bar or adjusting playback speed. As I also claim in this chapter, these features are optional (and may vary slightly across phone models and applications). But for the sake of simplicity, the arguments below continue to treat interaction as a feature and perhaps even the integral characteristic of smartphone spectatorship. Some types of interaction serve as responses to more complex social scenarios, for example, choosing a faster playback speed in order to fit a video sequence into a specific timeframe (e.g., time spent in transit). Others, like pausing the footage or moving the screen, are typically spontaneous reactions to momentary situations. Both categories involve the spectator's body and media-specific knowledge, but I will apply the theoretical framework in this section predominantly to interventions triggered by spontaneous reactions.

The smartphone's affordances trigger the spectator's interactions while watching audiovisual content and the combined effect of sensory input and bodily intervention trigger physiological reactions and contextual understanding. Besides embodiment, the other aspect that guides this section deals with how the sensory scope alters as a consequence of interactive spectatorship and the way these changes influence comprehension and the filtering of context-(narrative-) relevant and irrelevant information. Even though the visual range of spectatorship stretches beyond the screen itself in most screening platforms, the plasticity of this scope and eventual shifts in the proportion of the screen and the sensed surrounding space reveal the peculiarity of smartphones. I propose to explain this through the representational and compositional facet of mise-en-scène. I borrow the concept from stage design and moving-image direction and apply it to the case of the *mobile mise-en-scène*, to highlight the significance of the screen's changing visual angle.[46] Thus, my concept of the mobile mise-en-scène

[46] What this approach cannot acknowledge is the presence and role of sound. Although an integral part of both film narration and environmental distraction, sound is beyond the scope of this section. The reason is that I primarily pursue a theoretical and methodological framework that captures the multiple layers of stimulation and interaction, as well as the pliable ontological boundaries of the film/video image, the smartphone and media player interface, and the screen. Constructing a model of smartphone spectatorship through the spotlight of vision enables a detailed discussion and should potentially aid future studies not only of the visual attributes of post-cinematic spectatorship, but also their auditory scope. Still, I maintain that the viewer can shape her perception of the soundscape by adjusting sound volume (see Chapter One) and can respond to sonic distractors according to their urgency and relevance (as discussed in Chapters Three and Four).

acknowledges the malleability of sensory information, perceptual mechanisms, and spectator intervention. This means that the spectator perceives the artificial but composed (filmic) and the natural but plastic (real-life) visual perspectives and it is her attention and bodily position that defines the ratio between the two.

Embodiment and visual perspective provide a methodological framework for studying watching and interacting. I propose the model of *phenonarratology*, that combines phenomenology and narratology to link the spectator's opportunity to influence the movie presentation with the experience of perceiving a customized, at times even fragmented, narrative. The phenonarratology model is inspired by the physical properties of the smartphone, namely, that due to its size and weight, it is portable and mobile. But while it is the screen that frames moving-image stimuli and narration and divides the physical space from the diegetic space, my focus in this section is the spectator, whose physical contact with the screen enables constant reflection upon the content and the surrounding environment.

Smartphone users typically hold the screen in their hands and are aware of interaction mechanisms, such as motor gestures and computerized procedures. Moreover, being familiar with the specifics of the screening platform enables an immediate phenomenological link to the film or video stimulus, which is maintained until the viewer's attention wanders toward distracting stimuli—typically from external sources in one's surroundings or the smartphone interface. In this scenario, three factors are responsible for paying attention to and comprehending moving-image content: first, sensory alertness, the fact that one registers sensory data embedded in the sequence and the surrounding space; second, perceptiveness, the ability to contextualize information as relevant or irrelevant, which means that the observer decodes sensory stimuli and classifies them according to their relevance and urgency;[47] finally, cognitive processing, to interpret incoming information in a contextual framework that fits the narration or the viewer's position in the social structure of the physical space. This latter factor, in addition, induces conscious decisions and reflexive reactions that prompt motor interactions. In response to incoming narrative information, the spectator constantly evaluates whether it suffices, needs to be complemented (for instance, by replaying a scene), or diminished (by terminating the screening).

Viewing circumstances, such as haptic interaction, as well as distracting stimuli that relocate the viewer's attention, can affect the parameters of spectatorship—attention, perception, and comprehension. From another angle, this means that the viewer's subjective interpretation of the moving-image narration

[47] See more on distracting stimuli in Chapters Three and Four.

is the result of the following: perceiving sensory information from the screening platform and the surrounding environment, evaluating this information, pursuing eventual interventions, and perceiving the altered stimuli. In a potential real-life scenario, this statement implies that tactile and kinesthetic (or other bodily) interventions, such as pausing the footage or changing the visual angle are triggered by external distraction. These are stimuli that are likely irrelevant in terms of narrative comprehension and insignificant in the diegetic space or hinder the viewer's immersion, but nevertheless draw attention by their attributes, such as loudness or urgency (see Chapter Three). If the viewer performs an intervention in response, the sensory scope changes. For example, if the screen is moved further from the body to pay attention to a potentially urgent or intense stimulus in the surrounding physical space, the proportion, and consequently, the intensity of the movie or video stimulus decreases.

The other focus of the phenonarratology model is the screening platform, the smartphone, and its capacity as an ontological perimeter between the diegetic and the physical space that defines the proportions of the viewer's visual scope. Here, it is crucial to emphasize the substantiality of the device: the smartphone is an object and, as such, part of the physical space. It unequivocally follows that the sensory spectrum of the smartphone spectator, in other words, all information that she can sense, composed or natural, affects perception and therefore comprehension. While this is by no means exclusive to smartphone spectatorship, the idea propels the organization of my observations and my inquiry into aesthetic composition.

As the most important component of the visual organization of narrative information, the mise-en-scène exhibits momentary visual imaging according to the filmmaker's or addresser's (artistic) intentions, which includes locating objects and agents of interest. But if watching—and especially interactive watching, as in the case of smartphones—can be defined somewhat boldly (yet not without precedent, if we recall Bordwell, 1985) as subjective narrative construction, the collection of visual input can be treated as part of a single visual sphere. In other words, it is the mise-en-scène of movie watching. This assumes that any visual information, online and off-line alike, contributes to interpretation by the viewer though associations.

The representation of a movie's visual space is a constructed structure, a choreographed entity, that builds and maintains the spectator's attention and controls her senses according to a set of pre-determined mechanisms. Bordwell and Thompson (1979/2001) claim that the role of mise-en-scène is framing visual changes in movies, reflecting on the fact that the mise-en-scène guides the

viewer's attention "by means of changes in light, shape, movement, and other aspects of the image" (p. 189). The filmic mise-en-scène, the visual motives in the frame, and the rectangular determination of the recording and screening apparatuses express the meaning-making quality of visual depiction and its dependence upon cultural and social contexts. But such notions redundantly suggest a unity between camera and viewer. Relating filmic mise-en-scène to spatial situatedness is not without difficulty because it must take account of the role of the viewer's cultural reference point and characterize smartphone spectatorship as a bidirectional, metaphysical (cognitive) and physical (interactive) process.

Using the phenonarratology approach, I propose that it is necessary to let go of the preeminence of visual narration in order to understand the connection between a spectator's body and the film image screened on a portable device so as to elaborate a discourse bound to post-cinematic screening platforms. This is because, due to its material properties, the smartphone cannot be examined in the same way as the cinematic apparatus. Its size, weight, positioning, and other specifications make it an extension of the spectator's body, which highly influences viewing experience. Embodiment, that is, the involvement of the body not only on a neural, but also on a behavioral level, catalyzes a subjective reading of information that merges all that is sensed, filmic or otherwise. Extending the examination of the visual field into the space surrounding the screen and to haptic and kinesthetic intervention opens new perspectives in understanding the modes of smartphone spectatorship. Thus, instead of ontologically and theoretically separating the screen from the environment and its heterotopic (transformatory) qualities, my approach treats space in scenographic terms—with elements of different depth, size, and other sensory qualities (see Figure 7).

Theater and performance scholarship (see, for instance, McAuley, 1999) describes theatrical performance as something that actively involves the space beyond the stage, which is arguably correct if we consider the plasticity of live theater. The performance space is incorporated into the narrative both from the side of production and audience. The enactive nature of theater implies the bodily and perceptual presence of the observer and the real-time evolution of both intentional (narrative content) and subjective connections to the narration. The smartphone spectator's visual scope is a similar conceptual frame that inheres in the field of visual information from the screen and the surrounding. The conceptual frame of the theater is evoked to provide an example of the broadness of the observer's sensory scope—beyond the physical arrangements of storytelling agents and devices or visual dramaturgy. The somewhat

ambiguous definition of mise-en-scène in theatrical studies implies the intentionality of staging that embeds the narrative in the sensory elements of the physical space, and this holds true for cinematic and extra-cinematic mise-en-scène as well.

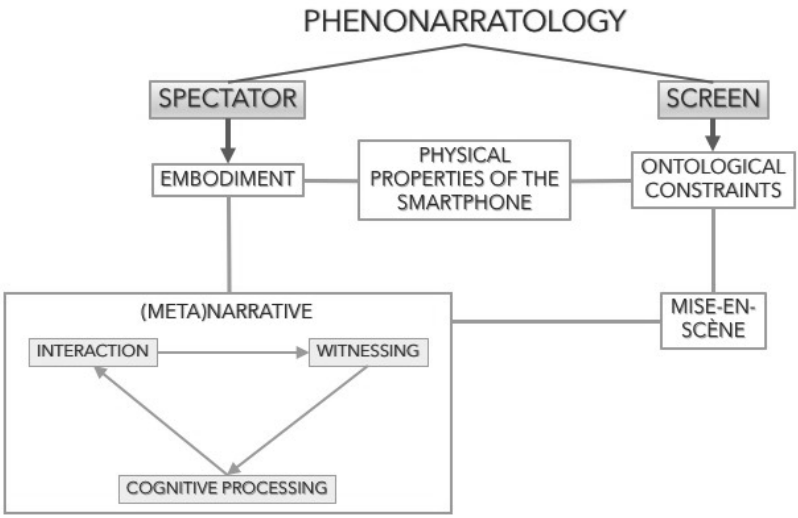


Figure 7. The outline of the phenonarratology model.

The action of controlling image and affect in smartphone spectatorship foreshadows a novel type of cognitive experience in both physical and social senses. This means that the presence of the spectator's body in an undefined social space (shared with objects, other participants, and their interrelations) significantly affects the content, temporal dimensions, and modes of consumption. Subjective viewing experiences, therefore, define the spectator's connection to the movie, the screening apparatus, and the social and physical relations with the surrounding space. This extends from comprehension and emotional responses all the way to social behavior and creates a framework for constructing meaning, in which the spectator's visual overview blurs the ontological and social borders between the movie and the physical space where her body exists. Associations that connect mediated and unmediated sensory stimuli draw upon mental connections between culturally coded concepts and agents and are affected by social patterns such as privacy and embarrassment; shame and fear of one's personal space, for example, the smartphone screen being exposed to others. Linking social behavior to the filmic narrative calls attention to the vanishing perceptual boundary between stimuli within and beyond the

screen. This can result in the endowment of physical agents with roles, personalities, and intentions in connection to the movie (and vice versa). A suitable example is being frightened by a suddenly appearing passerby while watching a stressful movie scene.

The smartphone screen is equally understood as an interface with an interactive outline—a physical device which has spatial dimensions, a given size and image quality—and as a space of representation, that is, an entity that visualizes (locates and frames) three-dimensional images on a two-dimensional surface. The dissonance between the flat glass surface of the touchscreen and the depicted images, which have depth that contain material indicators, such as textures, shapes, and sizes, is somewhat resolved in perception. Yet, no matter to what extent the three-dimensional image is explored through the touch of a two-dimensional surface, the viewer is nonetheless restricted to one point of view dictated by image composition (i.e., design perspective, camera angle). This apparently trivial quality of the flat screen implies that it is impossible to look "behind" the projected objects, unlike in a physical space. This feature also yields conclusions about the structural interrelation of the interface and the real world, as well as about the quality of this contact being defined by the properties of the smartphone. The screen surface marks the border between the physical presence of the user's body and the metaphysical, graphically or photographically rendered space. On a physical level, besides the conjunction of the user's hand and the device, bodily interaction entails exploring a three-dimensional sphere through a flat surface using the applicable gestures. On a mental level, mediated objects are endowed with physical properties irrespective of the angle of observation.[48]

The screen possesses an ability to circumscribe mediated and unmediated forms of representation, which ranges from artistic composition to spontaneously formed chunks of visual information (for instance, in any undirected natural scenes of everyday situations). However, the possibility of tampering with the images—and here, think of the simplest option of changing image size to the more advanced zoom function—eventuates a new "frame" within which it is no longer the filmmaker's intention and the camera that

[48] In his book, *Action in Perception*, Alva Noë (2004) explores a similar path by reflecting on perception, which, unlike seeing, is an active and skilled act of exploration and meaning making. Noë's enactive approach supports my inquiry into the smartphone spectator and draws attention to the eventual manipulation of the visual representation. As Noë also points out, this means that experience is constructed out of puzzle pieces such that the perceiver creates meaning from segmented sensory input: knowledge of an object supports the intake of sufficient information, which complements what is inaccessible due to the physical barriers of sensing.

defines image composition. Instead, the verisimilitude of the image is placed in the hands of the spectator and her exploratory role in spectatorship. Story*telling* shifts from the strict arrangement of visual elements to story *perception* and practical visual exploration; visual framing shifts from a strict, narrative, contextual composition to a subjective one; and finally, the analytical framework shifts from the concept of mise-en-scène and the camera eye to a nonconventional metaframe (Grodal, 1997).

As explained above, the visual design of a story told in moving images, that is, the placement of narrative cues in relation to the frame and each other, are elements of a statement enunciated by the recording apparatus. In Grodal's (1997) reading, this is also subject of sensory attention and cognitive interpretation. What his reading overlooks, however, is the incongruous coequality of the camera direction and the screen, more specifically, the spectator's perspective of the screen. For screens that enable interactivity in the sense of visual enactment within or outside of the frame—which the latest smartphone models by all means do—the plasticity of the frame in a metaphysical sense is in fact a valid dilemma. For instance, zooming in on an image (be it still or moving) and enlarging the object in the center of focus changes the mise-en-scène by occluding elements on the periphery. In this case, the user's intention guides composition; her interest in a particular segment of the image manifests itself in eliminating other, uninteresting segments, which now move beyond the frame of the screen. Similarly, changing the aspect ratio (e.g., holding the device in a vertical position) will increase the proportion of the user interface and decrease the frame of the movie or video.

Perceived visual narratives connected to moving images, screens, and viewing spaces of any kind are defined by the spectator's body and its ability to witness and interact in the virtual and physical spaces through organic and prosthetic devices. In this regard, four apparent constructs of information sources present themselves (see Figure 8): film-image composition, interface design, hardware design, and physical space. These constructs mark the intersection of visual stimulation with bodily involvement, that is, the watched narrative and modes of interaction. Below, I reflect upon these features using the phenonarratology model to sum up the fundamental factors of spectatorship and physical interaction.

Figure 8. The mobile mise-en-scène: a still from *Tangerine*; screenshot of the still displayed using a smartphone video player (MX Player); zooming in on a protagonist's face; and the same still on a smartphone's screen in an urban environment (photo: Olga Nikolaeva).

Firstly and chiefly, viewing experience is established upon the moving-image stimuli, in other words, a narrative rendered into sensory constructs. The storytelling function of audiovisual contents caters to the placement of visual, schematic, and sometimes even indexical references with quasi-universal meaning (see Bordwell, 1985; Bordwell et al., 1985). The spatial and optical organization of the narration guide the spectator's attention according to corporeal and neural mechanisms (see Chapter Three). The visual outline is closed into a rectangular frame according to the filmmaker's or addresser's intentions and appears much the same way on a wide range of screening media, which means that composition largely lacks screen-specific capacities. The visual composition of physical objects and characters induces embodied involvement at the level of film style, acting, physical reactions (Gallese & Guerra, 2012), and narrative schemata (Kiss & Willemsen, 2017).[49]

[49] Kiss and Willemsen's insights support the ecological connection between bodies and filmic mediation as they review the modal functions of narrative construction. Kiss and Willemsen call schemata the engine of sensory involvement that organizes narratives and visual representation into semantic units. Their definition emphasizes that these units are "bodily rooted dynamic patterns that internally organize our experience" and have an "influential role in the initiation and maintenance of *narrative schemas*, as formal gestalts through which one gains comprehensive access to different forms of (film) narratives" (p. 33).

This biosemiotic–ecological understanding of narrative comprehension is a substantial addition to the media-theoretical examination of smartphone screens. While visual composition presupposes a fixed observation point (which, in strictly sensory terms, is the camera's point of view), smart devices have specifically outlined designs that enable haptic intervention in response to movie stimuli. So, in light of touchscreen control, the next aspect of visual imagery refers to the interface and interaction mechanisms with the software, which enable manipulations of visual representation on a virtual platform within the space circumscribed by the screen's edges. Mobile operating systems and video player applications function through visual cues that localize areas for interaction, similar to the buttons on a game controller. In a media player application, such functional areas are either present at all times and can be activated by specific knowledge (e.g., areas that function as volume or luminance sliders in response to specific gestures) or appear as a control bar. In either case, touching the screen brings forth the visual outline of the video player's toolbox over the image. On-screen gestures offer manipulations incorporated into the screened image, and result in a customized screening of the film often disloyal to the author's intentions. Such modes of customization stem from distortions of the original mise-en-scène, for instance, changes in luminance, framing (zoom or aspect ratio), and the presence of a control bar or other icons.

In allocating the frame of the hardware in both a physical and a perceptual sense (screen edges, and object design and operational outline, respectively), the key roles the physical properties of the device play in movie watching are twofold. The screen attributes define image size and quality, while spatial (material) attributes of the smartphone, such as size and weight, delimit the possibilities for adjusting the visual angle. Thus, the perspective of the screened images alters by moving the screen; the dimensions of the surrounding space can be changed by physical actions, such as head or full body movement.[50] The connection between vision and haptic exploration manifests itself in the physical space of action and in perception, which means that tangible and visible qualities are equally embodied and spatially bound. Accordingly, optics, haptics, and the enactive interrelation between an organism and its environment imply a modal congruence in the body and visual stimulus in smartphone spectatorship, as explained above.

[50] Following a phenomenology course, this can be delineated to the realm of embodied visuality corresponding to Merleau-Ponty's (1968) chiasmatic (opto-neural) understanding of perception.

The composite of mediated and unmediated information entails a constant transition in the sensory spectrum in accordance with subjective interpretations of incoming data. Different spheres of sensory schemata are perceived in accordance with a personally assigned narrative that comprises medium- and context-specific knowledge and is based on the sensory constructs within and beyond the screen. The phenonarratology approach enables to understand the way spectatorship synthesizes sensory input and haptic interaction. It, furthermore, points to the role of an enactive narrative experience that is a hybrid of artistic and spectatorial intentions. The approach nonetheless avoids neglecting the pre-cinematic domain and reflects upon the cultural apprehension of motion, which connects mediated and unmediated visual input. The changing visual angle, the scenic view of the diegetic and physical space, and the visual references that connect them demonstrate a fundamentally exploratory viewing. Although the screen frames the composed mechanical representation of objects, actors, actions and their temporal and spatial unity, it also detaches the spectator from the multidimensionality of meta-realities. The situatedness of the gaze in relation to the smartphone merges multiple realities into one physically present, tangible input, while the multisensory interaction that smartphone spectatorship facilitates organizes indexical references in connection with the ambiguity of the staged space.

New Perspectives on an Imperfect Cinema

Juan García Espinosa's (1983) "imperfect cinema" gains importance in the abundance of digital media, the blurring boundaries of spectatorship and participation, and the increasing complexity and multitude of visual and auditory domains. When citing Espinosa, Hito Steyerl (2009) contemplates the social transformation of values in aesthetic representation, the institutional or medium-specific appearances of which sink into the "swarm circulation, digital dispersion, fractured and flexible temporalities" (para. 31) of consumption. This manifesto of abundance and involvement extends the boundaries of moving images to a space away from the shelter of cinema. The proliferation of mobile devices—smartphones, tablets, or phablets—provide but a few examples of this tendency to relocate not only the cinematic experience, but also the cinematic imagery to spheres of non-spectatorship. As prosthetic body parts, smartphones yield tactile and kinesthetic interactions with movies and video sequences that

are reinterpreted through subjectivity and embodiment. While synthesizing mediated and unmediated realities, spectators become physically immersed in creating experiences. This, as Steyerl notes, "transforms quality into accessibility, exhibition value into cult value, films into clips, contemplation into distraction" (para. 2).

The questions this section pursues consequently involve the sovereignty of smartphone video culture and the extension of embodied and technological trajectories toward the cultural ones in a fairly nascent, yet rapidly changing system of audiovisual consumption. In this system, the roles of content production and content access often merge and the increasing online presence and the broadening physical spheres of mobile encounters define the moving-image-exchange industry. My goal is to apply the theoretical discussion of this chapter to the practical realms of mobile film culture. To that end, I continue to examine the viewer's embodied and social presence, which has significance in the renaissance of mobile filmmaking and participatory culture in the late 2010s. There is no longer a definite difference in the quality of cinema and mobile media tools. The screen culture at present is deeply embedded in the values that Espinosa's imperfect cinema once meant for the Cuban film industry: the active presence of *ordinary* people behind the camera and in front of the screen.

You Want Me to Look into the Telephone?

"You want me to look into the telephone? This is the movie camera?" asks Vijay, the pizza-delivery boy and wannabe model in the first scene of *Rage* (2009), while trying out suitable poses for his interview with Michelangelo. He certainly has his doubts about the mobile phone camera, a tool of everyday people, a symbol of mundane interactions in the midst of the sparkling and creative fashion world. Michelangelo uses his mobile phone to capture all that is hidden behind the glitter and to translate this universe for the rest of the world to which he himself belongs.

Economic and social patterns of movie and video consumption on smartphones stand in a rather paradoxical constellation of spectatorship and self-curatorship. This diverges from Bordwell's (2002) *intensified continuity* and *post-continuity* or Shaviro's (2013) *accelerationist aesthetics* toward a representational nihilism where images are floating in the space of subjectivity and user intervention. In her contribution about speed watching on the platforms of *compact*

cinematics, Neta Alexander (2017) places movie consumption in the framework of temporality in both a social and physical sense. Alexander claims that time is not only mirrored in fast-forwarding a footage or skipping frames, which serves to compress a footage into the period designated for movie watching (and not the other way around, as in the case of non-participatory screening platforms, like cinema). It also becomes apparent in the ways the aesthetic traditions of commercial movies have changed. By analyzing the customs of playing moving-image content by compressing its length even by half, Alexander concludes that participatory spectatorship entails the viewer's simultaneous distancing from the narrative information and close involvement with it: missing frames or even scenes and not knowing what exactly was lost "puts the viewer in a limbo of watching and un-watching" (2017, p. 105). While this practice seems to mean an economical use of time, it implies a changing tendency in consumer behavior. Moreover, it highlights the trend toward gamification, both in terms of content and spectatorship. Saving time by speed watching or manually skipping scenes seems to be users' response to an abundance of content and a hunger for information and stimulation. It is sort of a life hack for using time more effectively, which replaces the *flâneur, the urban spectator* with the *flâneur, the curious observer* (Friedberg, 1993; Benjamin, 1997).

Besides compressing consumption, smartphones likewise entail bodily presence and intervention. Whereas the former initiates a cognitive model based on spectators' sensory and mental engagement in the pursuit of making sense of the content, intervention can even make this engagement a distancing force. Although this appears to be a paradox, it is in fact the viewer's body that can shape the direct phenomenological connection between the mind and the narration. As my phenonarratology model suggests, bodily intervention can transform the aesthetic frame of storytelling, for instance, when decreasing image size or moving the device away from the body, the viewer includes the interface and the surrounding space as outer frames of the artistic composition. In addition, interactions can divide or hijack the spectator's attention from the narration toward the smartphone or the environment.

Embodiment, the underlying logical tie behind the multiple aspects of new media consumption, separates, but at the same time, connects immersion and interaction in the dimensions of the smartphone spectacle. This spectacle is far away, but still so near Tom Gunning's (1986) *cinema of attractions*: although in the abundance of mediated representations, a smartphone screen *per se* hardly offers anything sensational, narcissistic power over moving-image consumption offers the pleasure of participation, which, in the end, couples with an impression of

tailor-made conformity. Active participation absorbs the viewer in the act of creating and the act of creating absorbs the movie and the (cinematic) spectatorial experience (see Alexander, 2017). Joost Broeren (2009) follows a similar argument, only to illustrate how YouTube and user participation signify a new era of attractions. He departs from the idea that, as opposed to narrative film, cinema of attractions at the turn of the twentieth century was hallmarked by technological limitations (or rather, specifications), which defined both the duration and content of films. Identifying exhibitionism as the other analogy between early cinema and social video sites, Broeren pronounces that home videos, commercials, and professional content alike are subject to the fast-paced consumer patterns of the 2010s. His argument suggests that participatory film culture paves the way for the evolution of a new moving-image trend, while following in the footsteps of cinema.

Nevertheless, when observing social behavior from the viewpoint of a platform that incorporates the alloy of online and off-line (virtual and physical) presence, it is necessary to acknowledge the role of commercialization and the commodification of media and culture—a process that relocates cinema onto the smartphone's screen and into the spaces of everyday life. The abundance of content indubitably poses questions of quality and quantity, but the preeminent point is how social and cognitive participation shapes movie consumption. In cinema, the physical distance from the screen initiates an immersive state based only on the senses, suggesting that being too close would make spectators vulnerable by their own bodily presence (see my discussion of proprioception in Chapter One). Ironically, this kind of presence is perhaps the reason why smartphones have grown a fan base for movie and video consumption: intimacy, subjectivity, and ubiquity play genuine roles in elevating small handheld devices to the niveau of cinematics.

A look back on Gunning (1995) and his other work on early film confirms the duality of enactment and detachment: he suggests that attraction requires a certain level of alienation—precisely because of the need to eliminate the uncanny feeling of the screening medium and to foreground the narrative. When the functions that the smartphone as a screen and communication device yield and the spectator's physical presence merge into a state of immediacy, the semantic meaning of the narration and the content itself become the central element in cognitive processing. But when the spectator's attention oscillates between content and the sensory and behavioral modalities of the device or the surrounding unenclosed environment, a loop of adaptation commences: she adapts both to the content and the screen size and eventual distractions, which

may project a less engaged viewing experience.[51] This, in the end, seems to explain filmmakers' harsh criticism of consuming movies on smartphones and other types of mobile devices.

David Lynch is one source of this critique. "Now if you're playing the movie on a telephone, you will *never*, in a trillion years, experience the film," he said in a video recorded for the DVD edition of *Inland Empire* (Lynch, 2007). "You'll think you have experienced it, but you'll be cheated. It's such a sadness that you think you've seen a film on your *fucking* telephone. Get real!" In 2003, years before the first iPhone was released, even Steve Jobs doubted that smartphones would be used for screening movies in the future. He candidly admitted he was "not convinced people want to watch movies on a tiny little screen" (Chen, 2010, para. 16). Although these statements were made over a decade ago, in a time of skepticism, they must be regarded keeping in mind Lynch's ironic and critical style of expression and without ignoring the fact that Steve Jobs eventually did make an obvious stand for mobile multimedia. During Jobs's MacWorld presentation about the new iPod edition in 2007, he live-streamed a short sequence of *The Office* (NBC) on a time-travel hack and then, played a few minutes of the feature film, *Pirates of the Caribbean: Dead Man's Chest* on the iPod, demonstrating the options for not only watching sequences but even changing aspect ratios. He introduced mobile-device-specific spectatorship to the audience's great amusement.

The unmistakable clash between optimizing feature films or television series to smartphones by the touch of a finger presented by a tech professional ("It works like magic!") and an artist's ultimate rejection of smartphone spectatorship is not exclusively about timing or technological developments; not even about a conservative advocation of the filmic medium. It rather concerns the assumed quality of mental and physiological processing of a movie and its narrative. More recently, the award-winner director of *Son of Saul*, László Nemes professed concern about his film being distributed on portable media platforms, outright calling smartphone spectatorship "the end of the world" (2016), thus privileging the physiology of the celluloid and the perpetuity of the cinematic experience.

The dystopian idea of film consumption rejecting artistic intentions in favor of a communication platform, where form, language, and all the aesthetic garnishing become nothing more than a crate for transmitting information, is certainly unsettling. Still, smartphone film spectatorship creates more opportu-

[51] For more details of technology adoption and its impact on smartphone spectatorship, see Chapter Three.

nities rather than being labeled threatening or disparaging to film art (see Byford, 2017). The legacy of cinephilia notwithstanding, this statement implies that participation has outgrown the movie fan's passionate interest in film art and cinema. Through the broadening arsenal of screens, spectators have received a reliable supply of tools for watching and interfering with the content simultaneously. In fact, portable smart devices have a great deal to do with establishing an alternative "film industry," where users are able to *create*, *access*, and *watch* films and videos, to *distribute* them through social media, to *browse* for additional information, and to *evaluate* contents—on the very same multimedia personal platform, at one touch. This means that anyone in possession of a smartphone can contribute to the wide palette of content by recording and sharing whatever they find relevant. Not least, users often assume the role of impromptu distributers and critics through various social media sites or other online forums.

By democratizing the individual interactive experience, smartphones have inspired not only consumers, but even professionals, such as Chan-wook and Chan-kyong Park (*Night Fishing* [*Paranmanjang*], 2011), Sean Baker (*Tangerine*, 2015), Jenna Bass (*High Fantasy*, 2017), and Scott Barley (*Sleep Has Her House*, 2017) to shoot films on smartphones or to make use of the aesthetic framework of mobile filming as Sally Potter did (*Rage*, 2009). Moreover, a great number of short and feature-length films, like *The Silver Goat* (Aaron Brookner, 2011) or *Roma* (Alfonso Cuarón, 2018),[52] have been distributed chiefly or even entirely on portable smart devices or streaming sites. The list is large and continuously expanding. However, there are spatial and temporal divisions in this trend, which separate films that were made with the intention or possibility of being released on portable devices from those that were not. This division is expected for movies made before the advancement of mobile distribution (most of film history). It is, however, more ambiguous in the case of contemporary releases, as I demonstrated through the example of *The Silver Goat* in Chapter One. Films shot on smartphones often proclaim the presence of the mobile device as an economic rather than an aesthetic feature. In other words, smartphones are more often used for decreasing budgets or for challenging the industrial framework of professional cinematography than for creating a new audiovisual language. There are notable exceptions, though, for instance, Jenna Bass's *High Fantasy*. In that film, the director uses handheld aesthetics (a shaking and fast-moving camera recording mostly from short distances) specifically to aid the

[52] Netflix acquired *Roma*'s exclusive rights for distribution: following festival releases and a brief (month-long) theatrical run, the film is only available on the streaming site since December 2018.

narrative and capture embodiment and intimacy when the characters wake up in the morning to find themselves trapped in each other's bodies.

Even though portable devices are often treated as unsuited or even detrimental to movie making and watching, these examples unquestionably elevate the discussion around mobile filmmaking from one of pure conventionality to that of artistic manifestations. Several artists have experimented with the limitations of the apparatus in the past decade, celebrating a *lo-res* culture (for some examples, see Fowler & Voci, 2011; Odin, 2012). But perhaps an equal number have exploited what modern mobile phones and their cameras can offer—completely obliterating evidence of a (quite intentional) lack of professional equipment.

In addition to the novel dynamics of mobile film and video industry, the options for simultaneous creating, transmitting, and watching disrupts users' roles, appointing them as a prosumers/produsers (Bruns, 2009), or viewsers (Harries, 2002). From another angle, the roles of production, distribution, and consumption merge in the domain of online content and convergent media technologies. Directed by the spheres of multimedia production and consumption, film and media experience become chiefly characterized as the flagship of information abundance, which results in an increasing level of stimulus and narrative intake. Despite of the disparagement of mobile multimedia, audiovisual content created, accessed, distributed, or consumed on handheld devices "enter[s] into a dialogue with the classic cinematic *dispositif* claiming the place of their compact cinematics in cinematic tradition" (Walden, 2017, p. 141). An interpretation of Walden's words is that, on the one hand, new media screens are still based on the cognitive processes rendered by the traditions of the cinematic apparatus (see Chapter One), and on the other, that a one-on-one media-archaeological comparison with cinema to define the genealogy of mobile movie platforms is simply missing an account of the present media landscape. The abundance of magazine articles and academic analyses of participatory consumption, the how-to-filmmaking tips, and online collaborative platforms[53] present a clear picture of the fact that there is a part of film industry that is actually based on participation and is shaped by the conflict between conservationist and more permissive agendas.

This point is well illustrated in Robert Rodriguez's *Two Scoops* (2013), made in partnership with BlackBerry and "You" (see Figure 9). The first version of the action-filled short film was shot and then presented online. Although the main storyline was complete, the film contained unfinished scenes and green screen

[53] For example, ScreeningRoom, a content-making community for independent filmmakers.

features and Rodriguez reached out to the public to participate in completing it. *Two Scoops* is about a pair of twin sisters, who run an ice cream truck during the day, but at night they search for their missing father and others kidnapped by a mysterious monster. One of the scenes that was to be completed by crowdsourcing depicts the twins' agent calling on a video phone to share information about their father's potential kidnapper. The agent appears on the convertible wall of the ice cream truck and on a BlackBerry tablet. The scene was filmed with a green screen to be later filled by self-recorded footage from any aspiring actor wanting to play the role of the agent. For this Rodriguez provided the script and suggested using anything from mobile phones to webcams. Apart from missing scenes and sections of dialogues, people were asked to help design the ludicrous monster and the weapons to fight it, as well as to send their photos to feature as missing people's pictures on the community bulletin and lamp posts. Public participation gave *Two Scoops* new aesthetic features.

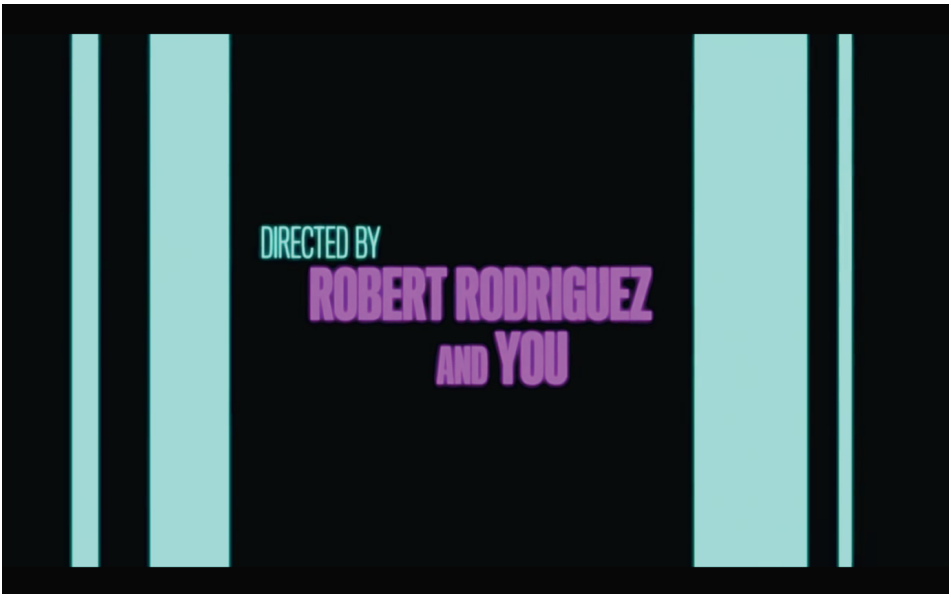


Figure 9. Director's credit at the end of *Two Scoops*, a crowdsourced short film. Screenshot from film.

Based on the results of interactivity, corporeality, and spatial presence, Engberg and Bolter (2017) define the aesthetics of mobile cinematics as polyaesthetic. Such a term captures the changing modes of representation which are catalyzed by the increased accessibility and participatory nature of moving-image culture.

This can be approached in two ways, from a sensory and from a representational angle. The first implies the previously discussed distortion of the visual field and synchronous attention to mediated and unmediated stimuli—which also is in the center of Engberg and Bolter's discussion of the immersive visual and sonic language used in VR films. The second approach introduces a transforming mediascape induced by the specifications of the most ubiquitous screening devices. I argue that these elements comprise both the aesthetic and cultural trends, chiefly because the formal characteristics of movies connect to the culturally constructed conventions of cinema, but also because online consumption, multi-window representation, the increasingly mediatized surroundings, and parallel activities all have an effect on film and video consumption and production.

When it comes to mobile video aesthetics, smartphones challenge even the basic cinematic terms of image capturing by often neglecting the aspect ratio that a horizontally held smartphone provides. *Wired* columnist Clive Thompson (2017) blames the ergonomics of the device. "It feels weird," he writes, to hold a smartphone horizontally (para. 4). Thus, while perhaps vertical images are no longer strange (if they have ever been), they clearly impede both image content and perception—especially in cases of content produced for horizontal screens. Since (unimpaired) human vision captures a landscape view, a vertical composition of moving images may cause an uncanny feeling and this bias in visual representation influences narrative comprehension too.

Charlie Lyne, in his video essay entitled *Frames and Containers* revisits Eisenstein's idea of the cinematic frame and its dimensions. Lyne's reflection on *The Dynamic Square* by Eisenstein vivifies a scheme for the plasticity of the film frame and the possible deviation of it depending on its container. Lyne's motivation is that, over eight decades after Eisenstein's lecture, this continues to manifest itself—this time in the abundance and flexibility of screening platforms that are yet "to embrace all the multitude of expressive rectangles" (Eisenstein & Leyda, 1982, p. 52). The idea behind the plasticity of the frame (even within a single movie) chiefly lies in filling the visual dimensions for the sake of storytelling and creating the illusion of immediacy. Thus, by the dynamics of visual narration, the spectator is detached from awareness of the screen's physical presence, while the frame as a tool for focusing gaze opens new opportunities for manipulating the immersive quality of the apparatus. The effective potential of visual representation, in other words, is remediated by technological development and consumer culture. Since his proposal, there have been numerous movies, video installations, and film projects to embrace Eisenstein's expressive

rectangles from *The Door in The Wall* (Alvey, 1956) to the Vertical Cinema project and *The Numberlys* (Blinkoff, 2015; see Figure 10), "the world's first *tall* short film," which was released alongside an augmented-reality app and a picture book with the same characters. What is specific about the case of smartphones, however, is that perspectival changes are ubiquitous, familiar, and spontaneous both while capturing and watching. This frames the blurring boundary between viewer and producer, amateur and professional, intentional and spontaneous, and between creation and consumption.

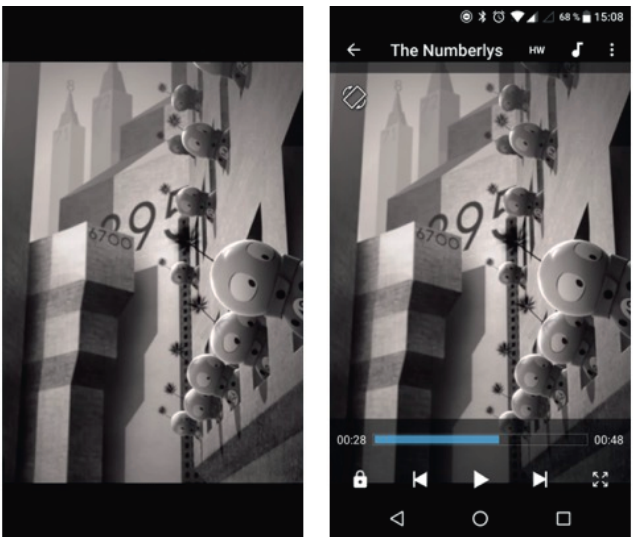


Figure 10. Screenshot of a still from the vertical animated movie, *The Numberlys* displayed in MX Player.

Considering the user's corporeal intervention, the urban aesthetics of *pocket technospaces* (Richardson, 2007) entail not only the emergence of participatory film and video culture, but also its diversion from the cinematic medium, which Manovich (2001) had foreshadowed as a colonizing force of identity and imagination, but which instead became the signifier of amateur versus commercial culture. As Bordwell (2009) notes in line with Eisenstein's aforementioned proposal, visual representation follows the technological framework of screening as opposed to the framework of filmmaking. But such disparities in reference to smartphone screens originate solely from its distance from movie traditions and promote the cognitive and phenomenological potentials of consumption. Thus, besides the actual domain of viewing experience outlined in the earlier sections of this chapter, two novelty factors play the leading roles in such a connective

sphere, namely, the inductive (such as, filmmaking, distribution, streaming) and the reflexive factors (as discussions, forums, fan culture). Mobile media content gained sovereignty as part of an institutionalized instrument of social, cultural, aesthetic, and technological components to hit cinemas and major film festivals and mobile platforms very near you.

PART 2 Technology Adoption and the Physiological Effects of Smartphone Spectatorship

Chapter III.

Neurocinematics and Portable Screens

Your Brain Wasn't Built for Movies

"Your brain wasn't built for movies," Jeffrey M. Zacks introduces *Flicker* (2015, p. 3), a study of the curiosities of moving images, which have hallmarked spectatorship since the time of the Lumière screenings and which, more than a century later, still signify cinematic spectacles. Indeed, it is movies that were built for our brains. Balancing between a viewer's and a neuroscientist's perspective, Zacks examines spectatorship building on Münsterberg's (1916/2014) analysis of film's affective capacity as a then-novel realm of psychological research. Münsterberg applied a psychologist's observations to early film screenings concentrating on comprehension and emotional and affective conditions. Through this, he explored human perception of audiovisual narrations, cherishing the aesthetic and technological independence of early 1900s screenings. Moving-images and cultural cues have changed greatly since then, along with the evolving narrative strategies and fidelity of representation. Münsterberg worked in the era marking the very beginning of cinema's independence from other theatrical entertainment. Zacks and his contemporaries, however, have access to sound and color film, as well as to an overview of how moving-image storytelling has developed new strategies for immersing viewers into a fictional presence. Zacks's opening statement is justified in the sense that people are able to contextualize and react to bits of information in movies without actively being aware of mediation. A person in a white coat with a stethoscope around his neck reflects on both the sensory and cultural domain of a doctor, irrespective of whether he appears on a photograph, filmed footage, a painting, or is standing in front of us in person. This is what our brains were built for: to interpret sensory and semantic information.

The question of comprehension can also be approached from the opposite direction: Sergei Eisenstein and Lev Kuleshov experimented with applying behaviorist theories to the development of filmic montage in the 1910s and 1920s. Their work lead to a conclusion that the function of moving-image narration (i.e., the formal organization of actions represented as moving images) is to provide a contextual framework for organizing semantic meaning. Along with others also contributing to the Soviet theoretical and filmmaking school of semiotic and dialectical storytelling methods, Eisenstein and Kuleshov advocated for the importance of contradictory shots to produce combined meaning. In his iconic experiment, Kuleshov demonstrated that juxtaposing images inclines spectators to base their understanding of them upon logical associations, even if those images (or in other cases, sonic cues) seem incoherent or astonishing. He presented a sequence in which the same footage of actor Ivan Mozzhukhin's motion- and emotionless face alternated with sequences showing a bowl of soup, a dead child, and a seductively dressed woman. Kuleshov observed that subsequent images affected viewers' perceptions of Mozzhukhin's character's emotions: through unconsciously connecting him with the rest of the footage, viewers saw evidence of hunger, grief, and lust, respectively, on his face.[54]

In light of the above statements, the synchrony between mental and bodily processes (nature) and learned skills and cultural knowledge (nurture) provides a foundation for processes while watching movies. My aim in this chapter is to first examine the evolution of skills related to spectatorship in general to, then, analyze the ways those skills become applicable to the changing patterns of post-cinematic moving-image consumption. This calls for a cognitive approach that connects spectator, content, and the precognitive and cognitive effects of moving images. A cognitive approach, here, aids in assessing the causal links between theoretical inquiries about storytelling and comprehension and empirical evidence from cognitive neuroscience, psychology, and associative and non-associative learning. Embarking from theses of how moving images stimulate semantic and contextual connotations, I will evaluate the roles of knowledge, as well as structural and audiovisual composition in comprehension. These will serve as building blocks for defining the cognitive, bodily, and behavioral

[54] Mobbs et al. (2006) remodeled Kuleshov's experiment and monitored brain functions through functional neuroimaging (fMRI) to confirm the role of contextual framing, that is, the role of associations in cognitive processing. By presenting subjects with images of men and women juxtaposed to emotional or unemotional sequences, they concluded that context affects socially coded judgments of emotional states and viewers label emotional impressions according to their contextual frames.

elements involved in learning smartphone-specific spectatorial formulas and, in the next chapter, for assembling a methodology to analyze empirically smartphone spectators' comprehension and behavior.

The modal nexus between movie narration and human cognition can be discovered in the absence of a duality of innate and learned skills, using examples from inexperienced viewers' narrative comprehension. Discontinuities in movies—such as jump cuts that connect sensory information in Kuleshov's demonstration—serve as formal links that organize and structure meaning. But, as Stephan Schwan and Sermin Ildirar (2010) demonstrate in an experiment, shifts that cuts induce appear rather like boundaries for those lacking knowledge of moving-image narration. Schwan and Ildirar observed the relevance of empirical knowledge in interpreting transitions in a narration to confirm that spectators are in fact incapable of understanding these transitions without having acquired sufficient experience with photographic image sequences (e.g., movies or television content). They showed short clips containing a simple plot each and one of the seven most typical narrative transitions (e.g., point-of-view shot, establishing shot, shot/reverse shot, crosscutting) to inexperienced adult subjects.[55] The clips were made specifically for this purpose and contained no references to popular culture. After interviewing each subject, they concluded that, without sufficient knowledge, subjects were capable of interpreting schematic information clusters such as people, objects, or places, but failed to detect spatiotemporal connections interrupted by cuts. For instance, after watching a footage which presents a house from the outside and then cuts to an interior shot, Schwan and Ildirar's inexperienced subjects were unable to tell that the interior belonged to the house if the depicted building or room was unknown to them.

Moving images have gained cultural autonomy as part of a storytelling medium and acquired specific structural and formal attributes, cultural codes, and industrial practices through methods for spatially and temporally organizing stories (Bolter & Grusin, 1999; Carroll, 2003). Narrative composition guides spectators' attention—directing it toward crucial and meaningful pieces of information. Moving-image narration also has a formal framework that is based upon the mental and physiological mechanisms involved in interpreting events and stories: these are the cognitive processing of cultural schemata through

[55] Schwan and Ildirar recruited participants from a rural area of Turkey where people lived in relative isolation and had no access to electricity or television sets. Although they were in possession of some photographs, none of them had previously seen moving-image recordings. Their responses were compared to participants with similar cultural backgrounds but some or substantial viewing experience.

previously acquired knowledge and precognitive mechanisms, such as reflexes, attention, and "the path of the eye" (Eisenstein, 1947/1975, p. 190). Eisenstein's words in *The Film Sense*, a work on composition and perception, confirm this remark as he reflects on the author's role in guiding the perceiver's senses along predetermined paths. What Eisenstein emphasizes is that the visual composition of narration informs, engages, orients, and even manipulates the spectator's comprehension and physiological reactions, such as gaze patterns. Moreover, viewers can only see a rather small portion of the frame at once, which means that elements containing important textual information must draw attention. This is of particular importance in the case of media deficiencies, for instance, distraction or in out-of-the-ordinary screening conditions associated with, for instance, mobile screens, virtual reality goggles, or unenclosed projection surfaces.

Murch (2001) and T. J. Smith, Levin, and Cutting (2012) provide another perspective on moving-image narration, namely, that the correspondence of storytelling and cognitive processes makes movies an appropriate tool for modeling information projection when studying cognition. This leads me back to the above proposed harmony between the cinematic illusion and the perception of real-world stimuli. The illusion of nonmediation, that is, the sensation of being physically present in a virtually projected sphere, lies in the fact that the mechanisms used to access moving-image narrations correspond to perceptions of reality. This means that, when immersed in the diegetic space, experienced spectators comprehend and respond to the depicted actions much the same way as they do to real-world ones, even if those actions are presented as stylized by editing or by the continuity system (T. J. Smith et al., 2012). Movies and screening media maintain the immediacy of content by storytelling methods developed in relation to human perception and cognition. Zacks (2015) clarifies, "our brains didn't evolve to watch movies: Movies evolved to take advantage of the brains we have. Our tendency to want to respond physically to them highlights this" (p. 4).

There are a number of legends about early cinema screenings in which frightened spectators recoiled, screamed, and even fled of the screening room. These stories presumably exaggerate reality—not to mention that films' fidelity was far from real-world and thus screenings were far from being truly immersive. But they nevertheless morphed into legends on account of the plausibility that inexperienced viewers in the late nineteenth century could potentially react to photographed actions and objects as they would to actual ones (see Gunning, 1995), where the mediated stimulus was inconsistent with

the spatial and cultural framework of social get-togethers. Precisely as works of art featuring optical illusions created with the *trompe-l'oeil* technique can deceive observers, objects depicted on a flat screen can appear as if they exist in a three-dimensional space.[56]

An oft-told story, and one of the most popular fables about the first years of cinema, features the Lumière brothers' *The Arrival of a Train at La Ciotat Station* (1895). Unlike in the case of trompe-l'oeil paintings, the train not only appeared on the screen, it also moved toward the audience! In reality, if a large locomotive comes racing headlong toward someone standing on the platform, her brain would calculate its approximate point of arrival and command the necessary muscles to perform actions for avoiding collision. Simply because of the knowledge of apparent motion and perspective, neural and motor reflexes can become activated even in response to fictional events (see also Bordwell, 1985).

The British short satire film, *The Countryman and the Cinematograph* (also known as *The Countryman's First Sight of the Animated Pictures*), directed by Robert W. Paul in 1901 and the Edison Company's American remake, *Uncle Josh at the Moving Picture Show,* by Edwin S. Porter in 1902 also provide examples that illustrate the illusion of nonmediation. The illusion arises from the viewer's alignment with the camera perspective, which evokes the countryman's or Uncle Josh's astonishment at cinematic screenings. These early (perhaps two of the first known) examples of dramatized *films within films* present the over-enthusiastic viewer, who engages with the characters and objects on screen (a dancing woman, a locomotive in reference to *The Arrival of a Train*, and a couple in a yard) and reacts as if what he sees is reality. He climbs onto the stage to dance with the dancer, runs and hides from the approaching train, and acts out his jealousy by trying to engage in a fight with the farmer on the screen (who, in the British version, is in fact himself). This leads to him tearing down the canvas and ending up in a fight with the projectionist. Through the caricature of the unsophisticated countryman's first encounter with the sensation of cinema, the two movies dramatize the audience's deception in a cinematic screening.

Such reactions are not necessarily far from reality, if we consider the affective qualities of movies. The perceptual illusion illustrated above is greatly dependent upon the viewer's sense of her position relative to the screen. Explained in Chapter One, proprioception plays a key role in estimating and adjusting bodily configuration to surrounding objects, even mediated ones. If

[56] A popular anecdote to illustrate the effect of trompe-l'oeil paintings is the story of the contest between two Ancient Greek painters, Zeuxis and Parrhasius: Zeuxis's still life of a bowl of fruit deceived a bird who mistook the painting for actual grapes and Parrhasius's painting of a curtain deceived the entire audience and his fellow painter.

the screened images fill a significant proportion of the spectator's visual field (the screen's size is larger than her body), precognitive reactions are more likely to cause a change of some sort in bodily posture. These movements, however, are possibly regulated by cognitive evaluation that reminds the spectator of the illusion of the movie. Yet, as Zacks (2015) recalls, even seasoned, twenty-first century spectators can be found recoiling or ducking when watching the Jabberwock's falling head in Tim Burton's (2010) *Alice in Wonderland*.

Is Film Viewing a Skill?

My scrutiny of the perceptual and cognitive processes related to spectatorship takes two directions before I engage in the details of the movie spectacle's captivating qualities and their manifestation on smartphone screens. First, I introduce the factors involved in recognizing and schematizing sensory information, which will be followed by an analysis of the broad narrative context that defines the elemental goals and strategies of spectatorship. Departing from narrative and cultural schemata and audiovisual stimuli, the aim of this composite overview is to demonstrate that watching movies requires specific skills and to identify the learning process for obtaining such skills.

As specified earlier, factors such as unenclosed viewing environments and smartphones' technical specifications that enable interactions may significantly impact viewing experiences. Hence, I presume that it is the network of behavioral and material elements that defines individual responses and movie and video consumption in general. In this section, I persist to argue that any momentum in the evolution of movies and screening platforms is closely connected to the human factor, and more specifically the spectator's brain and senses. The segments in the process of sophisticating moving-image storytelling signified by the countryman, Uncle Josh, as well as by Schwan and Ildirar's subjects reflect on the role of movie literacy, which will be examined in the following pages through the lens of the involved bodily, neural, and cognitive processes.

When it comes to labeling, Noël Carroll (2003) argues for the transmediality of movies, claiming the name, *moving images*, to describe the complex phenomenon that unites anything from various data media, styles, and forms to technological evolution. This side note, with which Carroll opens his essay collection, *Engaging the Moving Image*, is a crucial focal point here in establishing the theoretical outline of stimulus comprehension. Not only does it acknowledge

the manifold cultural legacy of what we can call the indexicality of the photographic image, but also, even more importantly, such terminology frees aesthetic practices from their containers (for instance, a screening medium) and bounds or likens them to exposure of animation in the physical world. This claim highlights a fundamental methodological path for studying new media spectatorship, including on smartphones. Yet, Carroll's argument about terminology loses its importance in the torrent of digital images. What is more, the term *film* now signifies a specific formal structure (as opposed to video) rather than only a data medium (as opposed to digital image).

Useful however is Carroll's identification of the container medium as an entity that legitimizes and organizes content—as did cinema with the film reel. This identification supports an understanding of the evolution of moving-image storytelling into a systematized form of communication with facilitation from the mechanisms of the human brain. This aligns with my earlier argument (and the main objective of this dissertation) and what Zacks (2015) refers to when recounting the semantic connections between mediated and physical-world stimuli: viewers perceive patterns of information on the screen as they perceive stimuli from real-life sources and it is the context that defines their relevance. Even though it seems dissonant at first glance, Carroll's defining the role of screening media as vessels for moving images justifies the distancing of audiovisual materials from the context of the screen. This provides a suitable way to analyze how visual and sonic cues guide spectators' attention and how their juxtaposition generates specific social contexts. Understanding the elemental communication channels between stimulus and perceiver leads to conclusions regarding the exact processes involved in adopting media-specific knowledge and practices.

Visual and sonic representations in movies aid comprehension only if spectators are aware of social conventions and aesthetic traditions. The very first step in achieving this requires the ability to process narratives of varying complexities using knowledge of sensory concepts such as motion and perspective. At the dawn of behavioral studies, Fritz Heider and Marianne Simmel (1944) conducted a substantial series of experiments to investigate the perceptual and response mechanisms used to understand the dynamics of social interactions. The specific answers their experiments were seeking concerned the interpretation of visually available information in a social context. Heider and Simmel aimed to explore the process in which an observer decodes configurations of visual stimuli as behavioral patterns, specifically in cases without explicit social cues, such as facial expressions, sound, or speech.

In the experiments, Heider and Simmel used short stop-trick sequences in which a circle and a large and a small triangle moved at different speeds in and around "the house," a rectangle with an opening on one side. This made sensory information limited to the movement of objects in relation to others. Heider and Simmel's anthropomorphized description of the scenes depicted by the three geometric shapes explain actions as "[the large triangle] moves toward the house, opens door, moves into the house and closes door," or "[the large triangle] seems to try to get out of the house but does not succeed in opening the door: [the small triangle] and [the circle] move in circles around outside of the house and touch each other several times" (1944, p. 245).

In each of the three experiments, participants were instructed to pay close attention and then interpret what they had seen. The first group of participants simply had to tell what they thought happened in the footage; the second was instructed to explain the actions as if the "characters" (the geometric shapes) were human; while the third group had a similar task as the second but was shown the clip in reverse. The questionnaire in the second (main) experiment contained questions about the shapes' personalities ("What kind of a person is the little triangle?"), behavior ("Why did the two triangles fight?"), and their actions ("What did the circle do when it was in the house with the big triangle?") (Heider & Simmel, 1944, p. 246). Heider and Simmel reported that nearly all participants in each group perceived the geometrical shapes as animate beings (most even as sentient beings) and constructed much the same coherent story out of the actions. Especially notable is the correspondence with which participants attributed the shapes with personalities and intentions in experiment two: in a statistically significant agreement, most participants described the large triangle as a strong and aggressive male (as it chases and hits the two other shapes and wins fights), the small triangle as courageous and brave (it hits and rapidly chases back the large triangle), and the circle as a frightened, but also resistant and gentle female (as it escapes and also closes the door locking in the large triangle).[57]

Motion is a social indicator that presumes intention and provides information about perspective (hence, the cinema audience alarmed by the train roaring along at top speed) and behavior (as the shapes in Heider and Simmel's experiments). Motion in abstract representations and moving images has the same role as in the physical world: a body in movement serves as a cue and provides

[57] Heider and Simmel's study highlights the impact of motion as a social cue. In a recent revisiting of the original study, D. S. Berry, Kean, Misovich, and Baron (1991) ran a similar test, only emitting motion from the sequences to demonstrate that perceiving geometrical forms as characters with intentions is derived from their movement.

information about intentions, social connections, and the spatial attributes of a person's actions. The process that leads to accessing this information is based on object recognition and on comprehending spatial relations by considering mechanical, intentional, and mentalistic agency, as well as an object's self-propelled motion and its reaction to external forces (Simion, Bardi, Mascalzoni, & Regolin, 2013).

Even though they emerge from concrete, physical experience, the set of skills needed to interpret visual information in a context with additional sonic and verbal cues play a significant role, even in the case of movies. The reason for this lies in the technical mechanisms underlying moving images: in the process of remediation, film inherited the indexicality of photography, and consequently, photography's direct perceptual relationship between object and interpretant. The direct reference of audiovisual content to physical reality may explain the analogy between perceiving unmediated information and movies. However, this theory neglects the question of screening media (including fidelity, visual angle, stimulus intensity, etc.), as well as the manipulations that pertain to edited and digitally created sequences, both of which will be addressed below. To this limitation a semiotic approach offers a solution, namely, that representation of an object (index) reflects on the presence of this object in a particular spatial and temporal framework, and not on the object itself. The process of interpretation requires recognition of different schematic units: an object (e.g., a train), its motion (it is approaching), and the relevant socially conditioned reaction (one should act to avoid being run over).

Departing from semiotics and the phenomenology of representation, one can approach the question of visual perception in a similar manner.[58] The relation between visual and cultural information is intertwined. This notion is clearly traceable in the experiments of both Heider and Simmel (1944) and Schwan and Ildirar (2010), where a context was either added to or removed from their participants' subjective understanding of what they saw. The ontology of vision can relate to narrative styles and material, technological, and art historical frameworks of representation. A formalist approach implies that the Kuleshov effect or Münsterberg's and Eisenstein's theories legitimately emphasize the presence of cultural and subjective layers of connecting sensory information to social and cultural formulas. Zacks and Magliano (2011) propose a similar thesis in their cognitive–neuroscientific approach to moving-image narration. They claim that spectators perceive a temporally and spatially coherent narrative of

[58] See, for instance, Vlad Ionescu's (2014) work in which he uses indexicality to explain that imagery cannot simply be rid of complex cultural meaning.

juxtaposed, but seemingly disparate shots by organizing them into information or event clusters that also serve as basis for predicting upcoming actions.

Capturing Attention

The common substantial ground for culturally coded narrative comprehension in spectatorship propels recognition, associations, neural processing, and these are all based on attention. To approach the question of attention, I employ two perspectives. First, a subjective perspective, which regards attention as one of the mechanisms that filter and amplify given semantic indices; namely, context-specific information related to characters, objects, and spatial features. The second perspective is compositional such that attention is part of cognitive processing and guided by the arrangement of discreet clusters of sensory information on a screen or in a soundscape.

Social knowledge guides viewers in seeking relevant patterns on screen or in sound. This implies that filmmakers can drive their attention according to the same logic. Lacking definite compositional elements and consequently specific attention-control methods, natural scenes capture attention according to social relevance. Motion, sudden onsets, changes in the visual or sonic spectrum, and similar elements can provide valuable information about a beholder's relation to the surrounding space and inform her about potential sources of interaction or danger. Natural scenes (e.g., footage of undirected passersby shot in an urban environment) evoke varying responses among viewers.[59] Accordingly, verisimilar cinematic techniques used in realistic compositions, such as deep focus, long takes, or open frame filming, induce a significant ambiguity in viewers' attention, and, thus, in their physiological and neural responses and narrative understanding. The more the composition of a movie frames and centers upon certain visual or sonic semantic elements (for instance, faces, speaking characters, vivid colors), the more unequivocally it grabs attention—introducing determined and highly congruous responses from the audience.[60]

[59] This theoretical premise is widely addressed by Bazin (1967) and Bordwell (1986), among others, and implemented in experiments run by Hasson, Nir, Levy, Fuhrmann, and Malach (2004), Hasson, Landesman, et al. (2008), Mital, Smith, Hill, and Henderson (2011), T. J. Smith and Mital (2013), among others.

[60] Such congruity was observed in feature films and television programs as opposed to natural scenes. See, among others, Hasson et al. (2004), Hasson, Landesman, et al. (2008), T. J. Smith and Henderson (2008), Mital et al. (2011). Consult T. J. Smith (2013) for an overview.

Edited and composed footage employs exogenous (externally compelled) control in viewers, whereas the absence of directing, as in the case of natural scenes, yield endogenous (internally compelled) control, which effectuates unpredictable and widespread focus. Exogenous and endogenous factors define the objects to pay attention to either by composition, stimulus properties (e.g., stimulus magnitude, as visual saliency), the observer's knowledge, interest, and intentions, or the attributes (e.g., speed) of selection (Theeuwes, 1994; Lauwereyns, 1998).

Stimuli accessed through exogenous and endogenous mechanisms are processed using different strategies, both in cases of stimulus-driven (bottom-up) and task-driven (top-down) processing. Bottom-up processing occurs when sensory information enters through the sensory organs and is recognized and interpreted subjectively through the multiple layers of the perceiver's knowledge and experience. In cases of exogenous control, sudden onsets or stimuli with high magnitude, such as salient objects, vivid colors, or loud noises, draw involuntary attention to certain objects. In cases of endogenous control, familiar objects activate certain types of responses; these could be, for example, pop-culture references which appear in movies in the form of objects or written texts. The other trajectory, top-down processing in movie spectatorship can be induced by specific tasks to, for instance, gain information about the color of a piece of clothing a character wears, but narrative hypotheses also induce top-down processing. Generating hypotheses regarding the outcome of an action or even an extended sequence activates search tasks for relevant (confirming or denying) information (see Bordwell, 1985). A search task (processed from the top down) overrides involuntary attention to low-level features, such as saliency or loudness.[61]

[61] Viewing tasks (e.g., to look for certain information in a sequence) also induce knowledge-driven spectatorship, meaning that a specific assignment could shift the viewer from an externally compelled exploratory mode to an internally compelled controlled mode of spectatorship. Christopher Chabris and Daniel Simons's selective attention experiment, which became known as "the invisible gorilla" test (Simons & Chabris, 1999) and "the monkey business illusion" (Chabris & Simons, 2010) exquisitely demonstrates attention shifts induced by specific tasks. In videos widely available on various video-sharing sites, people dressed in white or black t-shirts pass balls to others wearing the same colored t-shirts. The viewers' task is to count the number of times those in white pass their ball. Counting passes from one team member to another captures viewers' attention, masking other visual details, namely, that a person dressed as a gorilla walks into the frame, bangs on its chest, and walks away. Failing to capture attention because the viewer is looking for task-relevant information, the gorilla actor's presence (whose costume is the same color as those of the non-relevant group of players) goes unnoticed (Levin & Simons, 1997; Simons, 2000).

In his paper analyzing the impact of exogenous and endogenous control over attention, Johan Lauwereyns (1998) names four distinct strategies of visual selection based on the controlling function of optical information. Besides external and internal control, his division introduces the variables of space-based and object-based attention. This suggests that the direction of visual attention depends upon comprehension of ecological coherence. Lauwereyns's review of the then-existing paradigms recognizes the need and locates the possible methods for defining the role of expectation as a catalyst for paying attention to objects. Although Lauwereyns's work is limited to visual selection, which is likely augmented by other sensory modalities (at least sound) in the case of moving images, his method corresponds with my argument and initial steps for exploring the acquisition of smartphone-specific spectatorial behavior, more specifically the question concerning the viewer's differentiation between relevant and irrelevant information. Lauwereyns's work provokes me to make three assertions. First, in the spatial dimension, image composition directs the spectator's attention to specific regions of the visual field. Second, in the dimension of objects, a stimulus source—an object—contains units of information (for instance, adjacent surfaces with the same texture or color) which help understanding its spatial configuration. Third, exogenous cues are recognized faster than endogenous ones and they attract attention by their low-level qualities (e.g., luminance, color, magnitude, etc.) irrespective of semantic meaning. These insights contribute to my upcoming analysis of distracting extra-filmic stimuli originating in the spectator's surroundings while watching movies or videos in unenclosed spaces.

In spectatorship, exogenous control of attention reflects the filmmaker's or addresser's intentions in governing the location, size, and juxtaposition of semantically meaningful elements. Controlled sensory exposure results in an increased level of synchrony in viewers' visual and sonic attention and narrative comprehension. Exploring visual storytelling in movies, a study by Carmi and Itti (2006) reveals that frequent cuts increase the level of synchrony among viewer's attention. They attribute this to editing, since editing controls the visual outline and the time one can spend exploring the content. Endogenous control, by contrast, is volitional, and therefore attention is immensely dependent upon the observer. Consequently, disparities in spectators' personal interest decreases synchrony in terms of attention. In line with Lauwereyns's (1998, revisited in Lauwereyns, 2012) division of attention trajectories, the mise-en-scène in a movie or other moving-image content guides the spectator's attention in a space-based manner by—literally and semantically—framing the area that

contains meaningful information. Visual composition (again, the outline of salience, vivid colors, or pop-out visual elements) directs the gaze in such a way that attention will likely be paid to the element with the greatest stand-out quality. This is equally applicable to endogenous control, where spectators monitor the content of the frame for relevant or familiar objects or object-clusters, based on their own intentions or expectations. The latter, however, is not only understood in terms of top-down processing, namely, that specific hypotheses guide the search for pertinent stimulus sources. As explained above, expectations and familiarity can also induce bottom-up processing of schematic information referring to social concepts or even previous moving-image experiences.

Orienting attention to the most relevant details means that other stimuli or pieces of information are selectively perceived. When watching a movie, attention is paid to a limited number of stimulus sources. But, according to the above logic of spatial and object-targeted attention-capturing mechanisms, sudden new stimuli, such as changes in the visual field or unexpected noises, are likely to induce a response. This can be a motor response that focalizes sensory receptors on the source or one that activates physiological changes. The amplitude of these responses, however, decreases in cases of continuous exposure (Ravaja, 2004). This leads me to assume that if a new stimulus is semantically related to the narrative, it will induce epistemic links to it. If, however, it is unrelated, it interferes with narrative comprehension by either adding an additional (unintended) meaning or by drawing attention.

Zacks (2015) and Hasson, Furman, Clark, Dudai, and Davachi (2008), among others, suggest that, since movies guide spectators through the same cognitive and emotional processes as physical-world stimuli, they use social behavioral patterns to make sense of incoming sensory information. Taking the above discussion into account, I can say that the goal of spectatorship is identifying sensory information possessing coherent meaning and clustering it into schema blocks based on physical and cultural indices, such as object attributes, motion, spatial configurations, and semantic meaning. Images, sound, and other related or unrelated stimuli in spectatorship are connected or—in the case of distraction—disconnected by these factors.

Narration and Embodiment: I Sense What Others Do

The heading, "I sense what others do," is a translation and paraphrase of a quote by experimental psychologist, Albert Michotte ("je sens ce que l'autre fait," 1953, p. 88), who studied the ecological processes of perception in the mid-twentieth century. Sensing what others do (*doing* indicating both sensing and physical doing) is fundamental to spectatorial activity, to adopting corporeality in relation to the diegetic world's spatial features and establishing an ecological relation with a virtually created space. The embodied qualities of perception are equally grounded in the human body's sensory motor mechanisms (for example, proprioception), the cultural domains that surrounding objects represent, and the mechanics of moving images and filmmaking. Below, I continue by reflecting on the propositions made in the previous section, and analyze whether establishing a situated ecological connection with the diegetic space impedes attention to stimuli originating in the viewer's surroundings. To address this question, I reach for a set of theoretical tools and start by painting a picture of narrative engagement, founded on immersion, empathy, and embodied simulation.

In moving-image spectatorship, content, percept, and context intertwine as the spectator immerses herself in the depicted world. This is due to the epistemic honesty of movies and video recordings, which are capable of presenting a fictional reality in the form of sensory truth. Such an epistemological view of spectatorship opens two paths: one is sensory unambiguity, while the other concerns the contextualization of narrative actions. By sensory unambiguity, I refer to visual and sonic truthfulness or verisimilitude, the direct indexical reference to concepts accessible from other mediated and unmediated sources. Contextualization is the process of organizing these pieces of sensory information according to their meaning. The two paths align as the spectator interprets and engages with the movie.

Moving-image narration can "transport" (Gerrig, 1993) or "absorb" (Bálint & Tan, 2015) the viewer into a logically coherent sphere of narrative actions, which momentarily supersedes and masks the rules of the physical world (see van Laer, de Ruyter, Visconti, & Wetzels, 2014). Sensory truthfulness can increase the likelihood of this immersive experience, even if the movie's diegetic space operates with seemingly unrealistic rules. Science fiction movies or historical period dramas, for instance, often present a narrative world that conflicts with the laws of physics or certain operative elements of our present-day societies. In the film *eXistenZ* (Cronenberg, 1999), characters can escape

reality into a virtually created parallel universe, where they exist and act according to specific game rules. Players reach *eXistenZ* through so called "bioports" surgically implanted in their spines, which are their only physical (biological) connection to reality. The main characters, Allegra Geller and Ted Pikul, are immersed in the game's reality up until the point at which they are confronted by external forces, namely, an infection in Geller's bioport. The spectator's immersion in the diegetic space resembles a similar association–dissociation parallel: the feeling of presence lasts until the point when external stimuli (distracting noises or visual effects, for instance) or internal motivations (boredom, disgust, mind wanderings) disconnect the spectator from the narrative.

In *Experiencing Narrative Worlds*, Richard Gerrig (1993) aptly labels this phenomenon narrative transportation, which he uses to describe written literature, but which is equally applicable to any form of storytelling.[62] Observers achieve the sensation of being transported into a diegetic world by establishing an ecological connection to the fictional sphere and empathizing with its characters. The process and outcomes of narrative transportation resonates with the provisions of Noël Burch's (1982) *diegetic effect* generated by moving images. The diegetic effect theory assumes the spectator's identification with the diegetic space and an emotional connection to actions and characters, but only if she has access to the necessary semantic cues and pays continuous attention. Moreover, whereas narrative transportation presumes the existence of complex narration, Burch argues that the diegetic effect can be achieved even in the absence of classical (feature or documentary film) storytelling. Introducing yet another angle in *Emotion and the Structure of Narrative Film*, Ed Tan (1996) claims that the monocular perspective (also used in paintings and photographs) is the determinant of the diegetic effect. Tan explains that the monocular perspective "draws the beholder in a position that is defined in relation to an imaginary space behind the window formed by the picture plane and the frame" (1996, p. 52). This requires the perceiver to adopt the artist's or addresser's perspective when observing the diegetic space. The prerequisite is sensory illusion, some sort of a fantastic experience, like the one that perhaps frightened audiences in the anecdotes encircling the Lumières' screening of *The Arrival of a Train* (1895).

These concepts combine a formal approach to moving-image narration (representation) with the sensory illusion of being there (identification). But approaches that attempt to explain the illusion of presence through comprehending a broad narrative context seem to fall short of one crucial point, that of how to access the elements of the sensible diegetic space while watching audio-

[62] For movies, see, for instance, Bezdek et al. (2015), Bezdek and Gerrig (2017).

visual content. The diegetic space contains objects and characters which move, behave, communicate, and have roles and functions. Viewers must grasp these roles and functions, map and memorize their locations and spatial relations, and acquaint themselves with their physical characteristics. Vittorio Gallese and colleagues' neuroecological and neurophenomenological paradigm offers a solution. Gallese (2005; see also Gallese & Sinigaglia, 2011; Gallese & Guerra, 2012) introduces the concept of *embodied simulation* as an applied version of the mirror neuron paradigm to connect neural processes of engagement to the sensory representation of a diegetic space. Whereas the mirror neuron paradigm describes the process of identifying others' actions and responses on a neural level, embodied simulation reflects on the motor capacity of this identification with external agents.

Similar to Heider and Simmel's (1944) conclusion (from decades before the discovery of mirror neurons and mirroring mechanisms in the early 1990s), one of the basic theses of embodied simulation is that the observer anticipates and retrodicts others' actions and intentions from their bodily motion (Gallese & Goldman, 1998). Gallese (2003, 2005) also recognizes that the process of identifying these actions and intentions is nested in the observer's body. Thus, embodied simulation implies that action and representation equally concern sensory involvement and the correspondence between acting and sensing, even when observing fictional characters. The observed body has a phenomenological similarity to the observer's own body and the human (primate) emotional map, which promotes survival and the understanding of others.[63] Affective participation is induced by the synesthetic and kinesthetic qualities of spectatorship; the fact that the spectator imagines herself in a character's position and observes and senses the diegetic world through such an affective channel. Hence, she experiences the virtually presented diegetic space by perceiving a character's position as her own and sensing the elements of the diegetic space as if those belonged to her own environment. Mental representation of the diegetic space provides information about social and emotional interactions, evokes memories of certain sensations (e.g., the touch of a surface), which in the physical environment, would inform biomechanical interaction.

In their work on embodiment applied to moving images, Gallese and Guerra (2012) acknowledge intersubjectivity as the propeller of transitions between fictional and physical worlds. They examine the qualities of stimuli that evoke motor activations in the spectator's brain as a response to, for instance, the sight of a manipulable object. As the term embodied simulation implies, classifying

[63] For a phenomenological approach, see Michotte (1953) and D'Aloia (2012b).

objects' manipulability activates motor functions related to corresponding manipulations (e.g., touching, grasping, lifting, opening, etc.) in the observer's brain. If an object affords grasping, the neural network responsible for the act of grasping activates.[64] Manipulation schemes also evoke memories of corresponding sensations that connect different sensory modalities: in the case of moving images, vision and sound can evoke even tactile and olfactory memories. Spatial information, as Gallese and Guerra (2012) state, is premised upon the integration of different sensory modalities and sensory and motor reactions, which I will address later as one of the bases for analyzing the spatial integration of visual and sonic distraction.

The processes of interpreting characters' actions and reactions to objects and other characters in a movie involve additional key elements, namely, empathy and subjective emotional engagement. Immersion in a fictional narrative includes empathizing with its characters, assessing the emotional motivations of their actions and reactions, and embedding them into the topography of the viewer's own emotions. The common ground of comprehension and emotional engagement through embodied simulation can be well described by the mechanisms of perception, emotion, cognition, and motor action—what Torben Grodal identified, linked (1997), and abbreviated (2006, 2009) as the widely acclaimed PECMA flow model for visual aesthetics. These elements, according to Grodal, are the neural features which lead the tracing of "salient forms in the chaos of information that arrives through the eyes," and by which, "the brain receives a small emotional reward every time it discovers a significant form" (Grodal, 2006, p. 4). As Coëgnarts (2017) argues, this process benefits from embodiment and embodied simulations, as increasing intensity of stimulation means increasing activity in the limbic system, the brain's emotional center, too, which provides emotional rewards.

The PECMA flow model rejects the thesis that spectatorship (or even film analysis) solely relies on cultural formulas, an approach deeply grounded in the exclusivity of humanities-based traditions in the history of film studies. Grodal (2006), instead, emphasizes a general perceptual and emotional scheme with a root both in psychology and aesthetics. What he proposes is a "Grand Theory" (see also, Grodal, 1997) that establishes and applies a framework for a "law" of aesthetics, similar to the laws of nature or the laws of the universe. Grodal proposes this as a possibility because of the prevalent statement that perceiving

[64] The spatial constraints of embodied simulation recall Shaviro's (1993) phenomenologically oriented note in *The Cinematic Body*. There, Shaviro emphasizes the prosthetic qualities of cinematic sensory experiences and the existence of affective synergy between the physical body and virtual sensory experiences. See also Sobchack (2004).

filmed images corresponds to the perception of the surrounding physical space. This implies that separating the two different contexts requires higher level computing and the identification of something like "reality-status markers" (2006, p. 3).

To extend my argument concerning subjective experience and the physiological and social responses of the spectator, I move beyond the PECMA flow model and borrow Nico H. Frijda's (1986) approach to the embodied nature of cognitive processing. Frijda advocates for an understanding of emotional processing (the process connecting the exposure to a stimulus to a given reaction) as a system that consists of subcomponents for evaluating incoming stimuli in terms of context, relevance, seriousness and difficulty, and urgency. Frijda divides this process into computational mechanisms, namely, the *analyzer* (decodes the stimulus as an emotional input), the *comparator* (evaluates its subjective relevance), the *diagnoser* (situates the possibilities for action), the *evaluator* (classifies the stimulus event in terms of urgency and difficulty), and the *action proposer* (creates an action plan) (see Frijda's figure 9.1 of the emotion processes, 1986, p. 454).

Since the relationship between spectators and (fictional) characters is based chiefly on empathy, instead of on personal connection, the mechanisms responsible for encoding realness, relevance, and feasibility—components that determine the intensity of the emotional reaction—in Frijda's model are replaced by identifying with characters and action schemata. Spectators hardly have the exact experience as the protagonists of *eXistenZ*, yet they may be able to relate the actions on the screen to their own experiences of love, affection, fear, gameplay, immersive experience, and similar scenarios. This implies that the brain's associative networks connect incoming fictional information with memories of physical, real-life experience and emotions that correspond to narrative events. Moreover, memories of objects sensed through embodied presence (e.g., objects the characters touch) and sensations connected to them facilitate the evocation of the relevant set of sensory memories that help to establish an ecological connection to a fictional character's position in a diegetic world.

Think Big, Start Small: Learning Smartphone Spectatorship

Two men approach across a misty, withered field. In the background, castanets and a slow string theme conjure Spaghetti Westerns, a dramatic epitome of mystery augmented with the deep bass voice of the narrator. "In the West, there are two kinds of people. Those that are right and those that are wrong." As the men come closer, their silhouettes become distinct, as a landowner with a rifle and his captive. The landowner looks at his pocket watch when a train whistles in the vicinity. "Right on time!" he says and moves behind a rock to cock his gun as the music crescendos and the train whistle intensifies. "Cut!" A nine-year-old girl with a cartoon character on her hat shouts, expressing her frustration and assertively demanding the immediate delivery of her cup of tea. The two actors and the videographer with a smartphone camera in hand look at the young director with disappointment on their faces.

"Think big, start small" is the motto of the Mobile Motion (MoMo) Film Festival held in Zürich annually in early summer. Starting *small* is a pun, setting a parallel between the youth of the director in the video and the use of the easy-access smartphone, which not only makes the camera equipment small, but also cuts filmmaking budgets to an infinitesimal sum compared to ordinary shootings. The Western style of MoMo's 2017 promotional video and the rifle's central role is perhaps not a coincidence, either. "Hold the phone straight!" the director shouts after criticizing her crew, cutting off both the shooting of a movie and the gun, articulating that even a child can do all that better.

Children and adolescents are incredibly responsive to new technology, media, and especially smart devices. Several studies stand in agreement with those by Vittadini, Siibak, Carpentier Reifová, and Bilandzic (2014) and Loertscher et al. (2016), which discuss the importance of age in consumption and media use. Based on a social, cultural, technological, and economic approach, Vittadini et al. (2014) highlight that quotidian commitment to media chiefly stems from the quality and form of social belonging in a mediated or physical sphere. Moreover, they add, exposure is present throughout an expanding timeframe, as ubiquitous media devices are part of the repertoire for various activities from an early age. This largely defines a generation's media identity and the amount and intensity of their media usage. As Vittadini et al. argue, in different societies (delineated by countries, regions, or various social phenomena), media exposure ultimately affects generational differences in terms

of consumption and participation. It can be safely established that an extended learning curve results in increased participation in online mediated environments, and continuous presence yields a more stable social identity as a media consumer—even if motivations, interests, and level of engagement change with age.

In demographic terms, millennials (or Generation Y, the demographic cohort born between the early 1980s and the second half of the 1990s) are perceived as the leading adult generation in adopting and using smart technology. According to a 2018 study conducted by the American Pew Research Center, millennials surpass preceding adult generations in technology adoption, namely, ownership of smart devices, use of social media platforms, and exploitation of the internet (Jiang, 2018). The survey included respondents residing in the United States, but European data[65] shows a similar trend: millennials lead in daily internet usage, mobile and smartphone usage, daily online media consumption, and mobile centricity (the extent to which consumers use their smartphones for various functions instead of other, often analogue devices).

In a study investigating addictive behavior and social stress among smartphone users, van Deursen, Helsper, and Eynon (2015) found that those who exploit the social functions of smartphones (e.g., communication, social media), adopt smartphone-related habits in shorter time and are more likely to develop addictive behavior. The authors use the definition of habit by Oulasvirta, Rattenbury, Ma, and Raita (2012), according to which it is "*an automatic behavior triggered by situational cues*, such as places, people, and preceding actions" (p. 106). These habits result in increased participation in online communication induced by external or internal motivations, for instance, personal urge or the expectation of reinforcement through rewards. Because of the ubiquity of smartphones, dependency on being online more generally affects adolescent and post-adolescent generations (P. Nielsen & Fjuk, 2010). According to van Deursen et al. (2015), the most involved groups are between 15–25 and 26–35 years of age. They are not only faster to adopt the latest technology and practices, but also more focused on pleasurable experiences and immediate rewards for an activity. Moreover, these age cohorts are demonstratively more engaged in communication on online platforms (Howe & Strauss, 2000; van Deursen et al., 2015).

Technology adoption in my reading—in the smartphone spectator's case—does not concern the social or psychological consequences of addiction but rather the factors that establish and advance a habit. Two reflections are worth

[65] Via Google Consumer Barometer (2019); age group of participants born between 1984–1993.

mentioning in connection to van Deursen and colleagues' study, though. First, smartphone spectatorship and the related habits are independent of addiction, since it is the spectatorial activity that defines the modes of usage (e.g., attempting to maintain continuous attention). However, dependency is not completely negligible either: it supports the theoretical framework that explores the mechanisms for obtaining sufficient medium-specific knowledge. Learning to use mobile smart devices for different purposes, from the simplest acts of communication all the way to playing games and watching videos, requires the acquisition of both technical and cultural knowledge. Also, different age groups react to and interact with such gadgets differently, depending upon demographic indicators, such as level of education and country of origin or residence, as well as cultural indicators, such as interests and needs. Still, both the benefits and the dangers of smartphone usage affect millennials and post-millennials to the greatest extent of all the age cohorts using them. Besides being a curious matter, age is one of the indices that make smartphone video consumption a quantifiable indicator when it comes to empirical testing (see Chapter Four).

Applying the physiological and contextual mechanisms of spectatorship to the case of smartphones, I pursue the question of learning and seek an answer to the question of whether spectators can accustom themselves to using a screening platform that promotes audience effect in a loosely synchronized online sphere and copresence in a physical, however unenclosed and unspecified space. I begin from the thesis that viewing activities on smartphones take place in the midst of a temporally and spatially shifted virtual environment and a socially and culturally shifted physical environment. In other words, fellow spectators of online content may watch the same footage at another time at another location, while those who occupy one's space of spectatorship may be attending to other errands. The lack of institutional frame, which otherwise stands in agreement with the time, sphere, and behavioral protocols of spectatorship, ensures the presence of external distraction. The following analysis investigates the narrative role of perceiving visual, sonic, haptic, and other stimuli, with the aim of connecting the discussions of smartphones and the spectator's body and mind. The next section engages in a theoretical discussion of what leads to spectators learning to process a broad range of mediated and unmediated stimuli. The concluding section of this chapter applies this and the general theses of perception and comprehension to viewership on smartphones and discusses the impact of distraction on narrative understanding.

Adaptive Effect

The immersive quality that the latest smartphone models afford approximates that of cinema, television, or video games. By this I mean that smartphones engage users in watching familiar audiovisual content with effortless physical involvement, whether that entails a motionless seated position or active bodily engagement similar to operating a game controller. In any case, an active mental presence dominates physical activities. Regarding effortless and familiar acts of spectatorship, the question arises: what factors in spectatorship are responsible for enabling immediate connection to a narrative and how do viewers and smartphone users acquire sufficient knowledge to achieve such immediacy? In other words, how do users acquire smartphone-specific behavior in order to achieve a high level of spectatorial presence while watching movies or video content?

My inquiry into learning using media platforms combines two aspects, a mechanical one regarding pervasive social trends and customs and a more fluid one that departs from mental development. The common ground between the two lies in the effects of repetition and rewards, which together function as a base for automated consumption patterns and trigger cyclically reoccurring behavioral schemes. These are habits, in other words, automatically (sometimes even uncontrollably) executed behavioral patterns, which enable a medium to become transparent in pervasive and frequent usage (definition based on Oulasvirta et al., 2012).

A number of terms are used to define the processes involved in achieving automated behaviors—especially automatic mental or motor responses to sensory stimulation—with varying applicability to medium-specific learning. Habituation is a potential candidate, being often used to describe similar mechanisms. The term, however, does not denote a de facto learning process that grants users the ability to understand, systematize, and contextualize mediated and unmediated information. The peculiarity of medium-specific learning is that, similar to habituation, after repeated exposure, a stimulus is no longer perceived. But, unlike habituation, medium-specific learning processes are related to social behavior, instead of solely based on neural responses; connecting incoming stimuli to biological reactions which requires conscious training. Technology adoption here is also related to classical conditioning, as it is based on the process of establishing a modal connection between a non-biological, external stimulus (e.g., a ringing bell) to a physical reaction (Pavlov's dog's salivation) (see Rankin et al., 2009). While classical conditioning explains long-term neural changes, habituation touches upon perceptual effects (the way

perception alters as a response to a stimulus), for instance, getting used to a continuous disturbing noise.

As introduced in Chapter One, Michael Grabowski (2015) uses an ecological approach to *neuromediation*, which establishes a connection between the user's body and the sensation of medium transparency. Grabowski's approach implies that, due to neural changes in what he calls "interconnected neural networks," incoming impulses "create new connections between already established patterns" (p. 10). The sensation of immediacy signifies a direct connection between the stimulus and the perceiver's body, and conveniently disregards the screening instrument in both a theoretical and physical sense. This account suggests that the perceptual, cognitive, and behavioral schemes involved in processing incoming information require permanent or at least long-term neural changes. After significant exposure, media users develop skills and acquire the necessary set of context-relevant knowledge to interpret content on a given medium. If, for instance, a user is frequently exposed to narrative films, she will not consciously pay attention to the forms of storytelling. Instead, she will react to events happening in the diegetic space.

This proposition contradicts one of the key points of habituation, namely, that of spontaneous recovery, which is the termination of the habituation effect after stimulation is interrupted. Sufficient skills to contextualize moving images and use a screening and sounding medium (e.g., a smartphone, a video player application, and a pair of earphones) are gained through a learning process, which corresponds to the process of habituation as follows. In a review of the elements of habituation, Rankin et al. (2009) claim that repeated exposure results in a decrease in response frequency or magnitude. Moreover, repetition (even if it involves a series of spontaneous recoveries) and higher stimulus intensity increase the rate of habituation and decrease awareness. Beyond this, there is an observable decrement in response in cases of other stimuli with similar modalities. The latter process distinguishes habituation from other types of sensory adaptation and makes it suitable for development and acquisition studies (e.g., acquiring medium-specific skills). This is due to the fact that habituation research recognizes the organism's responsiveness to other (process-irrelevant) stimuli, such as distraction, that leads to dishabituation. Rankin et al. note that the exact modalities for a stimulus to undergo dishabituation are widely disputed in behavioral research, but the process and the possibility of re-habituation are nevertheless accepted, meaning that irrelevant, distracting stimuli can also induce habituation. In addition, Rankin et al. acknowledge the

existence of permanent, long-term habituation. Yet, calling a medium-specific learning process habituation still falls short in some aspects, as I explain below.

In ecological terms, viewers acquire the behavioral framework for spectatorship in a collective space (i.e., in cinema) by copying fellow viewers and applying institutional practices. As proposed earlier in the discussion of cinema's remediation, this social construction of collective dynamics is well-observable in the case of the smartphone spectator's online and off-line presence—in a virtual sphere and in a physical space. The modes of complying with collective behavioral schemes, although based on automatic processes to some extent, can be explained by the term I call *adaptive effect*. The term was initially introduced to describe biological processes on a cellular level, such as cell transformations catalyzed by medicine or hormones; however, it has been used in biosocial and social behaviors in response to external stimuli (Žuravljev, 2006). I find adaptive effect suitable here as, similar to classical conditioning, it allows for a collective (social) application, unlike the exclusively individual sensory approach of habituation. It, moreover, supports my account of the correlation between the entirety of sensory inputs and evolving neural and physiological responses, as well it can model the evolution of cinematic cognition that favors narrative comprehension instead of medium-specific processing.

Adaptive effect, in practice, leads to decreasing recognition of a medium, stimulus generalization, and finally the adoption of an efficient behavioral framework. In synchrony with decreasing conscious attention to irrelevant stimuli (that is, noise or visual distraction that arrives from sources other than the moving-image content), the process entails a generalization of responses. Accordingly, I presume that spectators perform similar behavior and respond in a corresponding manner whenever encountering movies, irrespective of screening apparatus, if in possession of sufficient knowledge and experience. These behavioral schemes originate from attention, engagement, and emotional involvement, which correlates with the findings of Chapters One and Two. However, two points must be revisited here in order to clarify adoption and learning in relation to the present discussion and describe the skills and cognitive processes involved in spectatorship, especially in smartphone spectatorship. By learning to interpret the distinct alloy of sound and vision presented as a *movie*, spectators acquire the competences that help put together an immediate narrative and navigate an abundance of characters, relations, times, and spaces. Acquiring screen-specific behavior, accordingly, is understood as a practical procedure in which the spectator becomes accustomed to the attributes of a new screen, its resolution, size, location, and the surrounding space.

The development of mobile phones and other communication gadgets followed an ecological path: users have formulated their preferences and manufacturers modified their products accordingly (Ling, 2004). In the process of broadening ranges of functions and improving usability, developments lead to a situation in which millions of people go about their businesses with pocket diaries, pocket messengers, pocket computers, pocket concert halls, and pocket cinemas; although sometimes slightly radical—the first ever portable phone, the first with a built-in camera, or the first with a customizable graphical interface. Still, certain developments seemed unpromising initially: around the time the first generation of Apple iPads appeared on the market, human–computer interaction expert Jakob Nielsen (2010) described them as "wacky." Nielsen criticized touch navigation for the lack of haptic feedback and the fact that the screen (the surface information is projected on) is touched which leaves greasy and dirty spots and easily spoils the visual experience. "Anything you can show and touch can be a [user interface] on this device," Nielsen first announced with mixed feelings concerning the iPad, which applies the same control mechanisms as smartphones. "There are no standards and no expectations" (para. 13). The lack of indicators to guide users to the areas of mechanical control is supplanted by a virtual sphere of interaction. The pseudo-aesthetics of pseudo-objects (including the visual representation of buttons and scrollbars) abandon the user in the midst of an all-clickable maze. The follow-up review (Nielsen, 2011) published a year later acknowledges improvement of the device, which became "decidedly *less wacky*" (para. 1). But with the current abundance of sensor-screened devices at the end of the 2010s, it is the argument about usability and dirty screens which seems wacky.

Although years had passed since the iPhone, the iPod Touch and various other manufacturers' touch-screen devices were released at the time the first-generation iPad appeared on the market, users were unaccustomed to the control metaphors and interface commands, labelled as inconsistent in Nielsen's first review. In retrospect, this opens a crucial inquiry in relation to my earlier analysis, which concerns usability aids that provide hints about the modes of interaction with a new or emerging technological invention, such as the touch-screen smartphone, as well as the way spectators learn to interpret mediated responses, such as visual, sonic, or haptic information. When watching movies or videos on a smartphone, two sets of skills are in play that, conforming to the framework of the adaptive effect, are acquired from previous encounters: one is for interpreting an audiovisual narrative and the other is for controlling and interacting with the device. Both are addressed in more detail below.

A narrative is constructed by combining various sensory stimuli. When it comes to spectatorship, in most cases, this applies to a movie's or video's imagery and sound and additional sensory information perceived simultaneously or recalled from memory. This sensory information is not limited to vision and sound: haptic, olfactory, and other stimuli sensed in real-time or recalled from memories can equally contribute to narrative interpretation, as, among others, Laura Marks (2000, 2002) explains while examining the tactile quality of film. In her previously discussed work, *The Skin of the Film*, Marks (2000) observes how the prevalence of visuality in arts is supplemented by proprioceptive information from non-visual senses such as hearing, touch, kinesthesis, or taste. In Marks's reading, instead of fusing into a conceptual multisensory totality, sensory modalities reinforce the limits of sensing: incoming information synthesizes and evokes memories that activate specific motor and sensory configurations. Thus, different senses contribute to momentary impressions, but also serve as sources of personal and cultural knowledge. The sensory configuration of moving images is closely associated with the authority of sight and sound: the formal idiosyncrasy of the "talkies" disregards other sensory cues, which, however, can alter a spectator's own experience of the narrative.

In my analysis, cognitive access to narratives is considered as a complex system of skills. Visual, auditory, olfactory, haptic, and other sensory information must go through a phase of interpretation before generating neural and motor responses. This is key to discussing smartphone spectatorship, which requires senses other than vision and hearing: touchscreen interaction involves tactile and kinesthetic interaction too. Tactition and kinesthetics serve as two of the mechanisms interactions require (changing the screen position or command the device) and, as argued in Chapter Two, add an extra sensory layer to the spectatorial experience in the form of haptic associations. So, when touching the screen surface and an image of a specific texture, associations are made between what is seen and what is touched.

As it stands in Chapter Two, in the case of smartphones, extra-filmic (non-narrative) sensory information can originate from the mobile device, which can provide visual, sensory, and tactile feedback even when a movie or a video is being played. A wide range of extra-filmic stimuli can also arise from the viewing environment. While the latter can equally apply to any screening media and environments, the former is specific to smartphones or other similar portable smart devices. Unlike multimodal movie screenings, such as

Smell-O-Vision, Odorama, and Aroma-Scope[66] that involve different scents to enhance immersion in the diegetic space, extra-filmic stimuli are deemed unrelated to the primary (movie) stimulus and therefore considered distracting. Distracting information can withdraw attention from the primary stimulus, but can also induce associations which can affect emotional state or even comprehension (see below in this chapter and Chapter Four).

One must note that while senses are embodied and innate, sensing is something cultivated in the process of social development such that it becomes a source of cultural knowledge.[67] Even those senses deemed the most grounded in innate systems are in close connection to contextual meaning creation from incoming stimuli while watching movies. Although movies are created and reproduced using a specific combination of sound and image, other sensory modalities—a familiar smell or texture, for instance—can resonate and thereby become part of the spectator's personal narrative experience or an emotional reaction that can affect immersion. The equation adds up in reverse too: visual input in a movie, such as a closeup of a freshly baked cake, evokes memories relating to taste, scent, and even haptic information about its texture or temperature. A haptic connection to the smartphone screen can further enhance this experience.

Smartphone-Specific Viewer Behavior

Movie spectatorship can be described as a jigsaw puzzle, where the individual puzzle pieces are sensory information originating from external sources, such as the screen, the speakers, the surrounding space, or from internal ones, for

[66] Several attempts were made to include sensory modalities other than vision and sound in cinema screenings and television programs, for instance, scent. Smelling screenings have been around since the sixties and have used scent-emitting cinema seats or scratch and sniff cards (see Gilbert, 2008).

[67] As an example, in discussing smelling as intrinsically primitive, Marks (2000) pronounces a cultural aspect for olfactory functions. She argues against an understanding grounded in Freudian psychoanalysis, claiming that odors are less important to humans than they are to hominoid primates or other animal species. In agreement with developmental neurobiologist, Pamela Hines (1997), Marks argues that, apart from a genetically coded sense for identifying the odors of danger or sexual arousal, humans also learn the meaning of smells via certain contexts: olfactory information reaches the brain through neural pathways that lead to the areas responsible for precognitive or noncognitive sensing (the hypothalamus, responsible for smelling) and for a cognitively wired understanding (amygdala, responsible for memory and emotions). Marks analyses the structure of the olfactory system to demonstrate that sensory perception for all the senses is highly pliable.

instance, the viewer's own memories (see Noë, 2004 and above). Assembling puzzle pieces can signify either data-driven (bottom-up) or expectation- and knowledge-driven (top-down) processing. This returns to an ecological approach to media, cognition, and content consumption according to which neural processes to ecological inputs are connected to construct novel neural pathways.

Lacking preconceived cognitive formulas, an observer perceives objects or organisms through sensory information, such as appearance, size, weight, scent, or the like. Then, one uses this information to hypothesize their function and the potential ways of interacting. According to the constructivist theory applied to spectatorial activity, narrative comprehension functions using the same computational mechanisms (Rescorla, 2017), as an active, goal-oriented act of inference. This suggests a momentary problem-solving task, yet, it can have long-term effects. The adaptive effect (in both the biological and cognitive sense) implies the immediate processing of a sensation and its long-term storage: I argue that a new piece of information is attributed with possible cognitive references (a narrative context, a hypothetical function or meaning), and an inference is made to be revisited in response to similar stimuli for both long and short term. For instance, in the demonstration of the Kuleshov effect, the viewer assigns the sensation of hunger to the shot showing a bowl of soup by creating a logical inference between food and hunger. This association is perpetuated even in the shot of the character's face and results in an interpretation of him appearing to be hungry. On a different scale, the same wiring or re-wiring process can be applied to learning the structural outline of narrative films and understanding a coherent narrative, even if it contains spatial or temporal discontinuities.

As I argued above, information previously gathered via vision, haptic, olfactory, and other senses, such as textures, scents, noises, and even flavors, can provide additional meaning to audiovisual representation by evoking memories of previous encounters. This experience can guide the spectator's expectations and, at least, affect viewing experience. Contextual information in spectatorship supports meaning-making in a way that it generates assumptions and the specific goal of searching for reinforcement for those assumptions. This brings me back to the question about how sensations independent of the movie, such as sonic or visual distractions, or even scents would affect viewing experience and, thus, comprehension.

During spectatorship, several neural paths must be constructed to interpret incoming information from an audiovisual narration, while attention is directed by both externally and internally compelled mechanisms. A suitable method for pursuing this inquiry is to explore the relevant skills for spectatorship and

applying them for smartphones. I therefore outline the cognitive functions in play in identifying task-relevant and task-irrelevant stimuli, while the viewer's primary task is comprehending an audiovisual narrative. The factors most important to this process are the viewer's memories, knowledge, and skills in (a) identifying the function, operability, and role of objects and organisms (characters) in a narrative context, (b) understanding audiovisual storytelling formulas,[68] and (c) employing the operational mechanisms of smartphones. Now, I briefly return to smartphone usability, then apply the theory of adaptive effect to filtering relevant and irrelevant stimuli.

In addition to touch control, the issues circling the first smart devices with touchscreens concerned size: too big to hold in one hand for an extended period of time, but too small to enjoy the high-resolution screen. Also, the lack of tactile feedback, which follows the push of a physical button, makes interaction rather perplexing. Even though the aim of universalizing touchscreens (applying either resistive or capacitive sensory panels)[69] was to create straightforward, easy-access usability, the apparent dilemma between scenic visualization and the functionality of integrated touch sensors was widely debated. Most areas offering haptic interaction are not specified on the hardware, but instead designated by the software. This grants interface designers greater freedom, but the absence of physical constraints, such as designated function keys, raises the question of adoption. Buttons, bars, hatches, or flat surfaces afford specific types of interactions (Gibson, 1977; Norman, 2013; see also Chapter Two), but, while lacking specified touch locations, so do touchscreens: they afford specific gestures, such as one- or two-finger touch or pinch, which are used fairly consistently across applications. A pinch usually initiates the zoom function, and on a wide range of video players, a single touch activates either the pop-up menu panel or pauses the video.

For smartphone use, kinesthetic involvement (holding and moving the device) and haptic input and output are inevitable. Even while watching movies or videos, bodily interaction with the device held in one's hand is fairly common.

[68] My theoretical model assumes a narrative context, that is, content with some sort of a narrative structure. Although I base this outline on narrative films (features or documentaries), I argue that other formats consumed on mobile devices, such as television programs (e.g., news programs, talk shows, game shows) or amateur video content (e.g., interviews, how-to content) have a narrative structure of some sort. Different kinds of content have specific narrative formulas, which must be employed in a similar manner as feature film storytelling.

[69] On a resistive surface, sensitive layers are placed on top of each other and their contact (by physical impact) leads to sensor activation. Capacitive screen panels use the human skin's conductive capacity, and it is the change in the screen's electrostatic field that registers haptic inputs.

Moreover, due to the fact that smartphones are frequently used in unenclosed environments, extra-filmic input, such as noise, smells, visual stimuli, or parallel activities are factors that induce some level of interaction that may affect the viewing experience. Revisiting this statement highlights the need to analyze whether adopting the necessary skills for object recognition, narrative contexts, audiovisual storytelling, and smartphone-specific operation would decrease the likelihood of occasional mind wanderings or attention oscillations from the movie stimuli induced by external sensory elements. This is important for formulating hypotheses to empirically test the effect of adoption on the feeling of presence, emotional engagement, and narrative comprehension in distracting environments. The basis for this speculation is that distraction originates from a clash between diverse behavioral protocols in unenclosed viewing spaces and the hypermedia qualities of smart devices. In unenclosed spaces, the attributes of copresence with others lack the framework of a common spectatorial activity, unlike in a movie theater or a room equipped with a home video system. Conflicting activities (e.g., watching and traveling, watching and dining, etc.) produce myriads of signals that have little relevance to the viewing activity and can cause distraction: for a viewer focusing her attention on watching a movie or video while sitting on a train, noises, visual stimuli, or even the task of traveling from A to B are deemed irrelevant up until the point when her attention moves back to the physical space (for example, when her station is announced). External stimulation also derives from the fact that the movie's sonic and visual stimuli lack the same level of exclusivity as in an enclosed screening space because of the screen's size, sound quality or level, and the involved social and behavioral scenarios.

The other source of distraction is the device itself. Unlike most other screening platforms, smartphones produce visual, sonic, or tactile effects and may visualize information of several running applications or notify users about online activities through pop-up messages on the very same platform where audiovisual contents are played. Moreover, the screen is mobile, which allows for momentary changes in position and distance. All factors considered, the movie presentation is augmented by the following: kinesthetic interaction that define the proportion of the screen and the surrounding space; tactile sensing and the contact with the smartphone; and various external sensory modalities, which originate either from the surrounding space or from the device. The role of adoption stands in filtering irrelevant stimuli.

Torralba, Oliva, Castelhano, and Henderson (2006) developed a computational method for predicting gaze positions, the so-called *contextual*

guidance model. Using this model, the authors aimed to establish a tool for computing the pathways attention and gaze follow. They also sought answers to how the combination of stimulus- (saliency-) driven and context-driven information processing affects attention. Based on their findings and earlier research on stimulus-perception (e.g. Parkhurst & Niebur, 2003), Torralba et al. (2006) assume that the attributes of visual attention are computable from the saliency of visual stimuli, but depends also on semantic and task-related context. Objects (in movies and the real world alike) appear in the configuration of hypothetical scenes, which embed them in a coherent semantic unit and provide cues for comprehending in context. This means that understanding the functions and roles of an object is based upon associations driven by a designated prototypical *context frame* (Bar, 2004). Context frames guide object recognition and feed the observer with expectations, as a result of repeated exposure to the same object and context. Context, Torralba et al. (2006) propose, is accessed through sensory features that attract attention, such as salient or vivid representations of objects, and a task to search for certain predetermined semantically meaningful objects.

In movies, the same connection is observable between visual language (see Cutting, DeLong, & Nothelfer, 2010; Cutting, Brunick, DeLong, Iricinschi, & Candan, 2011; Mital et al., 2011; T. J. Smith & Mital, 2013) and context-specific indicators, such as faces or buildings (Hasson et al., 2004; T. J. Smith & Henderson, 2008; Troscianko et al., 2012). Here, information processed through context dominates low-level features, such as salience. I introduced the theory of conceptual guidance to provide a model paradigm and to extend Lauwereyns's (1998) attention-division control framework mentioned above, which highlights objects' role in drawing attention to themselves. As I argued throughout the preceding sections, the fusion of moving-image stimuli and external sensory data functions as an integral unit of information, the components of which are detached from each other only on a conceptual level. In essence, this means that the observer synthesizes stimuli originating from the movie, the smartphone, and her surroundings into schematic clusters without necessarily separating the context of the film narrative or social situations. Instead, I hold, it is the *intensity*, *neutrality*, and *ecological relevance* of a stimulus that affect attention and possible responses.

Intensity, or stimulus magnitude, is a low-level feature that provides information about the source location of a stimulus; for instance, sound heard from a greater distance is softer than sound in one's immediate vicinity. Intensity is also related to the threshold at which one's sensory system initiates a response (see

Chapter One) and has a significant impact on attention and alertness (Nissen, 1977). This connects it closely to neutrality, which, however, is a contextual property, that is, a high-level feature.

Neutrality concerns the perceiver's interpretation of a stimulus, which is based on skills, knowledge, and experience. Depending upon its neutrality (or, from the opposite perspective, urgency), a stimulus evokes either the urge to respond immediately or no corresponding action. Natural sound or background music and still visual surroundings are less likely to evoke immediate action than, for instance, a ringing telephone (the urge to answer it), an approaching vehicle (the urge to move out of the way), or a pop-up notification on the mobile phone (urge to read an incoming message). In addition to sensory signals, such as sonic, visual, and tactile cues, social cues can also urge the perceiver to take action. Holding the door for someone or giving up one's seat to a fellow passenger are social formulas that are triggered by an impulse for immediate response.

Ecological relevance is also based on context; it appertains to the embeddedness of a stimulus in a specific narrative or series of actions. Bluntly put, ecological relevance signifies whether or not a stimulus makes sense in the context of the viewer's perception of a movie or observed events. For example, even if an external sound's intensity and apparent urgency should draw the viewer's attention from the movie, it may be overlooked if it is understood as cohering to the watched sequence: a honking car in the physical world might be ignored or perceived as part of the film when watching an intense car-chasing scene in *Drive* (Winding Refn, 2011) featuring a Hollywood stuntman, or *Duel* (Spielberg, 1971), where the ominously approaching tanker truck chases the main character in his Plymouth along the canyon roads of California. Similarly, seemingly incoherent narrative information from a movie can induce uncanny feelings, such as the nondiegetic visual representation of diegetic information.[70] Examples of this include text messages legible to the audience next to the character' face in ABC television series, *Nashville* (2012–2018) or Sherlock Holmes's theories and related lexicon or dictionary entries appearing on the screen as a manifestation of his thoughts in BBC's *Sherlock* (2010–2017).

This classification of stimuli, which I briefly introduced in Chapter One, serves the purpose of exploring the particularities of attention in the case of

[70] Incoherent sensory information, for instance, foreign-language speech or a seemingly inappropriate voice for a character (e.g., different age, gender, etc.); or laughter, a cheerful musical piece, or the scent of flowers in a crime drama would be highly dissonant. The spectator would be less likely to perceive such stimuli while immersed in watching the film and, if perceived, would impair engagement.

smartphone spectatorship in unenclosed spaces. It corresponds to the contextual guidance model proposed by Torralba et al. (2006) and their computer algorithm for predicting gaze behavior. Their theory, that sensory input is processed through contextual understanding, concurs with Levin and Simons's (Levin & Simons, 1997; 2000; Simons & Levin, 1998) argument that continuity errors in moving images and inconsistencies in the physical world are equally likely to be masked by narrative contexts. Observing people's reactions to inconsistencies, such as missing objects or inconsistent costume color, Levin and Simons found that responsiveness follows similar patterns when a scene is presented on film as when it is observed in real life.[71] This is attributed to the fact that observers fill in "perceptual blanks" with their own knowledge and experiences, which may be explained by their level of engagement with the goals of a task. Regarding film editing, T. J. Smith (2012), Shimamura, Cohn-Sheehy, Pogue, and Shimamura (2015), among others, claim that cuts are also inferior to narrative interpretation and therefore are rarely perceived consciously.

I propose that attention transitions between a movie or video and physical-world stimuli are similar to cuts in a way that they are largely not perceived (at least consciously), but that they connect diegetic and physical events through context-specific cues. Based on the intensity, neutrality, and the ecological relevance of the two sets of stimuli, context-irrelevant elements may be masked by the primary activity, that is, movie watching. Moreover, as my theory of the adaptive effect suggests, this process is the result of acquiring contextual and reaction schemes related to one's physical space, the movie, and the smartphone's hardware and software (or application) design. Thus, when immersed in watching a scene from *Drive*, for instance, on the road, in transit, or in a busy urban environment, it is likely that the viewer merges the motoric and urban noises of the movie with corresponding sounds from her physical surroundings: she may register both sets of sounds, but will assign them all to the context with which she is more engaged. Paying attention to real-world information—for example, sound, and also the sight or the smell of traffic—becomes secondary to the task of comprehending the main protagonist, the driver's stunt and criminal actions. So even if her attention oscillates between stimuli, the

[71] Levin and Simons (Levin & Simons, 1997; 2000; Simons & Levin, 1998) demonstrated that people fail to detect inconsistencies even when observing scenes in real life. In one experiment (Simons & Levin, 1998), a person approached a pedestrian and initiated a conversation. In the middle of the conversation, another person arrived carrying a door and while the pedestrian's view was blocked by the door, the first person walked away and another one stayed and continued talking. Even though the two people wore different colored clothing (and looked different), only half of the pedestrians noticed the change.

viewer might not notice the exact moments when her attention shifted (cut) from the movie to the physical world. Consequently, attention oscillation between movie and irrelevant distracting stimuli would be infinitesimal in the case of high information load that emanates from spectatorship (see Cutting et al., 2010).

The viewer's level of (emotional) engagement, associated with a feeling of presence in a story and an empathic connection to its characters, is highly dependent upon the type of stimulation. As Bordwell (1986) explains in *Classical Hollywood Cinema*, cinematic techniques direct viewers' attention and movies that most actively guide (manipulate) attention are those that distort reality the most. Empirical studies reported by Hasson et al. (2004), Hasson, Furman, et al. (2008), and A.-L. Cohen et al. (2015) confirm that the more a movie is based on a carefully controlled juxtaposition of images and sound and the less faithful it is to reality (for instance, lacking long takes or deep focus), the more manipulative it becomes. Hasson, Landesman, et al. (2008) refer to filmmaking schools, such as the Soviet montage and Alfred Hitchcock's legendary capability to manipulate viewers, in contrast to Italian neorealism or *La Nouvelle Vague*. Their approach to quantifying the extent of manipulation is based upon inter-subject correlation (introduced by Hasson et al., 2004), the index of how much synchronicity spectators' individual physiological responses show, where a higher correlation typifies a higher level of engagement.[72]

Temporary engagement with a narrative is a dynamic process in which the observer or media consumer incorporates narrative information into prior knowledge to create a personal story. This is closely intertwined with identifying with a protagonist's goals and means of achieving them (A.-L. Cohen et al., 2015). Cohen et al. conducted a study in which they showed an episode of *Alfred Hitchcock Presents* (1955–1962), *Bang! You're Dead*, to participants either as is or in a noncontiguous form with the scenes out of sequence. Hitchcock's crime drama is about a five-year-old boy who plays with his uncle's loaded revolver thinking it is a toy gun until his family sets off in a frenzied search to stop him. The participants' task was to lift a hand every time they heard the word "gun." Cohen et al. intended to measure engagement, following a hypothesis that with greater engagement (in the contiguous form), participants would more likely neglect the task. Two mechanisms operated in these trials. First, the task of identifying the moments when the word "gun" is vocalized induces task-driven processing. Second, Hitchcock's storytelling and narrative suspense cater to

[72] Inter-subject correlation has been measured from brain activity and eye movements. See also Mital et al. (2011) and T. J. Smith and Mital (2013).

both stimulus-driven information processing involving momentary problem-solving and knowledge-driven processing that yields assumptions about outcomes and engages viewers in the protagonists' actions and thus the narrative. (Note the famous Hitchcockian suspense!)

The outcomes of the study reveal that the frequency of correct responses to tasks decreases with time when greater engagement is achieved. These results suggest that engagement makes spectators less responsive to specific tasks not directly related to the "task" of narrative comprehension. As declining performance was only significant in the contiguous version and minimal in the noncontiguous version, Cohen et al. dismissed the possibility that cognitive load (the difficulty of the task) affected performance. Their conclusions are explained by the fact that participants forgot about the task as they became more engaged in the contiguous version of the sequence. This implies that as suspense increases,[73] so does engagement, which results in greater attention to the protagonists' actions and less of a likelihood of remembering the assigned task (Kuhl, 1984).

In relation of the above discussion, I suggest that spectators carry out the act of watching a movie in relation to the effects of short-term adoption (adopting the rules and logic of the diegetic space) and long-term adoption (adopting a medium-specific spectatorial framework). Long-term adoption involves three factors. The first is knowledge of storytelling and agency, that characters function according to specific intentions, as well as the rules of physics and social dynamics—as demonstrated in Heider and Simmel's (1944) experiments. Heider and Simmel found that even moving geometrical figures on a stop-motion clip can be interpreted as actors performing a series of actions, from the pace and direction of their movement.[74] The second factor relates to the filmic medium and the automatic processing of a narrative without attention to the modalities of moving images, such as discontinuities (see Schwan & Ildirar, 2010).

The third factor, acquiring the relevant knowledge of the screening medium, results in a focus on content without conscious consideration of the container medium. Chapter One proposes that, besides social dynamics, the most fundamental differences between cinema and smartphones lie in the magnitude of

[73] A.-L. Cohen et al. (2015) rely on a study by Bezdek et al. (2015) in stating that the evidence for increasing suspense is the manifestation of an apparent threat to the characters, which is a result of fearing negative and hoping for positive outcomes.

[74] Their conclusions have been the basis of various other experiments within the domain of social perception and agency that define the causalities of inter-personal relationships with fictional characters (see, for example, the entries in Rutherford & Kuhlmeier, 2013).

stimulation, which is the result of the smartphone screen's size, the earphones' sound quality, and the sensory presence of the surrounding online and off-line space. I also argue, that, despite these apparent differences, a spectator's attempts to engage with a narrative can still effectuate the sensation of non-mediation, meaning that the watched sequence is perceived as a sensory experience of real-life scenes. Adaptive effect, in terms of smartphone spectatorship, signifies a cognitively economic execution of movie or video watching with an intrinsic focus on narrative actions, which becomes potential source of immersion and emotional engagement. In other words, instead of continuous, mindful attention to the screening device, spectators engage with the narration up until the point when incongruous sensory information distracts them. Being unexpected, a distracting stimulus draws attention to itself when its ecological relevance to the diegetic space is insignificant, its intensity is substantial, and it requires an immediate reaction. This suggests that distractions induce problem-solving and decision-making independent from the primary (movie) stimulus.

It has been long known that the more intense a stimulus is, the higher the level of engagement it induces. As first stated in the early twentieth century, in Münsterberg's *Photoplay* (1916/2014), the desired intensity of visual stimuli is dependent on "the right" perspective of the screen, the mirror position of the camera. Dissecting all the possible aspects of the photoplay and its audience, Münsterberg assigned the prominence of moving images to the fact that they arrange sensory information and preconceived knowledge into contextual meaning and life-like representation. Cross-modal links between different sensory stimuli support this contextualization by incorporating the relevant clusters of schemata and activating neural gestalts. Thus, as a result of adoption, schema configurations construct a narrative out of sensory elements and relevant knowledge. This can mask the processing of context-irrelevant (but sufficiently intense) distracting stimuli, which can be explained by the framework used for testing inattentional blindness and event-awareness studies (see, for instance, the invisible gorilla illusion). Moving-image narration signifies the careful organization of shots that inform the viewer about appearances and relationships. Although transitions (cuts) differ from natural perception, viewers adopt modal links to interpret discontinuities. This makes movie sequences condensed instances of actors, objects, spaces, and times.

Chapter IV.

The Physiological Effects of Smartphone Spectatorship

Screen Size, Physical Connection, and Unenclosed Viewing Environments

The peculiarity of smartphone spectatorship lies in its pervasiveness, the fact that it has little in the way of cultural, behavioral, temporal, or spatial constraints. The following experiment[75] aims to complement the preceding theoretical analysis and take the first steps toward assembling a broader body of experimental validation of mobile media consumption. Its long-term goals lie in approximating a thorough understanding of the ways in which smartphones and other portable personal devices require new viewing strategies by users and possibly new strategies from content producers.

As I proposed throughout my theoretical analysis, viewing experiences on smartphones differ from cinematic and home video experiences in a number of ways. These differences can be classified as to three main factors that provide the basis for this experiment. First, the viewing space is predominantly an unenclosed space where various activities and social norms exist parallel to spectatorship. Second, the screen is smaller, which is assumed to alter gaze behavior and impact attention. Lastly, the viewer has a bodily connection to the device—both its hardware and software, which is frequently adjusted through kinesthetic and haptic interaction. Interaction with the screen pertains to adjustments performed by moving the screen to change the visual angle and by executing manipulations on the graphical user interface through touchscreen interaction. To further inspect these claims, I designed this experiment to

[75] Measurements were carried out at Aalto Behavioral Laboratory, Aalto University, Finland in 2017.

measure the effects of these factors on attention, engagement, feeling of presence, and narrative comprehension.

An unenclosed, actively engaging environment functions according to specific material and social properties, which often fail to cohere with viewing activities. While watching movies or videos in such spaces, an abundance of sonic and visual information and behavioral frameworks divides the spectator's attention among the different sets of sensory stimuli and of social and bodily enactment, as well as multiple independent narratives. The fictional narrative is augmented by the viewer's physical presence. Accordingly, stimuli and behavioral plots originating from the screen and the surrounding space, as well as from the act of viewing and parallel activities, are fused together, forming mental schema clusters. This implies that visual, auditory, tactile, and other types of sensory information acquired from one's surroundings can appear as auxiliary information to the primarily stimulus, the movie. For instance, if a sonic referent is perceived as relevant, even if it is not a composite part of the audiovisual content, the spectator may interweave the sound with the movie narration. Thus, the stimulus becomes an acoustic cue. If, however, the stimulus is unsuited to the movie narration—for example, when it is foreign or extraneous to the diegetic space, the temporal frame, or the on-screen actions—it can disconnect the spectator from her engagement with the narration. This leads to two methodological directions, which are addressed one after the other: first, classifying sensory stimuli as secondary information sources (see also Chapter Three), and second, analyzing them through the lens of immersion and engagement to assess their potential as sources of distraction. These points are important for modeling unenclosed spaces in a quantifiable manner and measuring the effects of distracting stimuli. Although distracting stimuli in everyday environments can be of any sensory modality, they primarily belong to the visual scope and soundscape, which leads this study to concentrate on visual and auditory distractions.

Visual and sonic referents are intertwined in spectatorship: visual information provides a framework for sounds and sonic cues, such as words, voice changes, music, or sound effect onsets, guide visual attention (see T. J. Smith & Henderson, 2008; T. J. Smith, 2014). Also, according to Annabel Cohen's (2009) conclusions drawn from viewers' subjective evaluation of sound and emotional reactions, attention and emotions are driven by the integration of sonic and visual information. An extension of Cohen's thesis exists in eye-tracking research conducted by Auer et al. (2012), Coutrot, Guyader, Ionescu, and Caplier (2012), and Wallengren and Strukelj (2015): adding sonic stimuli to

images influences physiological reactions, which consequently affects the overall information intake and viewing experience. The combination of sound and image directs attention, affects emotional responses, generates anticipation, and influences narrative meaning. In another behavioral study, Escera, Alho, Winkler, and Näätänen (1998) measured involuntary attention shifts resulting from changes in the acoustic space. Using different sound effects with varying relevance to a task and varying familiarity to participants, Escera et al. concluded that novel, unexpected auditory effects, as well as smaller changes (i.e., the introduction of a "deviant" tone), capture attention involuntarily, which is attributed to biological mechanisms for detecting changes in the sonic environment. This leads to the assumption that context-irrelevant stimuli can divert observers' attention from a movie.

As I established in the previous chapter, visual and sonic stimuli in the surrounding space can be classified in relation to movie sound and visual language as something either bearing or omitting *ecological relevance* to them. Ecological relevance is a context-dependent indicator pertinent to a stimulus' embeddedness in a specific narrative framework. In other words, ecological relevance marks the extent to which the stimulus in question coheres with the narrative assembled from the visual and sonic cues a movie provides.[76] The two other principles for classifying extra-filmic stimuli are *magnitude* (intensity) and *neutrality* (urgency or the lack of need for a conscious response). Stimulus magnitude provides information about the source location of a stimulus (distance and direction) and, to some extent, its neutrality as well. This means that a stimulus of large magnitude (e.g., a loud noise or blinding light) is likely to capture attention, not only on account of it possibly being unpleasant, but also because it implies a need for intervention. Neutrality, however, is independent of magnitude; it indicates the need to intervene on the basis of conceptual interpretation as well. For instance, to recall the discussion in Chapter Three, a still visual field, an abstract image, or background music will likely pass unnoticed and without corresponding action, whereas a ringing telephone, a fire alarm, the sight of a written text, or an approaching vehicle induce immediate reactions depending on social and cultural cues. Based on ecological relevance, magnitude, and neutrality, stimuli originating from the surrounding space have varying effects on viewers' attention (see Table 1 and Chapter Three), which can be predicted to affect immersion and even comprehension.

[76] See also E. Colin Cherry's (1953) and Anne M. Treisman's (1964) corresponding research on text-relevant and text-irrelevant stimuli.

Table 1
Classification of Stimuli

Classification	Stimulus quality	Positive outcome (attention stays on the movie content)	Negative outcome (attention is shifted to the source of distraction)
Ecological relevance	Nature and meaning of stimulus	Related to the movie narration	Unrelated or irrelevant to the movie narration
Magnitude	Intensity (loudness, saliency, etc.)	Low intensity	High intensity
Neutrality	Response urgency	No immediate response required	Immediate response required

As a personal device with significant physical proximity, a smartphone has an additional quality that can cause distractions. The fact that viewers have control of screenings and screen content in general can impair continuous focus on content, which can also affect engagement and narrative comprehension. This stems from the device's hypermedia qualities; pop-up windows, message notifications, and opportunities to switch between applications and screen configurations may interfere with movie or video viewing. Moreover, personal preferences and media consumption habits may cause detachment from the audiovisual narration: adjusting screen position, playback speed, luminance, and volume, or in other ways changing the balance of filmic and environmental stimuli can increase the risk of oscillating attention.

In Chapter Two, interaction was attributed to two factors, namely, externally controlled distractions and internally controlled involvement. In the first case, interaction (in other words, disruption of attention) stems from external motivators; external stimuli, such as incoming messages or distracting elements in the surrounding space propel interventions. In the second case, these are internal motivations, curiosity or interest in engaging with the narrative more intensely than previously. For instance, moving a screen closer to one's eyes expands coverage of the viewer's visual field; increasing the volume increases one's presence in the diegetic auditory field; and repeating sequences, freezing images, or decreasing playback speed can provide additional details about certain segments of a movie's or video's content. Interaction is presumed to be in a tight interplay with internal motivations and external distractions to adjust viewing conditions to accommodate one's surroundings.

Navigating the smartphone's interface through the touchscreen induces a higher level of involvement and stronger emotional reactions than non-haptic screens do (see IJsselsteijn et al., 2000; Oakley, Brewster, & Gray, 2001). Moreover, frequently used objects and interfaces, such as the control panel of the various video player applications or the smartphone's interface itself, become transparent (Suchman, 1987; Gauvain, 2001): adopting the related control mechanisms allows for automatic control (see Hirose & Nishio, 2001). The possibility of adjusting sound and image presentation enables a personalized viewing that is responsive to both the environment and the spectator's needs. For this reason, one can predict that touchscreen interactions either lead to a high level of immersion catalyzed by an "ideal" presentation of narrative information (van Laer et al., 2014) or to increased media awareness due to oscillating attention between the narrative, smartphone, and surroundings. Visch et al. (2010) claim that highly immersive conditions (in their case, the virtual reality CAVE) and a screen close to the viewer intensifies both the appreciation of the movie (artifact emotions; E. S. Tan, 1996) and emotional engagement to its characters (fictional world emotions). This is due to the VR stimulus being perceived as another reality. In addition, emotional responses correlate with the level of immersion (Gross & Levenson, 1995), and emotional engagement is the result of empathy toward characters' personalities and acts (M. D. Slater & Rouner, 2002).

Options for interaction lie in close connection to social factors, too: the spectator's presence in an unenclosed space amid manifestations of social and physical references can blur the clear boundary between the diegetic and physical space. Mental connections that support comprehension and that are generated between the sphere of narrative actions and the spectator's bodily and social presence are affected by social domains. These social domains can be based, for instance, on privacy and embarrassment over private actions (i.e., screening on a private screen) being exposed to a larger crowd. The spectator's presence in a socially, culturally, and physically unenclosed space may establish associations between physical entities and fictional characters or objects. Getting frightened by a sudden external stimulus (for example, a sound effect, a moving vehicle, or a passerby), while watching an uncanny movie scene would be a possible scenario and suitable example here.

Media properties can affect the spectator's feeling of presence depending on the intensity of visual stimuli and the screen's size. This evidently raises the question of engagement and distraction in that images (through, for instance, close-ups or wide-angle shots) change from being enlarged from its size in real

life to being compressed, which may alter a movie's affective qualities. Even though the physical parameters of smartphones allow for plasticity in the visual scope, the small coverage of the visual field can lessen immersion in the movie or video due to the limits of screen size and consequent substantial access to the surrounding physical space. Screen size and position foreshadow impact on the following factors: first, emotional and contextual involvement in the narrative in case of environmental distraction; second, the vividness of the image and its conformity to reality, which depends on the distance between the screen and the spectator's eyes and the size of the film image on the screen (based on software and application layout); third, the proximity, number, and intensity of the involved sensory modalities; and fourth, immersion, that is, the sensation of being in a direct sensory connection with the diegetic space without being aware of the screening medium (M. Slater & Wilbur, 1997; Wirth et al., 2007).

In an experiment comparing viewers' reactions to footage watched on different screens, Troscianko et al. (2012) pursued answers to whether a larger screen would result in better experiences. They used two screens of different sizes ("big" screen, 1.5 meters x 1.125 meters, with a 2.61 meters distance vs. "small" screen, 0.89 m x 0.668 m, with a 1.55 m distance) and compared the level of immersion in certain scenes of the movie, *The Good, the Bad and the Ugly* by measuring reaction times and pupil dilation, and obtaining real-time subjective presence ratings.[77] Their results revealed that large screens produce higher subjective presence scores, especially during scenes depicting faces.[78]

Whereas Troscianko et al. (2012) used a constant visual angle (when watching the smaller screen, subjects were also seated closer to the screen), in Tim Smith's (2014) demonstration mentioned in Chapter One, the corresponding usual viewing angles were used for eye movement-comparisons of both BBC reporters tested on a laptop screen and in an IMAX theater. Similar to Smith's experiment, investigating smartphone spectator's behavior and physiological reactions can also benefit from viewing angles that are typically used in these settings. Therefore, participants in the present study were given the liberty to adjust the screen to their personal preference, visual abilities, and the environment.

[77] Troscianko et al. (2012) used a line bisection task to obtain real-time ratings at certain points in the movie: when an indicator light was turned on, subjects marked a line with the appropriate level of immersion.

[78] In others' research, "better" performance was indicated by improvement in completing a visual task (as in D. S. Tan, 2004) or a higher level of self-reported presence (as in Lombard & Ditton, 1997, for instance).

The Present Study

This experiment aims to distinguish the impact of smartphone spectatorship on attention, immersion, and overall narrative understanding, compared to other, widely used screens in commercial consumption, for example, home video or cinema. Cinema, television, and even computer screens are not only larger, but are also most commonly placed in a fixed position in spaces which are set to enhance viewing experience. Contrarily, smartphone users often consume moving-image content on their handheld devices as a secondary activity and in spaces not necessarily designed for movie watching. An essential part of smartphone spectatorship is the active presence of the surrounding physical space, the proportion and sensory presence of which are largely based on the viewer's positioning of the screen and adjusting of the auditory channels. Instead of being an idle and permanent screen-centered space, it is filled with various independent actors, actions, and cultural networks, inducing constant attention oscillation between physical stimuli and the diegetic space.

To determine the impact of these, volunteers were recruited for a laboratory experiment, during which they watched clips from a feature film. While watching, their physiological reactions, eye movements and skin responses, and narrative comprehension were measured and their subjective ratings obtained.

Eye movements provide valuable information about viewers' attention (Duchowski, 2007; Mital et al., 2011). As feature films exogenously control attention to a significant extent, synchrony across viewers is generally expected to be high, which allows for testing the major effects of smartphone spectatorship on gaze behavior (T. J. Smith, 2006; Mital et al., 2011; T. J. Smith et al., 2012). Eye tracking records ocular movement in real time by monitoring and calculating gaze locations based on the positions of infrared light beams bouncing off from the surface of the eye. These measurements, however, do not provide sufficient information about information-seeking strategies, or about the role of physical engagement with a stimulus. Therefore, in this experiment, they were complemented by other methods. The most fundamental reason is that, although gaze positions and attention often coincide, gaze data cannot ensure that attention is paid to the object or the area upon which the gaze is concentrated. Another reason, more specific to screens and moving images, is the center bias: gaze tends to shift to the center of the screen after cuts and when there is nothing attractive elsewhere. Spectators are also apt to fixate on the center of the screen when watching content with rapid changes. This can affect the correlation of gaze positions and attention. Focus on the center also implies that the gaze rarely shifts to the edges of the screen (Tseng, Carmi, Cameron,

Munoz, & Itti, 2009; Mital et al., 2011), which is highly significant when measuring the effects of screen size on visual exploration. In addition, smaller screens have a tendency to effect a decrease in the amplitude of eye movements; spectators are likely to move their eyes within the center of a small screen, since the sharp area of vision (the central two degrees in the visual field) covers a larger proportion of the screen (see T. J. Smith, 2014).

Besides eye tracking, complimentary biometric measurements, self-reporting, and performance-based tests were used. The combination of these help determining not only *where* spectators look, but also *how* they acquire the necessary narrative information as well as what elements of the narrative support understanding and induce emotional reactions. In other words, these measurements provide information about differences in attention, emotional engagement, and the feeling of presence. Electrodermal activity (EDA) is a sensitive marker for emotional engagement. Confrontation with emotionally loaded stimuli induces changes in, among other things, pulse and thermoregulation, and activates sweat glands. Electrodermal activity measures monitor these autonomic, unconscious, changes in skin conductance, providing information on emotional arousal, reactiveness, attention, and immersion (E. S. Tan, 1996; Boucsein, 2012).

The self-reporting questionnaire and narrative comprehension test used for this experiment were designed to complement the biometric measurements (see M. Slater & Wilbur, 1997) and to gain information about subjective impressions of the overall viewing and narrative comprehension. These provide feedback regarding the average value of engagement for each trial (subjective ratings of the overall experience) and attention patterns for specific sections of the trial (comprehension of details presented once at given time segments).

Taking screen size and environmental distraction into account, the common specificities of smartphone spectatorship were recreated and compared to the circumstances that larger, stationary screens offer. Such a comparison was expected to reveal whether screen type (small smartphone screen or fixed large screen) and the presence or absence of additional sonic and visual distractions (interrupted and uninterrupted viewing) would skew responses in favor of uninterrupted, large-screen viewing. These effects were hypothesized as follows:

Hypothesis 1: There is a significant interaction between the effects of screen type and distraction on participants' physiological and subjective responses.

Hypothesis 1a: There is a significant main effect of screen type on physiological and subjective responses, where the large screen produces higher indices of engagement relative to the small screen.

Hypothesis 1b: There is a significant main effect of distraction on physiological and subjective responses, where interrupted watching reduces indices of engagement relative to uninterrupted viewing.

Hypothesis 2: Screen size affects gaze dispersion in a way that fixation points are more centralized in the case of smartphone viewing relative to large-screen viewing.

Hypothesis 3: In interrupted viewing conditions (interrupted smartphone and large-screen conditions), there is a significant main effect of the different forms of distraction on participants' physiological responses.

Hypothesis 3a: There is a significant main effect of a distractor's type on physiological responses.

Hypothesis 3b: There is a significant main effect of a distractor's neutrality on physiological responses.

Hypothesis 3c: There is a significant main effect of a distractor's ecological relevance on physiological responses.

Method

Experiment Design

In order to recreate the distracting nature of unenclosed environments and isolate the effects of both screen size and external stimulation, the experiment followed a two-by-two incomplete mixed factorial design. In this design, screen type (mobile screen and stationary projector screen) and the presence or absence of external stimuli ("distractors") were used for comparing the values produced by physiological responses, subjective ratings, and comprehension test. Distractors were used for establishing a near-natural viewing environment and recreating the occasional interfering stimuli that unenclosed spaces entail, while maintaining control over the stimuli. This design delivered four conditions,

mobile condition with and without distractors and projector condition with and without distractors. To these conditions, two separate, but corresponding, movie clips were assigned. Each participant was tested individually and was appointed to two of the four conditions and watched both clips, one on each type of screen, with and without distractions (interrupted mobile and uninterrupted projector *or* uninterrupted mobile and interrupted projector). The order of the clips was randomized but counterbalanced to minimize sequential effect,[79] while producing an equal number of trials for each combination.

The large-screen setting, the projector condition, was designed to provide baseline data for viewer behavior. In this setting, a stationary screen was used with a fixed viewing distance. The counterpoint measured smartphone spectatorship: the mobile condition recreated a typical smartphone viewing setup. In the interrupted conditions, additional light audio and visual effects were played at determined points in time in correlation with the sequence. The time marks for distractors were assigned to specific narrative elements with meaningful or high emotional content and were the same for each and every participant, independent of their subsequent intervention (e.g., pausing the footage) on the smartphone.

The specific distraction effects were chosen to model any unenclosed viewing space and, although they went off unannounced, created no more physical discomfort to participants than any stimuli in any natural environment. Semantically independent distractions are perceived in a bottom-up manner, meaning that, by their surprise effect, they draw attention quickly by means of low-level qualities, irrespective of semantic meaning. However, stimuli with ecological relevance to the diegetic space are considered less likely to induce notable reactions than unrelated stimuli. It is assumed that conceptual processing of external stimuli depends upon their connection to the primary information source such that stimuli unrelated to the on-screen content more likely induce physiological reactions, for instance, shifts in gaze in an attempt to find the source of the stimulus, than distractions with reference to the movie's diegetic space. Moreover, a complex and urgent stimulus is considered more likely to be perceived than a simple, neutral one.

In order to verify these propositions, three sonic and two visual distractors were chosen as secondary stimuli, with varying ecological relevance, source locations, durations, complexities, and neutralities. The specific distractors included a city sound with traffic noise (played at 7:15–7:29), a ringing tele-

[79] On reactions being biased due to knowledge of a sequence, see Noton and Stark (1971a, 1971b).

phone (7:41–7:52), a written literary text (8:32–9:00), and natural sound accompanied by an animated two-dimensional rectangle (9:01–9:10). Distractor type (i.e., distractors 1–4, see Table 2) was a repeated measures variable, which was analyzed for the subset of conditions containing distractions.[80]

Table 2
Types of Distractors Used in the Interrupted Conditions

Distractor		Stimulus	Ecological relevance	Neutrality	Source
1	Sound	Traffic noise	Diegetic space	Neutral	Diagonal speaker (back)
2	Sound	Ringing phone	Physical space	Urgent	Speaker (front)
3	Image	Written literary text	Neither	Urgent	External screen (front)
4	Sound	Chirping birds	Neither	Neutral	Speaker (front)
	Image	Animated rectangle	Neither	Urgent	External screen (front)

The first sound effect (traffic noise) was chosen to create an acoustic experience with ecological relevance to the diegetic space, given that the movie clips were set in an urban environment on a weekday morning. This made the sound effect function as an external diegetic feature. The second sonic distractor, the ringing telephone, functioned as a non-diegetic feature. The ringing sound had ecological relevance to the physical space where subjects either used a smartphone (mobile conditions)[81] or were aware of the presence of one in the room (projector conditions). The first visual distractor was a Shakespeare quote in English, which gradually appeared line by line with black font in front of a white background on a separate screen. Given its urgency (for reading) and the fact that it had no connection to the movie clip, this distractor was set with the expectation of attracting attention and causing participants to oscillate their focus between the primary screen and the external screen. The last distractor featured a sound sequence of birds chirping and, simultaneously, a dark green rectangle shifting from left to right against a white background on the external

[80] Before the final experiment, a low-sample pilot study was conducted to monitor the validity of the methodology and to select the distractors. During this study, participants had to choose the stimuli they perceived while watching the movie clip from a list of possible sound and visual effects.

[81] After the mobile viewing, one participant reported that she felt the urge to respond to the "call" when she heard the sound, even though there was no indication of it on the screen on which she was watching the movie clip.

screen. The sound evoked no need to take immediate action and the visual effect required only a short period of uninterrupted attention. Neither sound nor image had ecological connections with the physical or the diegetic space. Moreover, the dynamics of the figure were also independent of the movie sequence.

Participants

Thirty-eight volunteers, aged 24–37 (M = 28.6, SD = 3.52), were recruited for the experiment. Although this demographic group is far from representative of all smartphone users, it is archetypal of the demographic cohort of the millennials. Demographic cohorts differ both in terms of smartphone usage and mobile video consumption. Millennials belong to the adult generation that most widely exploits the functions and options that smart devices offer: studies show that millennials are leading in technology adoption, mobile media and social media usage, and mobile centricity.[82] As I argued in Chapter Three, the indices that describe millennials' smartphone usage qualify this age cohort as highly significant in characterizing the process of rapidly adopting smart media players at a post-adolescent age; familiarizing themselves with the widest range of functions, such as movie and video consumption.

All the participants were required to have normal or corrected visual and hearing abilities and to possess sufficient skills in English in order to fully comprehend both the instructions and movie sequences, as subtitles or synchronized voice translation could have affected their responses. After being recruited, each volunteer gained access to a short online survey that recorded demographic data, user habits, and experience with smart devices and mobile video player applications. Those who reported a lack of experience with such devices (i.e., no access to or less than two months of experience using smartphones; no consumption of audiovisual content on any portable smart devices) or other biasing factors were not considered for the experiment. Eligible volunteers were, then, assigned a time slot for measurement and were randomly assigned to a group with predetermined stimuli and viewing conditions. The anonymity of all participants was controlled according to relevant legal regulations; personal data was handled pseudonymized.

[82] For demographic data on European and United States populations, see van Deursen et al. (2015), Jiang (2018), and Chapter Three.

Material

For the experiment, a contemporary Hollywood-style feature film,[83] *The Walk*, (Zemeckis, 2015) was used. *The Walk* is based on the true story of Philippe Petite, a French artist, who in 1974 performed a tightrope-walking act, completing several crossings illegally between the tops of World Trade Center's towers. The movie opens with Petite's early years as a street performer, then depicts the planning of the performance between the Twin Towers, and climaxes with the act itself referred to as "the coup." The movie uses narrative techniques that hold a significant amount of exogenous (externally driven) attention control. Unlike natural scenes, action-filled feature films from the recent past tend to guide viewers' attention by juxtaposing a carefully selected set of semantically meaningful elements, light, vivid colors, and other sonic and visual effects which induce eminently congruous bodily and mental responses across audience members (see Chapter Three and Hasson et al., 2004; Hasson, Landesman, et al., 2008; T. J. Smith, 2013; T. J. Smith & Mital, 2013).

The storytelling style of classical and postclassical Hollywood films serves as a suitable starting point for investigating viewer behavior and narrative information acquisition on smartphones. In addition to this, representing the most significant factors for exogenous control, the following criteria were taken in account when choosing the movie and selecting the relevant clips for the experiment. Besides the movie's relative obscurity, yet up-to-date visual style (it needed to be recent and/or set in our present time or a relatively near past), another requirement was that it features details that maintain and control attention in an analogous way for all viewers. These details include short, fast-paced shots, semantically meaningful elements, such as facial expressions, landmarks, animate and moved objects, and congruent cultural references that induce identical and synchronous reactions (see Hasson et al., 2004; Itti, 2005; Carmi & Itti, 2006; T. J. Smith & Henderson, 2008; Cutting et al., 2011; Mital et al., 2011; Zacks & Magliano, 2011).

Compared to stationary screens at home and in designated public screening spaces, viewers typically spend less time watching narrative content (e.g., movies, television programs, user-generated content) on smartphones in one sitting, although the share of longer movie or video content (20+ minutes)

[83] For the definition of Hollywood films, see Bordwell et al. (1985) on pre-sixties cinema and Bordwell (2006) on the "postclassical" era of the late-twentieth and early-twenty-first century.

watched on mobile devices is growing.[84] Therefore, two short, each approximately nine-minute, sequences were used for the experiment from the final section of the movie, where Petite, with the help of his "accomplices," installs the wire and performs his walks. The two chosen clips were selected to fulfill the above criteria by both including semantically meaningful elements, a variety of saliences and shot lengths, a wide range of emotions, and cross-media references (e.g., written texts) to the same extent. The selected parts of the movie were also required to evoke strong emotional reactions, without being violent or showing disturbing content. Despite the fact that they are mild enough not to cause discomfort, the two clips can evoke concerns for the protagonist or even moderate symptoms of acrophobia caused by the sight of the tall buildings or the deep void beneath the WTC towers.

The clips are non-cohesive and depict different episodes of the main protagonist's feat (one each of Petite's several walks on the wire); still, they are analogous in containing both diegetic and non-diegetic speech (dialogues and first-person narration) and alternating short and long takes. Short takes seek control over the viewer's gaze, whereas long takes provide space for exploration and engagement with the protagonist's dangerous performance and the dynamics of the sequences. The clips are set mostly at the same locations—on top of the Twin Towers with additional locations in New York City in both darkness and daylight, indoors and outdoors.

The first clip is part of a storyline depicting Petite making the last preparations for the coup with the help of his collaborators and completing his very first walk across the wire. Whereas the opening scene is dark with few salient objects (set indoors and at dawn), the scenes on the top of the South and North towers, on the wire, and the surrounding streets are shot in daylight and are grey-dominated with a few striking colors against brighter backgrounds. The scenes set on the surrounding streets include a colorful mass of people and objects (e.g., yellow taxis) as well as written signs (e.g., a large clock and street signs). The tower sequences present figures chiefly in medium close-ups or wide shots against industrial or urban backgrounds. These elements draw viewers'

[84] As I argued earlier, shorter (though, perhaps more frequent) viewing sessions can be attributed to the fact that smartphones are used within the temporal framework of other activities. Still, consumer statistics show that the share of long-form (at least 20 minutes) video content of all content watched on smartphones was over 50% globally in 2018. Of all time spent consuming movie or video content on smartphones in the same year, 57% was spent watching content for 20–40 minutes and 45% watching content exceeding 40 minutes (Ooyala, 2018). However, the number of short videos (maximum 5 minutes) is still the highest (Google Consumer Barometer, 2019).

attention depending on the length of the shot and the motion of the depicted objects or figures.

The second clip starts during the last stages of the coup and presents Petite's last walk across the wire. It opens with the artist on the wire, approximately half an hour of plot duration after the end of the first clip, surrounded by policemen on both towers demanding that he finish the act and give himself up. The presence of the police and the games Petite is playing with them generates tension, in which the wire walker's success and also his life are at risk—similar to the tension during the first walk in the other clip. The second clip includes the same portion of written signs (e.g., newspaper headlines, street signs), faces and animate objects, and camera movement shooting Petite in medium close-ups and wide shots against the towers and the same industrial and urban background. The clip concludes with the end of the act, Petite's brief captivity in the lobby area of one of the WTC buildings, and an evening dinner scene where he and his friends celebrate their overwhelming success.

The two sequences depict two separate segments of Petite's nearly one-hour-long series of passes back and forth between the two towers. Both include moments of rapidly rising tension and both have a clear line of resolution with a successfully concluded walk. The semantic content of the two clips is notably similar. During most of the action in the two sequences, the visual language concentrates on the protagonist on the wire while the vertical lines of the steel beam structure of the towers present a downward dynamic, inviting spectators to follow them with their eyes into the void. Slow pans over the wire, close-ups on Petite's feet or medium close-ups on his upper body provide information about his physical and mental state (pride and fear, most typically) with eventual cuts to his accomplices and other observers. The balance between dark- and bright-toned images divides the pre- and post-coup events from the actual wire-walking, as do day and night. Whereas the first clip opens with events taking place at an indoor space at dawn and continues with the performance in daylight, the second one presents the act first and then finishes indoors in the evening. There are some narrative cues that suggest the sequential order of the two clips in the movie when watched in its entirety; still, each sequence presents a stand-alone storyline without clearly referencing the other. This made the order of the two clips reversible and suitable for measuring

participants for the same kinds of reactions while avoiding repetition and biases due to sequential effect.[85]

Apparatus and Setup

Participants in the projector condition were seated in a shielded and dimmed (but not completely dark) experiment room at a fixed distance of 180 centimeters from a 47.3-inch (120 cm by its diagonal) canvas. The movie clip was projected on the canvas at a 32.4-degree horizontal[86] and 18.55-degree vertical angle, and eye level was set to approximately the middle of the screening area. The visual angle was set in a way so as to exceed the range of angles for the mobile condition, even if participants hold the smartphone close to their eyes. For sound presentation, a pair of headphones was provided with mono sound, so that the two speakers of the headphones transmitted the exact same sound, which was necessary for technical reasons addressed below. The headphones featured no noise-cancelling function.

Modeling the parameters of smartphone spectatorship in general, the mobile condition was designed to recreate ordinary mobile viewing settings. For this reason, participants were permitted to adjust the viewing distance according to their will at the moment within the approximately 60 centimeters of distance between arm length and their eyes. The viewing distance varied between approximately 30–60 centimeters, which resulted in a horizontal angle of 11.52–22.8 degrees and a vertical angle of 6.49–12.93 degrees. For this setup, a five-and-a-half-inch (13.9 cm diagonal) OnePlus 2 smartphone was provided, running Android 6.0 with 1080x1920 pixels of screen resolution. The phone was set to airplane mode, so that the device could not generate any unforeseen distraction. The movie sequences were played on MX Player Pro video player application. MX Player Pro enables continuous playback without interruption from commercial material or pop-up windows and supports a wide range of settings and interactions with a footage, from searching to zooming in on the

[85] To further ensure the comparability of the sequences, an emotion map was created, based on test viewers' reactions, prior to choosing the final material for the experiment. For this, the concluding section of the movie, featuring Petite's coup, was divided into units containing self-standing narrative and emotional themes. The emotion map was used to determine how moving, emotional, scary, and stressful (suspenseful) the individual segments were perceived to be. The two clips were subsequently classified according to emotional valence and amplitude.

[86] For the sake of comparison, a viewer seated on the "prime seat" of a movie theater (the area sold or occupied first, approximately at the back two-thirds of an auditorium) sees the screen with an average horizontal angle of 45 degrees (Allen, 2000).

images. Although not all participants claimed to have used MX Player before, due to its similar design to other widely used mobile video players, such as the YouTube smartphone application (with which all of them claimed to be familiar), it was not deemed necessary to give separate instructions about the video player's configurations. The volume of the audio was synchronized to match that of the projector condition,[87] and the same headphones were used with mono sound.

Separate speakers and a screen were used to play the distractors. To maintain, or even enhance, subjects' sense of presence in the acoustic space of the movie, the first distractor (traffic noise) was played from a parametric (directional) speaker, which threw sound in a relatively small, concentrated area, toward where the participant was seated. Sound arrived from behind and to the left of the participant. The second and the final sonic distractors (ringing phone and chirping birds) were presented from another, regular speaker in front and to the right of the participant. A thirteen-inch external screen was used for the visual distractors (literary text and animated rectangle), which was placed in front of the participant on the left (see Figure 11). The luminance of the screen was set bright enough to be sensed, even if it was not in the viewer's visual range.

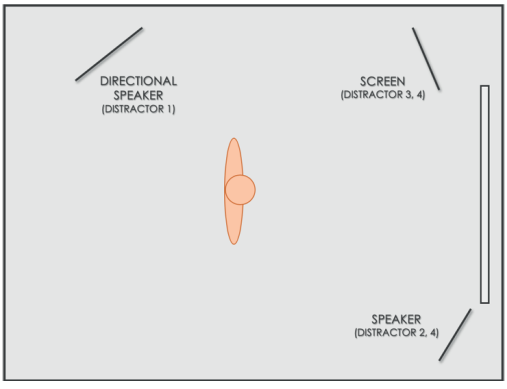


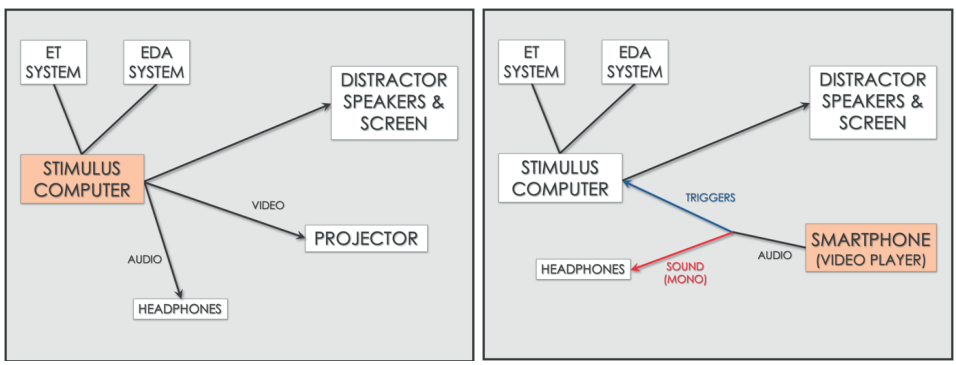
Figure 11. Experiment room setup with stimulus and distractor screens and speakers.

For the projector conditions, the primary (movie clip) and secondary (distractor) stimuli were presented from the same computer using Presentation, a stimulus-presentation software by Neurobehavioral Systems. The playback scenarios were coded to play the movie clip and, to avoid latencies between the different

[87] As subjects were free to interact with the smartphone in their hands, in one case, the volume was adjusted slightly during the trial.

types of data, send time triggers to the respective measuring software for eye tracking and electrodermal activity. These triggers signaled the start and end of the movie clip as well as the time of distractors.

As shown in Figures 12–13, for the smartphone conditions, Presentation scenarios contained all the respective stimuli, but the stimulus computer only controlled the playback of the distractors, while the video was controlled by the participant holding the smartphone. Moving the playback control to the smartphone was necessary in order to provide opportunities for interactions, such as pausing or rewinding the clip, otherwise, such interactions would have tampered with the timing of the distractors. Time triggers were embedded into the respective clips and signaled the stimulus computer and measuring software. The triggers were audible sound effects (a short ticking sound) that were edited in the movie clip's soundtrack. Thus, the original sound had to be modified in a way that one of the two sound channels only contained the triggers at given points of time, and the other contained the original (mono) sound of the movie. In order to exclude the trigger sounds from the headphones, while providing the exact same sound quality as for the projector conditions, the cable inserted into the smartphone's audio jack was split into two channels. The trigger channel was connected to the stimulus and measuring computers and the audio channel was connected to the headphones, with both speakers containing the same mono sound.



Figures 12–13. Technical setup for the projector and smartphone conditions, respectively. The primary stimulus source is marked with the orange block.

Procedure

Before the experiment, the procedure was introduced to each participant, covering the physiological measurements and the questionnaires participants would be asked to complete about their experience and narrative comprehension after each test session. In addition, participants received the necessary information about safety, and were told they could end the experiment at any time without being obliged to provide an explanation.

Each participant was tested individually to rule out the effects of social behavioral patterns, such as adjusting to others' presence or copying their reactions. In the projector conditions, participants received no specific instructions, other than to pay close attention to the movie sequence. In the mobile conditions, they were also instructed to pay close attention to the sequence, but were given the opportunity to exploit the functions of the video player application, interact with the device, and adjust the presentation of the sequence if and whenever they wished or felt the need to do so. This also allowed for changes in screen position.

Following the oral briefing, participants were seated in the experiment room to watch their assigned clip in the assigned condition and the measuring tools (eye-tracking glasses and EDA skin sensors) were applied. The eye-tracking appliance was initially calibrated with one (central) calibration point, and if the participant's gaze points showed at least approximately 0.5 degrees of deviation from the control fixation point, an additional, three-point calibration was used. Clips were presented with five seconds of black screen at the beginning to prepare the participant and to keep the screen dark to maintain the calibration of the eye tracker (the fixation points used to calibrate the eye tracker also appeared against a dark background). Another black screen appeared for five seconds after the movie clip to signal the end of the trial.

After watching each assigned movie clip, participants were asked to complete a questionnaire that measured emotional engagement and presence and to answer questions regarding their comprehension of the movie content. Participants were given no specific time within which they had to complete the questionnaire. This was to avoid pressuring them from rushing through the questions, and to make sure their priority was focusing on completing the survey according to their best knowledge. In addition, this short break between trials was used to change the setup between conditions, as well as to recalibrate the measuring devices when necessary. In total, the experiment took no more than

45 minutes per participant including briefing, the two trials, and filling out the questionnaire after each trial.

Measurement Methodology

Physiological Factors of Engagement

During the free viewing task, two sets of physiological data were monitored in a noninvasive way: oculomotor behavior (eye movements) and electrodermal activity. Measuring physiological factors of engagement with the narrative, the following indices were considered: the amount of time participants' gazes were on the respective screen, the frequency of fixation points that fell outside of the screen, the amplitude and frequency of saccades, the dispersion of fixation points, and changes in skin conductance. These data were recorded to be analyzed either for an entire trial or at specific time intervals, which were hallmarked by participants' adjustments of the screen content in the mobile conditions and the four distractors in the interrupted conditions. Due to the very low occurrence of participant–device interactions, however, changes in physiological responses during interactions were excluded from the final analysis.[88] Seven trials from the analysis of oculomotor behavior and 14 trials from the analysis of electrodermal activity were excluded due to technical errors or insufficient data.

For measuring oculomotor behavior, a pair of head-mounted SMI 1 mobile eye-tracking glasses were used with a sampling rate of 30 Hz. As opposed to screen-based eye trackers (measuring gaze in relation to screen content), the mobile eye tracker enabled participants to move freely and interact naturally with the smartphone, while registering both on-screen and off-screen gazes at objects near and far. It tracked binocularly, meaning that it simultaneously recorded gaze data from both eyes. This eye tracker model fits on a participant's head like a regular pair of glasses and the lens design allows for adjustments to correct near- and far-sightedness up to 4.00 diopters. Wearing contact lenses has no impact on measurements or quality of data. Participants' behavior and

[88] Interactions observed in this experiment were limited to two events: obtaining film playback details (e.g., timeline) and changing sound settings (e.g., volume), which occurred in seven trials. The timing of these events showed no correlation, and coincided with external distraction in only one case.

device use was recorded by the high definition video recorder built into the eye tracker, additionally, activity on the smartphone screen was monitored through screen capture, using AZ Screen Recorder Android application.

As a result of sonic and visual distractions and the small size of the screen, participants in the smartphone and interrupted conditions were thought to less likely maintain constant focus on the movie. Consequently, scanpaths were expected to leave the screen for a longer proportion of the trial time in the presence of distractions and when watching the footage on the small screen, as opposed to the large one. This was expected, on one hand, as a result of visual and sonic distractors drawing attention away from the screen based on ecological relevance, duration, and neutrality. On the other hand, this assumption was grounded in the idea that the smaller the visual angle a screen covers, the less focused viewers are on the screen.

In order to measure the likelihood of participants' gazes leaving the screen, a single dynamic area of interest (AOI) was defined that covered the respective screen on the eye tracker's recording, irrespective of head movements and changes in the visual field. The respective AOI for each trial was set manually and adjusted frame by frame to follow changes in position as there was a lack of linear or automatically predictable movements. Being present throughout the entire trial, the AOI enabled distinguishing among all gaze activities that fell on or outside of the screen.

Based on hypothesis 1 of this experiment, that smartphone spectatorship in unenclosed environments produces a lower level of engagement than it does when viewers watch larger, isolated screens, the total time gazes spent on the screen during a trial (dwell time in the AOI) was expected to be shorter under the mobile and interrupted conditions. This assumes that the small screen and the distracting stimuli induce more intense attention oscillation between the screen (the moving-image content) and the surrounding space and other stimulus sources. Correspondingly, the frequency of off-screen fixation points[89] was anticipated to be higher in the mobile and interrupted conditions.

Saccadic amplitude measures the distance the eye travels between two fixation points and is of particular interest in assessing skewness in participants exploring the screen and off-screen areas. Also, a decrease in saccadic amplitude can reflect upon the difficulty of cognitive tasks (May, Kennedy, Williams, Dunlap, & Brannan, 1990), for instance, difficulty focusing on the screen's content in interrupted conditions. This is true if the gaze stays on the screen during a distractor, and can be a potential response when observing static

[89] Gaze events with a minimum duration of 80 milliseconds were treated as fixations.

images or objects (e.g., when reading). However, when following dynamic visual information and looking up from the screen to seek the source of external stimuli, saccades travel longer to a fixation point outside the screen, meaning that saccadic amplitude increases with off-screen gaze. According to the latter assumption, the minimum, maximum, and average saccadic amplitude values were expected to be lower during uninterrupted, small-screen viewing.

Saccadic frequency (the number of saccades in a second, also known as saccadic rate) suggests conclusions regarding cognitive load and arousal, such that the increase of cognitive load decreases saccadic frequency and a higher level of arousal increases it. An increase in cognitive load signifies a more demanding and difficult task, but in terms of arousal, increasing saccadic frequency reflects a higher level of immersion and involvement in either diegetic or real-life events. Cognitive load may lead to unreliable conclusions in the case of dynamic visual stimulation (i.e., moving images). However, arousal can affect saccade frequency such that the smartphone and interrupted conditions produce lower values of saccadic frequency.

According to the second main hypothesis of this experiment, fixation points were assumed to be more concentrated in the central area of the image on the smartphone than they are on the large screen. To quantify and compare the variation (dispersion) of fixation coordinates, the standard deviation of all fixation coordinates was calculated for each trial. Here, a lower standard deviation value reveals that these points are distributed in a smaller area around the central point. The dispersion of fixation points decreases with smaller screens, so mobile conditions produce lower fixation dispersion.

Electrodermal activity measures changes in skin conductance, which are closely related to emotional arousal, immersion, and attention. EDA was measured with sensors attached to participants' fingers, which were connected to a digitizer (MegaWin ME6000 Biomonitor) with a sampling frequency of 1000 Hz. The sensors were placed on two fingers of a participant's non-dominant hand, on an area with a high density of sweat glands that would not interfere with carrying out the experiment tasks. Emotional arousal concerns a state of being reactive and capable of processing information.

EDA produces a high variability of baseline levels in skin conductance in and between individuals depending on physiological responsiveness and skin type. For this reason, relative differences were calculated between a baseline value and individual data points throughout the trials. The baseline was the average EDA value of a five-second window (5000 data points) immediately preceding the start of the trial, when participants were not engaged in any tasks and were

looking at the black screen. Electrodermal activity measurements provided two sets of variables for skin conductance levels: overall EDA score, the mean value per trial relative to the baseline value and EDA scores during distraction in the interrupted conditions. A comparison of the corrected overall scores across conditions was anticipated to determine changes in arousal that originated from engagement with the movie clip and the diegetic events. This suggests that similar to oculomotor indices of engagement, EDA values would be lower in the mobile conditions and during interrupted viewing.

To identify event-related changes in skin conductance, separate EDA scores were calculated for the specific distraction events using the same method. Instead of drawing further conclusions about engagement, these variables, along with the event-related clusters of off-screen fixations, were used to compare the effects of the four distractors. For these cases, no particular hypotheses were developed for viewing conditions; these tests served to differentiate the effects of distractors based on their modality (sonic and visual), ecological relevance (in the diegetic or physical space and no ecological relevance), and neutrality (neutral and urgent).

Subjective Ratings and Narrative Comprehension

To complement the physiological measurements' deficiency in determining the quality of emotional reactions (i.e., whether they are negative or positive), post-screening questionnaires were used. In the first questionnaire, participants evaluated their subjective impressions of their viewing experience on a 10-point Likert-type scale with values ranging from *true* to *not at all true*.[90] Wording followed first-person statements with phrases as "*I* empathized with the characters." or "*I* felt…" This questionnaire aimed to reveal engagement with the narration through the following indicators: presence in the diegetic space (at Philippe Petite's coup), empathy toward the characters, and levels of feeling scared, moved, and nauseated. Here, narrative engagement (presence), emotional devotion (empathy) and the mental and bodily manifestations thereof (fear, being moved, nausea) were used for drawing conclusions about participants' immersion into the diegetic space.

To evaluate subjective ratings of engagement, each item of the questionnaire was analyzed as an individual variable. An additional variable was calculated to

[90] The questionnaire is based on Gross and Levenson (1995), Witmer and Singer (1998), Qin, Rau, and Salvendy (2009).

determine individual averages of these ratings. A reliability test revealed an adequate consistency between these items with a Cronbach's alpha of 0.772. According to the related hypotheses, all these variables (rating of feeling present, empathy, feeling scared, moved, nauseated, and the average of these) were expected to have lower values for small-screen and interrupted trials.

The second questionnaire measured narrative comprehension and was used to reveal information about the effects of distraction and attention to the narration at specific points in time. This questionnaire featured statements relating to semantically meaningful narrative information and details that were obscured by the distractors in the interrupted conditions. Statements, for instance, included that "The wire-walker had an injury on one of his feet." or "The clock on the wall showed past 8 o'clock." The selected details were either primary or secondary to the main story line (walking on the wire) and were presented in the given sequence only once. The possible answers to each question were "yes," "no," and "I don't know." The "I don't know" option was necessary to avoid obliging participants to aim to give the expected answer when they failed to perceive that particular segment of the clip (perhaps because their attention shifted from the screen).

Answers were classified and analyzed as "correct," "incorrect," and "I don't know." Participants failing to choose the correct answer reflects being distracted either by the respective distractors (in interrupted conditions) or by the screen size (in mobile conditions). To compare participants' performance, an overall score was calculated for each trial: each correct answer equaled one point and "I don't know" answers equaled zero point; for incorrect answers one point was deducted. The overall score was normalized (where necessary, missing values were replaced with the participant's mean score). As lower scores mark less engagement with and concentration on the movie clip, as well as divided attention, participants in interrupted conditions and on the smartphone were expected to score lower.

Results

The Effects of Viewing Conditions on Physiological Responses and Subjective Ratings

To test hypotheses 1 and 2, a Generalized Linear Mixed Model analysis was performed, which included screen type and the presence or absence of distractors as independent variables. This analysis served the same purpose as a standard two-way analysis of variance, but it also allowed for including participant as a random effect to control for the incomplete design (each participant being measured in two of the four conditions) while maximizing statistical power. The analysis was run for each dependent variable of physiological responses, self-ratings, and narrative comprehension (hypothesis 1) and for dispersion of fixations (hypothesis 2). This determined the following effects of viewing conditions (for a summary, see Table 3):

Dwell time on screen. Running the model on gaze data, results showed no significant interaction between the effects of screen size and the presence or absence of distraction on the time spent viewing the screen ($F(1, 65) = 0.004$, $p = .949$). Screen type ($F(1, 65) = 1.958$, $p = .166$) and distractions ($F(1, 65) = 1.25$, $p = .268$) had no significant main effects either. Results of this test indicated that dwell time was not affected by screen type or the presence of distractors.

Frequency of off-screen fixations. The frequency of off-screen fixations is a complement to dwell time and measures the proportion of gaze events that fall outside the screen. Similar to dwell time, testing frequency of off-screen fixations showed no significant interaction ($F(1, 65) = 2.213$, $p = .142$) and no significant main effects of screen ($F(1, 65) = 0.791$, $p = .377$) or distraction ($F(1, 65) = 0.773$, $p = .382$). The frequency of off-screen fixations was not affected by screen type or the presence of distractors.

Saccadic amplitude. Saccadic amplitude data was measured to determine the distance the eye traveled between fixation points and was divided into three variables, minimum, maximum, and average saccadic amplitude. Average and maximum values showed no significant interaction between the effects of screen

size and the presence or absence of distraction ($F(1, 64) = 0.003$, $p = .953$ and $F(1, 65) = 0$, $p = .998$). On average and maximum values, neither screen type ($F(1, 64) = 1.658$, $p = .203$ and $F(1, 65) = 0$, $p = .996$), nor distraction ($F(1, 64) = 1.06$, $p = .307$ and $F(1, 65) = 0$, $p = 1$) had significant main effects.

Although average and maximum values were independent of viewing conditions, a significant main effect was observed between screen types when comparing the minimum values for saccadic amplitude ($F(1, 65) = 4.137$, $p = .046$). The minimum values of saccadic amplitude were significantly higher for the projector conditions than for mobile conditions. No significant interaction ($F(1, 65) = 0.872$, $p = .354$) and no significant main effect of distraction ($F(1, 65) = 1.696$, $p = .197$) were observed for minimum saccadic amplitude.

Saccadic frequency. Measuring the number of saccades per second, the results revealed that variables interacted in their effect on saccadic frequency ($F(1, 65) = 4.306$, $p = .042$). Average saccadic frequency was significantly higher for the mobile screen during uninterrupted viewings, meaning that participants performed more saccades when watching the clip on the mobile screen than on the projector screen in the presence of distractions.

Electrodermal activity. Changes in electrodermal activity compared to the individual baseline values measure arousal. EDA values showed a main effect of screen ($F(1, 58) = 5.78$, $p = .019$), where the average EDA level was significantly higher for the projector conditions than mobile conditions. This result indicates that participants were more aroused during projector watching. No significant interaction ($F(1, 58) = 0.014$, $p = .906$) or main effect of distraction ($F(1, 58) = 0.071$, $p = .791$) was observed for this variable.

Subjective ratings of viewing experience. The model was run for the separate items of the self-report questionnaire (sensation of presence, empathy, feeling scared, moved, and nauseated) as well as the individual average scores. In terms of the average value, distraction approached a significant main effect ($F(1, 72) = 3.611$, $p = .061$): average ratings were higher in the uninterrupted conditions than in the interrupted conditions. Results showed no significant interaction ($F(1, 72) = 0.01$, $p = .922$) and no main effect of screen ($F(1, 72) = 1.55$, $p = .217$).

In the case of presence and empathy ratings, significant main effects of distraction were observed: ratings for both items were significantly higher in the

uninterrupted conditions than in the interrupted conditions ($F(1, 72) = 4.644$, $p = .034$ and $F(1, 72) = 6.645$, $p = .012$). However, results showed no significant interactions between the effects of screen type and distraction on presence and empathy ($F(1, 72) = 1.102$, $p = .297$ and $F(1, 72) = 0.017$, $p = .898$) and the screen had no significant main effect on these items ($F(1, 72) = 1.247$, $p = .268$ and $F(1, 72) = 3.331$, $p = .072$).

Subjective ratings of feeling scared ($F(1, 72) = 0.004$, $p = .948$), being moved ($F(1, 72) = 0.032$, $p = .859$), or experiencing nausea ($F(1, 72) = 0.375$, $p = .542$) showed no significant interaction between screen size and the presence or absence of distraction. Neither had screen type ($F(1, 72) = 0.477$, $p = .492$; $F(1, 72) = 3.731$, $p = .057$; $F(1, 72) = 0.18$, $p = .673$) or distraction ($F(1, 72) = 0.212$, $p = .647$; $F(1, 72) = 0.722$, $p = .398$; $F(1, 72) = 0.314$, $p = .577$) main effects on these variables.

Narrative comprehension. For testing narrative comprehension, the overall scores indicating individual performance were compared. Narrative comprehension scores showed a significant interaction between screen size and distraction ($F(1, 72) = 4.811$, $p = .032$). Participants scored significantly higher in uninterrupted conditions than interrupted conditions when watching the movie clip on the smartphone.

Gaze dispersion. In order to test the second hypothesis, the effect of screen size on the dispersion of fixations, the same model was run. As expected, screen type had a main effect on gaze dispersion ($F(1, 64) = 26.229$, $p < .001$). Fixations were spread on a significantly larger area when watching the projector screen than when watching the mobile screen. No significant interaction ($F(1, 64) = 0.502$, $p = .481$) or effect of distraction ($F(1, 64) = 0.024$, $p = .877$) was observed.

Table 3
Generalized Linear Mixed Model, Interactions and Main Effects of Viewing Conditions on the Physiological Indices of Engagement, the Subjective Indices of Engagement, Narrative Comprehension, and Gaze Dispersion

	Results			
Measure	Interaction	Main effect of screen	Main effect of distraction	Interpretation
Dwell time	F(1, 65)=0.004	F(1, 65)=1.96	F(1, 65)=1.25	Dwell time is independent of viewing conditions.
Frequency of off-screen fixations	F(1, 65)=2.21	F(1, 65)=0.79	F(1, 65)=0.77	Frequency of off-screen fixations is independent of viewing conditions.
Average saccadic amplitude	F(1, 64)=0.003	F(1, 64)=1.66	F(1, 64)=1.06	Average saccadic amplitude is independent of viewing conditions.
Minimum saccadic amplitude	F(1, 65)=0.87	F(1, 65)=4.14*	F(1, 65)=1.70	Minimum saccadic amplitude is significantly higher in the projector conditions than in the mobile conditions.
Maximum saccadic amplitude	F(1, 65)=0	F(1, 65)=0	F(1, 65)=0	Maximum saccadic amplitude is independent of viewing conditions.
Saccadic frequency	F(1, 65)=4.31*	F(1, 65)=0.07	F(1, 65)=0.05	Saccadic frequency is significantly higher in the uninterrupted mobile condition than in the uninterrupted projector condition.
Electrodermal activity	F(1, 58)=0.01	F(1, 58)=5.78*	F(1, 58)=0.07	EDA level is significantly higher in the projector conditions than in the mobile conditions.
Experience rating – presence	F(1, 72)=1.10	F(1, 72)=1.25	F(1, 72)=4.64*	Presence ratings are significantly higher in the uninterrupted conditions than in the interrupted conditions.
Experience rating – empathy	F(1, 72)=0.02	F(1, 72)=3.33	F(1, 72)=6.65*	Empathy ratings are significantly higher in the uninterrupted conditions than in the interrupted conditions.

Experience rating – feeling scared	F(1, 72)=0.004	F(1, 72)=0.48	F(1, 72)=0.21	Ratings of feeling scared are independent of viewing conditions.
Experience rating – feeling moved	F(1, 72)=0.03	F(1, 72)=3.73	F(1, 72)=0.72	Ratings of feeling moved are independent of viewing conditions.
Experience rating – feeling nauseated	F(1, 72)=0.38	F(1, 72)=0.18	F(1, 72)=0.31	Ratings of feeling nauseated are independent of viewing conditions.
Experience rating – average	F(1, 72)=0.01	F(1, 72)=1.55	F(1, 72)=3.61	Average ratings are independent of viewing conditions.
Narrative comprehension	F(1, 72)=4.81*	F(1, 72)=0.04	F(1, 72)=4.19*	Narrative comprehension scores are significantly higher in the uninterrupted mobile condition than in the interrupted mobile condition.
Gaze dispersion	F(1, 64)=0.50	F(1, 64) =26.23*	F(1, 64)=0.02	Gaze dispersion is significantly higher in the projector conditions than in the mobile conditions.

Note. Eye tracking measures were performed using a sample size n = 69, electrodermal activity measures n = 62, and experience ratings and narrative comprehension n = 76.
* $p < .05$

Distraction Effects

To observe the specific role of distractions on physiological responses and test the third hypothesis, two additional sets of data were calculated: frequency of off-screen fixations and electrodermal activity changes (compared to baseline values) during the periods of the four distractors. For the former, the number of fixations were counted for the time period of each distractor; for the latter, a calculation of deviation preceded the comparative tests. This was necessary as EDA responses often appear following a one- to three-second delay after the onset of a stimulus and values during these delays must be excluded from the analysis. To generate an accurate comparison among distractors and their interaction with screen type, a Generalized Linear Mixed Model analysis was performed, where the distractors (distractors 1–4) and screen type (mobile and projector) were the independent variables. Conditions without distractors were

excluded from this analysis. The analysis measured the interactions between screen types and distractors, as well as their main effects on off-screen fixations and EDA values. Results are summarized in Table 4.

Table 4
Generalized Linear Mixed Model, Interactions and Main Effects of Distractors

	Results		
Physiological reactions	Interaction	Main effect of screen	Main effect of distractor
Frequency of off-screen fixations	F(3, 132) = 20.97*	F(1, 132) = 27.51*	F(3, 132) = 47.54*
Electrodermal activity	F(3, 116) = 1.92	F(1, 116) = 22.99*	F(3, 116) = 3.00*

* $p < .05$

Frequency of off-screen fixations during distractors. The significant interaction between screen type and distractors ($F(3, 132) = 20.965$, $p < .001$) revealed that the frequency of off-screen fixations was significantly higher in the projector condition than in the mobile condition during the third distractor (see Table 5). However, during the first distractor, the opposite tendency was observed: the frequency of off-screen fixations was significantly higher in the mobile condition than in the projector condition. In the mobile condition, distractors 1 vs. 3 and 1 vs. 4 were significantly different: as shown in Table 6, distractor 1 produced lower values than distractor 3 and 4. In the projector condition, every distractor but 2 and 4 were significantly different; distractor 3 produced the highest and distractor 1 the lowest frequency of off-screen fixations.

Table 5
Frequency of Off-Screen Fixations Mean Values by Distractor Types

					Mobile		Projector	
Distractor type		n	M	SE	M	SE	M	SE
Distractor 1	Traffic noise	35	0.04	0	0.07	0	0	0
Distractor 2	Ringing phone	35	0.16	0.06	0.16	0.08	0.16	0.08
Distractor 3	Written literary text	35	0.90	0.08	0.32	0.11	1.48	0.12
Distractor 4	Chirping birds and animated rectangle	35	0.29	0.06	0.23	0.06	0.35	0.10

Note. M = mean value; SE = standard error.

Table 6
Frequency of Off-Screen Fixations Mean Values by Screen Types

				Distractor 1		Distractor 2		Distractor 3		Distractor 4	
Screen	n	M	SE	M	SE	M	SE	M	SE	M	SE
Mobile	60	0.20	0.04	0.07	0	0.16	0.08	0.32	0.11	0.23	0.06
Projector	80	0.50	0.04	0	0	0.16	0.08	1.48	0.12	0.35	0.10

Note. M = mean value; SE = standard error.

In terms of stimulus modality (sound or image), these results revealed a pattern of no difference as these differences are applicable to all modality combinations only for the projector condition (see Figure 14 for differences between distractors). In the projector condition, the visual distractor (distractor 3) produced significantly higher off-screen frequency than sonic distractors (distractors 1 and

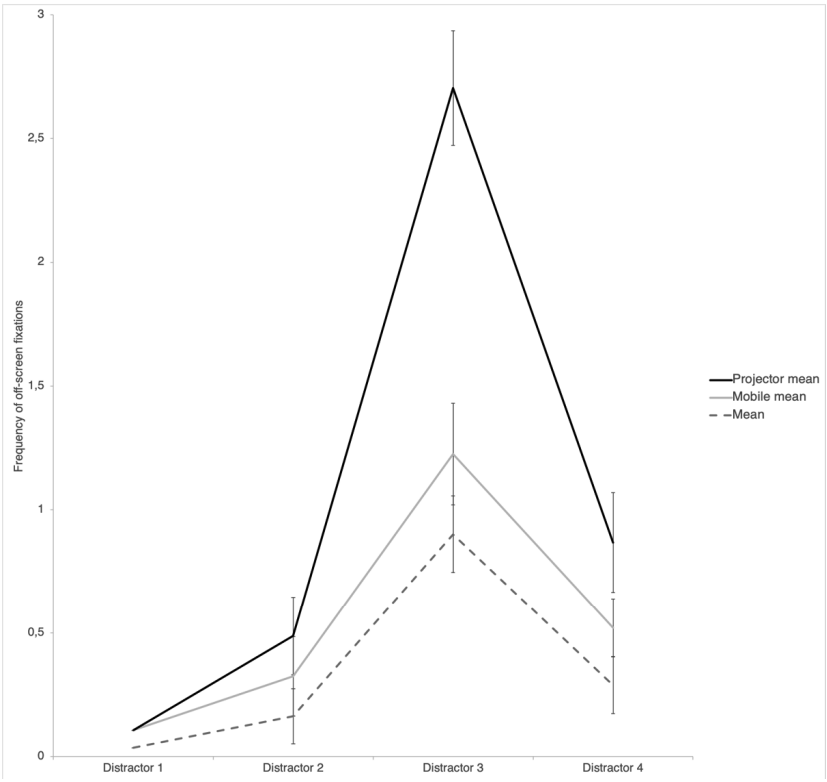


Figure 14. Differences between distractors' mean values of off-screen fixations. Error bars show ± 1.96 SE. Points are offset vertically so that error bars are visible.

2) indicating that participants looked off the screen more when the distractor was visual compared to sonic. Similarly, regarding the urgency of the distractors, the differences were only significant in the projector condition. There, urgent distractors were more distracting: the neutral distractor (distractor 1) produced a significantly lower off-screen frequency than urgent distractors (distractors 2, 3, and 4).[91] The results showed a general tendency with regard to ecological relevance for both screen types: the distractor with ecological relevance to the diegetic space (distractor 1) produced significantly less fixations off the screen than distractors with no ecological relevance (distractor 3 and 4). In addition, in the projector condition, the first distractor even produced a significantly lower off-screen frequency compared to the distractor with ecological relevance to the physical space (distractor 2).

Electrodermal activity changes during distractors. As presented in Table 4, EDA values showed a main effect of distractors ($F(3, 116) = 3.001$, $p = .033$), where these values were significantly lower during distractor 1 than during distractor 3 and distractor 4 (see Table 7; for differences between distractors, see Figure 15). Correspondingly, screen type also had a significant main effect ($F(1, 116) = 22.992$, $p < .001$): EDA values were significantly higher for the projector condition than the mobile condition (see Table 8).

Table 7
Electrodermal Activity Mean Values by Distractor Types

Distractor type		n	M	SE
Distractor 1	Traffic noise	31	0.05	0.02
Distractor 2	Ringing phone	31	0.09	0.02
Distractor 3	Written literary text	31	0.09	0.001
Distractor 4	Chirping birds and animated rectangle	31	0.12	0.01

Note. Since the Generalized Linear Mixed Model analysis showed no significant interaction between screen type and distractor type, only mean values of distractor type are reported. M = mean value; SE = standard error.

[91] The comparison was also applied to distractor 4, where both urgent and neutral stimuli were used. This is because the urgent stimulus was assumed to overwrite the neutral stimulus; therefore distractor 4 was treated overall as urgent.

Table 8
Electrodermal Activity Mean Values by Screen Types

Screen	n	M	SE
Mobile	52	0.05	0.01
Projector	72	0.12	0.01

Note. Since the Generalized Linear Mixed Model analysis showed no significant interaction between screen type and distractor type, only mean values of screen type are reported. M = mean value; SE = standard error.

The pairwise comparison revealed a significant difference only for ecological relevance, where the distractor with ecological relevance to the diegetic space (distractor 1) produced significantly lower EDA values than the third and fourth distractors which had no ecological relevance.

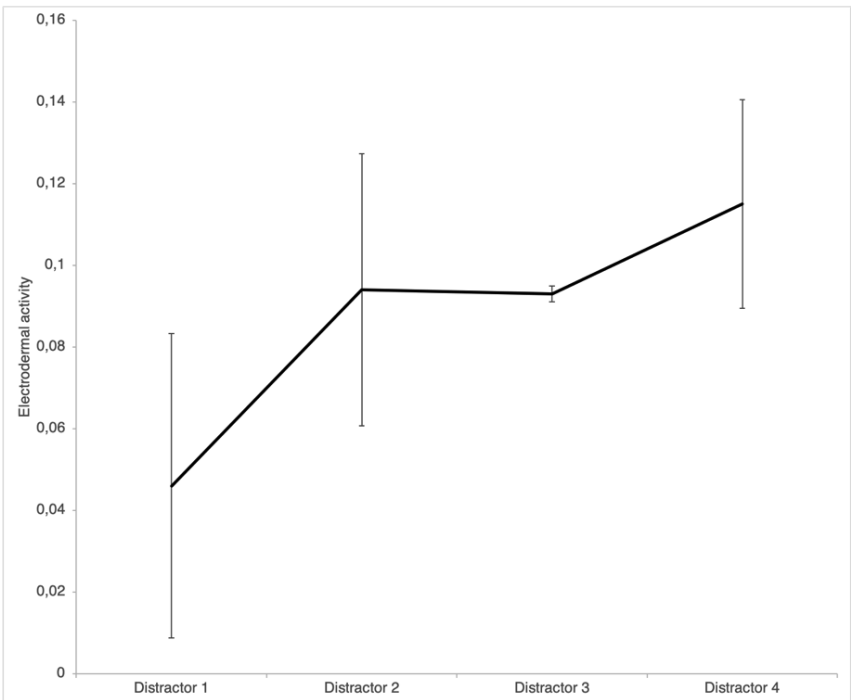


Figure 15. Differences between distractors' mean values of electrodermal activity. Error bars show ± 1.96 SE.

Findings

The main goal of this experiment was to present a potential methodological framework for assessing the behavioral formulas of smartphone spectatorship and to test whether the smartphone's small size and spectatorship in unenclosed spaces influence physiological responses, subjective ratings of engagement, and comprehension of an audiovisual narrative. The hypotheses of this experiment were built upon the theoretical discussions in the preceding chapters with the aim to point toward preliminary conclusions and explanations regarding the extent that smartphone spectatorship employs a learnable skill based on references to earlier screening and media devices.

According to the first hypothesis, there is a significant effect of screen and environmental distraction on physiological responses (gaze behavior and electrodermal activity), subjective ratings, and narrative comprehension. In terms of gaze behavior, the results revealed that viewers are similarly unlikely to transfer their visual attention to the surrounding space from a small portable screen as from a large stationary screen, even in the presence of environmental distractions. A possible explanation that coincides with the main point of Chapter Three is that viewers accustom to screens and surroundings that fosters immersion into the diegetic space. This may also refer to the fact that smartphone viewers can enhance their viewing experience by choosing the best possible screen positioning. Observing the frequency of fixations that fell outside the screen through the viewpoint of distractions supports this explanation: participants kept their attention on the screen to the greatest extent during the distractor with ecological relevance to the diegetic space. This effect was weaker with regard to distractors that presented urgent and visual stimuli.

Immersion and narrative engagement rely upon the extent to which viewers shut out their surrounding physical environments, and may be measured by the level of arousal in response to narrative events. Indices of arousal (saccadic frequency and electrodermal activity) showed that screen size has an impact on emotional engagement. Arousal was expected to be lower in the mobile conditions than in the projector conditions. For average EDA levels, the results corresponded with this expectation.[92] The results of comparing saccadic

[92] Skin sensors are sensitive to changes in electrodermal activity caused by changes in one's thermoregulation (e.g., sweating caused by emotional arousal), but also by muscle movements. Although participants wore the sensors on the fingers of their non-dominant hand which were less likely to be in motion during the trials, it cannot be ruled out that some EDA values are results of muscle movement.

frequencies revealed the opposite tendency in uninterrupted conditions: saccadic frequency was higher in the mobile condition than in the projector condition in the absence of distractors, while no significant effect was observed in the presence of distractors. This could signal a divergence between EDA and saccadic frequencies in terms of indexing arousal, or even the effects of other factors on saccadic frequency, such as the properties of visual information: the fluctuating luminance of a visual sequence can affect gaze behavior, which was not accounted for in this experiment. These factors could also include cognitive load: smartphone viewers faced a different task than projector viewers, given their involvement in shaping the attributes of viewing or simply the less comfortable viewing situation. In addition, this may be attributable to the fact that viewers' gazes travel shorter distances on a small screen: screen content can be explored while making smaller changes to gaze position, which can increase the frequency of saccades. The dispersion of gaze corresponds to this latter conclusion: the location of fixation points on the screen was more spread out in the projector conditions. This seems to comply with the fact that, even though larger displays produce a larger retinal image, a greater proportion of the image stays outside the fovea, the area which provides sharp vision.

Results showed no direct impact of screen size on narrative comprehension; yet, it became clear that distraction can decrease comprehension of and attention to narrative events. Distractors had a similar effect on the subjective assessments of presence and empathizing with the fictional characters. The effect of distractors on these indices confirms that environmental stimuli (especially if they lack ecological connection to the diegetic space) make it more difficult to maintain constant attention to a movie than during uninterrupted viewing and that unrelated sound or visual effects can disrupt affective involvement with the content. To account for such momentary attention deficits during smartphone spectatorship in unenclosed spaces, smartphone-specific audiovisual content may demand updated storytelling strategies, either to eliminate these effects with more engaging sensory presentations or to make up for attention oscillation with, for instance, repeated information or longer takes.

Whereas the conclusions above are plausible, the methods used in this experiment cannot rule out that subjective ratings and comprehension measurements are content specific and cannot be generalized beyond the movie sequences used here. In order to determine the specific effects of certain semantic information (faces, bodies, urban environment, or even the effects of acrophobia), additional tests are required. It is also important to note that discrete (post-experiment) questionnaires are also highly subjective. Moreover, while a narra-

tive questionnaire provides valuable proof about how much information a viewer absorbs, due to the interim between watching and recollecting, it may not capture objective information and respondents may be biased by social and cultural expectations. The same applies to subjective ratings: completing a complex task, for instance, watching a movie sequence, requires meaning construction containing both emotional and semantic components. Therefore, mental abilities, such as memory and language knowledge, can consequently influence self-reporting.[93]

Given the low number of events when participants interacted with the smartphone, this experiment did not provide conclusions on the effects of interactive watching outlined in Chapter Two. This low number can possibly be attributed to the laboratory environment and the fact that the smartphone used did not belong to the participants, which factors most likely affected their behavior. The limited cases of interaction lead to no general conclusions on the role and effect of touchscreen interaction; however, they can serve as a starting point for further empirical research that account for the role of interactivity.

The laboratory environment can also affect the validity of results; results from an experiment conducted in situ, in a natural environment, may result in a higher level of generalizability regarding viewer behavior. Nevertheless, due to a lack of control over distractions, experiments conducted in natural environments are difficult to recreate, and no matter the method used, the observer effect still cannot be ruled out. Measurements of environmental effects produce more verifiable results in laboratory settings: all participants are exposed to the same stimuli and all stimulus channels can be entirely reproduced.

In experiments of this sort, even the measuring devices' size and design can influence responses. The mobile eye-tracking glasses' temples may have blocked the view of some angles; when, for instance, a participant was looking downwards at the smartphone screen, he or she may have had difficulties in noticing visual distractors appearing on an external screen at approximately a 45-degree

[93] Employing methods for collecting ratings simultaneously with the task (during the movie clip) disrupts the viewing experience and consequently viewers' state of immersion, which also biases responses. Drawing conclusions about feelings of presence in real time using reaction time measures by, for instance, instructing participants to raise a hand when certain information is presented on screen (as in A.-L. Cohen et al., 2015) or by marking a scale (e.g., Troscianko et al., 2012), as well as more autonomic channels, such as facial expression, provide solutions to this problem in similar studies. Yet, these methods only lead to valid results for certain segments of a sequence and would not capture a participant's overall feeling of presence.
Offering a different solution, developments are ongoing to create applications and add-ons for portable smart devices enabling real-time measurements. See, for instance, the AttentiveVideo interface for monitoring the effects of video advertising (Pham & Wang, 2017). For an overview of non-real-time measurements, see IJsselsteijn et al. (2000).

angle upwards. This was compensated for by setting the brightness of visual distractors to be observable even when a participant's head was not turned toward them.

This experiment aimed to measure individual viewing strategies. However, another potentially significant factor of unenclosed viewing environments is the proximity and social presence of others. Testing viewers in groups would likely produce different results.

Hollywood-style movies—*The Walk* is no exception—have a globally known style that includes unequivocal visual and sonic references, which imply strong exogenous control and therefore similar behavior. Still, ethnic and cultural differences across participants can also bias the results. All the participants in this experiment were affiliated with a Finnish university and had permanently lived in Finland or other European countries for at least several months, which implies that they were more or less equally exposed to Western-type cultural products. Nevertheless, it is not impossible that their background may have influenced both subjective ratings and physiological responses.[94]

Oculomotor data proved that the smartphone's small screen influences gaze dispersion. Moreover, smartphone viewers are more likely to be affected by external distractions, which effect was foremost observable in terms of subjective ratings and narrative comprehension. These results confirm the importance of regarding smartphones as distinct media tools that encompass specific practices and clear-cut impact on viewing experiences. Yet, overall, the results show limited effect of screen type on the indices of engagement, which is attributed to the result of extensive smart device usage which particularly defines the age cohort represented by the participants of this experiment.

[94] In a few cases, participants of South Asian or African origin were non-responders in terms of electrodermal activity. Although there have been studies measuring the number of sweat glands and level of skin conductance across skin types (e.g., L. C. Johnson & Landon, 1965), no studies show clear evidence that skin tone affect electrodermal activity measures or that finger sensors require sensitivity adjustments. In these cases, conductive gel was applied on the skin, but trials that nonetheless produced no measurable electrodermal activity were excluded from analysis.

Discussion: Screen culture 2.0

Summary

Perceiving a three-dimensional space projected on a two-dimensional surface that is located in another three-dimensional space induces a peculiar spatial and affective experience, caused by the illusion that one is inside and outside the projected space simultaneously. Being one of the most fundamental characteristics of spectatorship, this is the eternal clash that accompanies the study of moving images. What, however, is propelled by technological innovations of mobile screens is that engagement is no longer restricted to mental and emotional devotion, but linked to tangible bodily involvement too.

When operating smartphones, users interact concurrently with physical and virtual objects. Gestures are defined by the device's properties and laws of physics, whereas interactions with user interfaces follow the principles of a constructed, programmed sphere. Users' fingers move over a blank, smooth surface that nevertheless induces complex plots. Interaction with a smartphone's interface goes hand in hand with the abstraction of this uncanny relationship that connects the user with content through multi-level involvement.

By combining media theory, cognitive, phenomenological, and behavioral approaches, this dissertation has followed a methodology specifically developed to investigate smartphone spectatorship, including interactive watching, usability, changes in sensory perspectives, consumption and participation, as well as viewing in unenclosed spaces. Following this methodological path, I have concluded throughout that smartphone spectatorship is demonstratively based on two key factors: medium- and platform-specific knowledge and the spectator's bodily and mental presence. Mental presence not only mobilizes cognitive processes to access and assemble a coherent narrative, it also informs the means, time, and degree of bodily intervention.

Initiating my study of smartphone spectatorship, the theoretical discussion highlights the cinematic heritage of spectatorship and the ways in which spectating on smartphones reinterprets and reestablishes the temporal, spatial, and cultural frameworks of screenings in movie theaters. I argued that references to other media are essential features of smartphones that enable various functions through modal or cultural references to other computerized or physical tools—among others, video players and movie screens. Moreover, references to other screens define how spectators adopt and update cognitive mechanisms according to their viewing circumstances, while striving to initiate affective connections with the screened sequences. These insights reveal the technological and social mechanisms that define smartphone spectatorship and the framework through which existing media provide references for establishing new behavioral schemes and (re)contextualizing them in light of consumer demands. In other words, new platforms emerge in reference to older ones, providing initial frameworks for developing media practices. But even though smartphone spectators encounter specific spatial and temporal dimensions, which clash with both social and cognitively effective behavioral formulas, an immersive state can be attained by creating quasi-cinematic settings and (literally or mentally) closing out elements irrelevant to narrative comprehension.

While cinematic references are integral to viewing experiences, they are also complemented by games and interactive screenings. This is an essential point, given that smartphone spectatorship is hallmarked by deviances from cinema's institutional and material foundations, but nonetheless resembles the illusion of control in interactive screenings and the interactive setups of video game play: movie watching becomes personalized in response to viewer preferences and environmental demands. Video game players react to the fictional features of a game's narrative and perform virtual actions using motor gestures on physical devices, such as keyboards, joysticks, or touchscreens. Likewise, spectators watching content on mobile screens often alter its presentation. These interventions are performed through gestures such as changing the screen's position, adjusting the sound, image, and playback settings, or attaching or detaching ear- or headphones. The motivations for these interactions were identified following two strategies: responses to external distractions (e.g., social obligations or distracting sound or visual effects) that distance spectators from the content and interrupt immersion, or responses to internal motivations, such as curiosity, which increase involvement.

Observing smartphone spectatorship through the lens of multisensory and interactive viewing provides an overview of how spatial signals trigger and form a synergy between device, stimuli, and spectator. Lacking clear social or material boundaries, smartphones as movie or video screens are characterized by parallel activities and the spaces to which these parallel activities belong. Investigating interactions and smartphone spectators' active bodily participation, I specifically aimed to connect the options and motivations for influencing screenings to the encounters with personalized, yet fragmented narrative presentations. Arranging these elements into the phenonarratology model of narrative presentation on platforms enabling interactions revealed two fundamental points: the pliancy of the spectator's sensory scope and the sensation of physical involvement in self-curating the screening. Based on the integration of sensory modalities, the spectator can interact with the smartphone in a way that changes the spatial composition of sound or image. In the case of sound, attaching or removing headphones or adjusting sound settings can result in changes in the intensity or surround effect, which modifies the balance of the movie's sound and sonic indices from the physical space. Visual perception is circumscribed by the outline of visible elements, which include the composed, filmic mise-en-scène, screen and software design, and the surrounding environment. This outline changes when the viewer adjusts the screen's position, visual angle, aspect ratio, color, luminance, or other visual characteristics. Accessing extra-filmic information from the surrounding space also affects associations, neural and motor responses, and narrative comprehension.

Subjectivity and the role of the user's body shape participation not only in content consumption, but also in content creation. Smartphones of the late 2010s are equipped with cameras and microphones capable of recording in near-professional or professional quality. This fact is well documented by the number of professional filmmakers who experiment with mobile filmmaking for full-length feature films. Participatory culture, however, is chiefly marked by amateur users who use smartphones equally in reflexive and inductive ways. While the former points to an abundance of discussion forums and the widespread manifestation of fan culture (for example, in social media) the latter stems from streaming to distribution and, most crucially, includes recording, too. Besides colonizing technology and moving-image recording, the presence of amateur users on the production end deconstructs the formal foundations of the cinematic heritage in favor of an accessible new cult of storytelling. When recording moving images on a smartphone, perspectives are no longer defined by the camera and its attributes; it instead conforms to the framework of the

screen. Vertical sequences shot on smartphones are fitting examples of this. Amateur content favors quantity, accessibility, and condensed information transmission—framed by the screen and the way how a smartphone conforms to one's body and hand. The very notion that filmmaking and spectatorship, as such, can accommodate mobile platforms (and vice versa) is highly debated by film professionals. Nonetheless, the emergence of mobile film and video content confirms that consumers are capable of adopting the social and behavioral frame of spectatorship under a variety of circumstances.

Meaning-making in accordance with the sensory scope and mechanisms of storytelling is informed by motion, social interactions, film literacy, and medium-specific knowledge. Among the prerequisites for spectatorship are sensory abilities, textual understanding, and an immediate phenomenological connection to the screening platform. Sufficient knowledge of a medium automatizes interactions and thus, realizes a direct perceptual connection between the viewed content and the viewer's body and mind. Following this, I developed a theory, termed as adaptive effect, which denotes the process by which spectators or users acquire a strategy (skill) to filter out context-irrelevant stimuli. Context-irrelevant stimuli can be anything that originates from the design, customs, or technological features of the screen, which are independent of and fail to cohere with the audiovisual content. Somewhat parallel to classical conditioning, adaptive effect entails a learning curve along which neural and cognitive links are established to filter relevant stimuli and contextualize them within the framework of a given narrative. One assumption behind this theory is that spectators apply "old" habits to new scenarios, such as affective devotion to a cinematic narrative during smartphone spectatorship. Yet, stimuli that often originate from the non-conformity of smartphones or an unenclosed viewing space can indeed influence narrative comprehension, either by distracting the viewer or by inducing associations, for instance, when sounds, scents, lights, or people in the environs are perceived as linked to the movie narration.

Explaining environmental sensory information and social associations, I identified and examined three attributes of external (non-filmic) stimuli in relation to moving-image content and spectatorship, to contemplate whether (and if so, how) stimulus properties define the likelihood of attention oscillations between a watched movie and other stimuli. The first attribute is a stimulus' intensity, which provides information about its location and distance. Intensity, such as the magnitude of visual or sonic effects, can inform the viewer of the stimulus' spatial position in relation to the movie's diegetic space or its relevance—for instance, whether it is close enough to have an impact. This latter

remark is point of departure for the second attribute, namely, ecological relevance, which is the contextual value of a stimulus to the narrative. In other words, it signals the extent of embeddedness in the diegetic space. Industrial noise, for instance, bears little ecological relevance to a documentary film about wildlife, whereas traffic sound coheres well with an urban action scene. The third attribute is neutrality, which indicates whether the stimulus requires immediate attention and a corresponding reaction. For instance, a ringing telephone requires a reaction, unlike, for example, background music. As the empirical results from the experiment I conducted on this topic show, stimuli with ecological relevance to the diegetic space are less likely to distract viewer's attention than urgent stimuli, if they are similarly intense.

The most essential aim of the experimental verification was to isolate the characteristics of spectatorship on smartphone screens in environments that may introduce interruptions with visual and sonic distractions in comparison to large, stationary screens in enclosed viewing spaces. I created the methodology with the most prevalent form of smartphone movie consumption in mind; watching non-mobile-conforming content (a feature film) in an unenclosed space. Data were collected to measure attention, narrative comprehension, and feelings of presence: these comprised measurements of physiological reactions (eye tracking and electrodermal activity), a narrative comprehension questionnaire, and self-evaluation.

The experiment results revealed the effects of viewing settings on oculomotor behavior (minimum saccadic amplitude and saccadic frequency), electrodermal activity, self-reports regarding engagement (feeling of presence and empathy), as well as narrative comprehension. The fact that screen size and the presence or absence of sonic and visual distractions did not significantly influence every index of engagement with the narrative is attributed to technology adoption and extensive smartphone usage in which millennials (the cohort representing the experiment participants) greatly surpass other adult generations. An effective technology adoption also suggests that frequent smartphone use make users confident adjusting viewing parameters for comfortable, enjoyable, and cognitively efficient viewing. Distraction nevertheless had an effect on some indices, namely, participants rating their feelings of presence and empathy toward fictional characters lower and scoring lower in the comprehension test when there were distractors. However, it may be that obtaining subjective ratings and narrative comprehension only after each trial is responsible for these results, as the presence or absence of distractions had no significant effect on physiological responses.

The physiological impact of screen size was validated by fixation points covering a smaller area of the screen and by the least distance participants' eyes traveled between fixation points being smaller when watching content on the smartphone. Yet, the overall findings of this dissertation are that smartphones' small size can be compensated for by positioning the screens and making other adjustments to create an uninterrupted viewing experience; but that smartphone viewers are nevertheless susceptible to distraction, which causes changes in viewing practices, interaction with the device, and the cognitive processing of incoming information.

Understanding the Smartphone Spectator

This dissertation began with a promise to review movie and video consumption on portable devices through the case of smartphones. The importance of studying moving-image consumption on smartphones is attributed to the rapid technological advancements and the fast-growing number of users, who use, shape, and profit from smartphone spectatorship. It was necessary to develop a new framework, first, to summarize and question current new media discourses and, second, to develop new approaches for inquiring into the cognitive and behavioral elements of spectatorship. Along my analysis, I evaluated embodiment and the role of the spectator's body in interactivity, learning and usability, and the cinematic roots of spectatorship, to contextualize patterns of sensory exposure, attention, narrative comprehension, and engagement.

With a focus on the ways in which viewers adapt to novel viewing circumstances, I proposed that spectatorship is composed of schematic information clusters that create meaning out of sensory information and conceptual knowledge. Information originating from sensory stimuli, usability formulas, and technical features is accumulated into batches of data. During this process, some information is intentionally perceived, while other information is processed unconsciously and this amalgam is what makes for comprehension. Narrative comprehension rests on the ability to make sense out of schemata and schema systems accessed through the senses and prior experiences, including earlier sensory encounters and social formulas. Perceived information is filtered according to relevance: information deemed relevant becomes embedded in an overall narrative, whereas irrelevant information is treated as a distraction. In practice, this means that information (consciously or unconsciously) considered valuable

is processed, regardless if it originates from a movie's narration or any other source. In the case of irrelevant stimuli, stimulus properties, such as intensity or neutrality, as well as contextual indices, as ecological relevance, define the level of attention the viewer pays to them.

The attributes that distinguish smartphones from other screens or screening apparatuses include size, weight, and software design. The screens are notably smaller than most other screens used for screening moving-image content, which spectators frequently compensate for by holding them close to their eyes. Smartphones' small mass allows for mobility, that is, they can easily be moved as much as the spectator's body enables. Smartphones afford portability too; they are easy to carry and can be used in any spaces where one goes about. This leads to another fundamental point: apart from the small size of images and limited sound capacity, it is also the—often unsuitable—environment that defines viewing experiences on smartphones. Bodily (haptic and kinesthetic) control thus manifests in nonce configurations of stimulus sources and sensory organs, on the one hand, and in phenomenological contact with content, on the other. The former factor affects perception in a way that viewers execute changes to the attributes of image and sound that, for instance, alter stimulus magnitude or immersive effect. In other words, as a matter of illustration, if a screen is held close to one's eyes, the visual stimulus is perceived to be more intense (of increasing magnitude) and covers more of the visual field than if it is held far away. Consequently, it blocks out a greater proportion of irrelevant, distracting elements from one's vision. This, moreover, effectuates increasing immersion and lessens the effects of the surrounding environment. Or, if raising the volume, sound from the speakers, earphones, or headphones are more likely to block out environmental noise.

Two trajectories must be distinguished to delineate the factors in play in smartphone spectatorship and to give room for extending this research: one targets the spectator as the agent who adopts medium- or content-specific behavior and the other focuses on visual and sonic compositions and the ways in which academic and industry research could, in the long run, examine and facilitate mobile content optimization. This is crucial, as content production is now increasingly influenced by consumers and consumer culture. Due to the thriving popularity and growing impact of participation, there seems to be a need to map the fast-changing portfolio of genres and film form.

Human Factors

When it comes to the human component, spectatorship is characterized by an affective connection between spectator and screen and screened content. Smartphone spectatorship, again, requires the complex, embodied presence of the spectator, who, besides perceiving audiovisual content in connection to a story being told, defines both the material and contextual framework of the screening. Affective presence, in both social and phenomenological terms, gives rise to two spectatorial routines: the attempt or desire to improve access to narrative information and the obligation to conform with the social and behavioral benchmarks of physical spaces. The former point reflects the traditions of spectatorial behavior, which have roots in earlier, institutional, cultural, and behavioral practices. This implies that spectators strive for an ideal setting that accommodates sensory access to a movie's narration. Designated screening rooms operate with a specific spatial arrangement designed to guide attention to the visible and audible elements of the narration. The lack of such arrangements in unenclosed spaces is compensated for by direct bodily control over the presentation of stimuli. The spectator moves the screen and sets the volume so that they balance out these deficiencies. Nonetheless, the social composition of unenclosed spaces holds constant reminders of the roles and duties that the spectator's presence requires, by transmitting sensory and conceptual indicators for interacting with people and objects also inhabiting them. The elements of physical space are fundamental to social interactions and inform the spectator of what are considered suitable reactions and interactions. Even so, those elements may not harmonize with the fictional narrative and this contrast between tasks and roles can cause distraction. As a contrast, in a cinema, the spectator's main role is to spectate and a darkened room decreases access to other sensory information and behavioral frameworks.

In regards to smartphones, actions and reactions refer to moments of interaction and adjustments afforded by the design, usability, and practices of the device and the various video and media player applications. Bodily intervention through moving the screen and adjusting the sensory channels also affects the quantity and type of content, as well as the place and time of moving-image consumption. Different viewers react to given situations in different ways. Yet, focusing on the tangible properties of stimuli, narrative formulas, storytelling, viewing, and mental processing allows for observing and analyzing the spectator as a biological and social entity, defined by various mental processes and behavioral frameworks. Whereas certain reactions are salient in only a limited

number of cases (depending on cultural or individual formulas), others have wider applicability.

Movie Content, Mobile Content

Besides the smartphone spectator's skills and familiarity with moving images, mobile screening platforms and their interfaces also define spectatorial practices. Content, however, often fails to fit into the specific configuration of the visual and auditory scope, as viewing practices often fail to conform to the content. At the time of the widely referenced Cannes Netflix-gate of 2017, Pedro Almodóvar said that "the size of the screen should not be smaller than the chair you're sitting on," as that would impair the spectator's humble position in relation to moving images (Donadio, 2017). The reason for the scandal was that two titles produced by Netflix were presented at the Cannes Film Festival without being (and without plans to be) included in theatrical distribution. The response of Netflix Chief Content Officer, Ted Sarandos was a reference to the group of movie consumers "that has seen every great movie ever made on a phone" to justify the presence of online streaming companies in the line-up of such a prestigious film festival (Donadio, 2017). Cinema is obviously not cinema anymore, if films are watched on five-inch screens while sitting on a train or strolling the streets of cities. Moreover, by enabling such freedom, spectatorship increasingly distances itself from the substantial and economic features of cinematic spectatorship. Viewing platforms (whether portable screens or streaming sites) make profit-oriented responses to the increasing demand for constant entertainment—however, not to replace, but to extend cinema's boundaries.

The technological transformation of spectatorship is evident: film in the form of a celluloid roll has almost ceased to exist, and was replaced by digital formats, computer files, which often have no physical containers. Analysis of film and cinema should also follow this transition. Cinema is more than a sole screening location; it is the screening apparatus, the institution, behavioral norms, just as much as spectatorship is about more than just watching a movie. Spectatorship includes the choice of a movie and engaging with it through many layers, including competences and cognitive processes, which layers are now extended by streaming sites, personal devices, and online fan networks.

Future Trends of Smartphone Spectatorship: The Birth and Growth of Mobile Media

Smartphones and smartphone spectatorship may not be around for as long as cinema has been; they may be transitioned in a shorter period than we might assume. But, then, where are the trends of cinema and smartphones taking us? Can we ever regard movies as points of departure for media discourses or smartphones as the Swiss Army knives of entertainment? Nicolas Maigret, Clément Renaud, and Maria Roszkowska (2015–2018) propose conceptual designs of fantastic devices that combine everyday tools and objects with mobile phones to reflect upon the trajectories of technology, mirroring user demands all around the world (for some examples, see Figure 16). The *Swiss Knife Phone*, for instance, features the actual blades of an army knife, the *Razor Phone* mixes a mobile phone with an electric razor; other designs incorporate a lighter, a radio, or a video projector, or are designed to fit in a small pocket or to operate a sound system with Chinese communist songs. The idea behind the project, entitled *Shanzhai Archeology*[95] is to reflect upon technological anticipations in a new way. By combining past and future through blending old-fashioned, (now considered) low-quality devices with contemporary demands, it aims to challenge over-standardized, Western approaches to innovation.

Figure 16. Mobile phones from the *Shanzhai Archeology* project. Screenshots.

Shanzhai Archeology criticizes Western traditions that strive for an unnecessary uniformity in mobile phones and, thus, questions the legacy of these devices as something that encapsulate collective user needs. Fake electronic products made

[95] The Chinese expression "Shanzhai" refers to pirates or outlaws working in remote mountain villages and is a reference to counterfeit consumer goods and trademark infringement.

of cheap or reused parts have been produced in unimaginable amounts. While the term *shanzhai* is rather disparaging, these products are often romanticized on the Western market, even though they are closely connected with unfair labor conditions and a disregard for environmental issues. The designs of our mobile devices are the fantasy of individuals; still, they form social dynamics. The product concepts in Maigret, Renaud, and Roszkowska's work refer to marginal needs, while deliberately avoiding the fabrication of new needs. Whereas smartphones constantly update patterns of immersion, devotion, and addiction by featuring novel sets of functions requiring a growing amount of attention, the devices presented in the *Shanzhai Archeology* collection reflect a desperate craving for disconnecting from the online sphere.

Although *Shanzhai Archeology* projects a negative image of the present social embeddedness of portable devices, neither ubiquity nor scholarly interest is in doubt when it comes to smartphones. I will leave the judgments and analysis of the dangers of overwhelming information intake to other studies. What, however, is crucial to conclude with is the potential methods required for movies to address audiences in an era of audiovisual abundance. Will the regime of the "flatties" end soon? Will visual and sonic compositions aim for hitherto-unseen (and unheard) realism to continue competing with the physical world in satisfying human senses and mental processes? Cinema and movies have changed cyclically by introducing darkened screening rooms, surround sound, three-dimensional screenings, and then returned to realistic representations and 3D again and extended the boundaries of spectatorship to personal and portable devices and immersive virtual reality. Despite frequent statements concerning the abnormality of watching movies on smartphones that were or were not produced on or intended for such a platform, smartphone spectatorship is nevertheless an inherent part of post-cinematic inquiries. This is the case not only because smartphones are omnipresent in the Western world, but also because they affect content production and film consumption patterns. New trends in consumption reflect changes in the amount of time spent in front of screens and moving-image content, which has an impact on narrative comprehension, cognitive processes, emotional engagement, and may have long-term effects on learning, imagination, and cognitive skills, which must be detected and constantly monitored.

Central to cognitive film studies and the analysis of spectatorial behavior and narrative processing is the process of learning. These entail previous encounters with moving-image storytelling, contextual knowledge of objects, characters, and structural frameworks, and familiarity with screening platforms, with the

latter fast-forwarding our society toward immersive media. The same processes typify smartphone spectatorship that is a notable example of rapidly evolving technology and corresponding social and behavioral processes. Smartphones infuse film and media literacy with portability and mobility to induce physical and mental involvement in shaping both content displays and viewing modes. Thus, the individualism, with which we react to our hyperstimulating environment and which shapes our momentary connections with moving images, challenges the uniformity of smartphone media.

Smartphone manufacturers now are ready to update smartphones' hardware to optimize them for increasingly immersive but not less comfortable media experiences. By working toward standardizing screen aspect ratio from the most common, 16:9, to 18:9,[96] manufacturers have started to produce smartphones that are longer and slimmer to accommodate larger screens, but not at the expense of ergonomics and usability. The aspect ratio of 18:9 is that of Univisium (Storaro, 1998), proposed in the 1990s as a universal film format to unify theatrical, television, and new media film formats and the post-theatrical distribution of movies on various platforms without sacrificing screen content. The new aspect ratio for smartphones is highly efficacious for applications using vertical visualization (such as social media apps or browsers). What is more, in landscape mode, it improves movie watching and even virtual reality experiences by presenting larger images (and therefore, more content) for each eye, when using VR headsets designed for smartphones.

To achieve an increasingly immersive viewing experience and compensate for the drawbacks of various viewing environments (including unenclosed spaces and even the lack of fellow spectators), providers offer virtually created viewing environments for watching movies. The Oculus Cmoar VR Cinema application enables choosing from a variety of virtual screening locations, as an ordinary screening room or IMAX theater, while Netflix VR provides a home cinema experience in a virtual living room, where ambient lights dim when the movie starts and turn back on when it ends or the screening is paused. Moreover, users can connect with other users to communicate and watch content together, in the same virtual environment, even if they are in different physical locations. These developments make using non-space-bound, physically private, portable media players and streaming platforms more like traditional and typically social cinematic and home video experiences. Virtual screening room applications run

[96] Manufacturers such as Huawei, OnePlus, and LG (including the Google Pixel series) have released smartphones with an 18:9 aspect ratio since 2017. In early 2019, Sony announced production of new models with even longer (or wider) screens, with an aspect ratio of 21:9, which is a screen size marketed as being suitable even for professional filmmaking.

on pocket-sized portable devices and relocate large-screen viewing into the space of portable media, while blocking out the physical environment.

My research has explored different aspects of smartphone film and media experiences and offered tools for analyzing moving-image consumption, user interventions, narrative comprehension, immersion, and even provided a sample of the latest trends in smartphone film production. These tools, besides heralding the ubiquity of both smartphones and moving-image content, illuminate the plasticity and personal quality of spectatorship—in time, space, form, and modes of viewing. This draws attention to the fast-changing media landscape that occupies an ever-greater role in everyday life and offers a bedrock for film and new media research to keep up with the pace of change.

Svensk sammanfattning

Smartphone Cinematics: En kognitiv studie av filmupplevelse i smartphones

Den här avhandlingen diskuterar den vanligt förekommande, men samtidigt lite undersökta, filmupplevelse i smartphones, och ger ett teoretiskt, empiriskt och metodiskt bidrag till studiet av digital media och filmkultur. Det speciella med filmupplevelse i smartphones jämfört med upplevelse på biograf är att smarta mobiler har få institutionella, sociala, tidsmässiga eller rumsliga begränsningar. Genom att kombinera medieteoretiska, kognitiva, fenomenologiska och beteendemässiga metoder undersöker denna avhandling interaktiv filmupplevelse, den smarta mobilens användbarhet, förändringar i sensoriska perspektiv, filmkonsumtion och deltagande samt filmtittande i miljöer som inte ursprungligen är planerade för just detta.

Inledningsvis förs den teoretiska diskussionen genom att filmupplevelses cinematiska arv lyfts fram och därmed frågan om hur smarta mobiler förhåller sig till de tidsmässiga, rumsliga och kulturella ramarna för filmvisningar på biograf. Smarta mobiler har modala eller kulturella referenser till andra medier eller fysiska apparater, bland annat videobandspelare och filmdukar. De cinematiska referenserna definierar en affektiv relation till filmen på skärmen, vilket möjliggör att stänga ute element som inte är relevanta för berättelsen.

Filmupplevelsen i mobil är baserat på tittarens kroppsliga och mentala närvaro. Hur en audiovisuell berättelse spelas upp varieras ofta på mobila skärmar: i enlighet med yttre påverkan bestämmer tittarens behov och möjligheter att förändra filmen. Denna interaktion utförs genom välbekanta handrörelser, till exempel genom att ändra skärmens position, justera ljud, bild och uppspelningsinställningar på enheten eller ta bort hörlurarna. Behovet av interaktion kan ha

två olika källor, dels individuella behov, såsom nyfikenhet, vilket ökar känslan av närvaro, dels externa distraktioner (t.ex. sociala skyldigheter eller distraherande ljud eller visuella effekter från smartphonen eller det omgivande rummet) som distanserar tittaren från filmen och avbryter känslan av närvaro.

Komponenterna i meningsskapande processen är rörelse, sociala interaktion, filmkunnighet och mediespecifik kunskap. I allmänhet genererar tillräcklig kunskap automatiskt interaktion och innebär en direkt perceptuell koppling mellan innehållet och betraktarens kropp och sinne. För att testa denna process utvecklade jag modellen ”adaptive effect,” vilket beskriver hur tittare eller användare antar en effektiv strategi (kunnighet) för att filtrera bort stimuli som inte är relaterad till filmens narrativ.

För att kunna förklara miljöns sensoriska och sociala påverkan på tittandet och för at kunna definiera sannolikheten för att uppmärksamhet flyttar mellan filmen och annan stimulus, identifierade och granskade jag tre egenskaper hos sensoriska stimuli. Den första egenskapen är stimulus intensitet, vilken ger information om stimulus källa och avståndet till den. Intensitet kan också informera om relevans (till exempel, om stimuluskällan är tillräckligt nära för att påverka). Denna anmärkning är utgångspunkt för den andra egenskapen, nämligen ekologisk relevans, vilket är stimulus kontextuella värde för berättelsen: ekologisk relevans berättar hur väl en stimulus kan bäddas in i filmens diegetiska värld. Industriellt ljud, till exempel, har liten ekologisk relevans för en dokumentärfilm om vilda djur, medan trafikljud väl stämmer överens med en scen från en actionfilm. Den tredje egenskapen är neutralitet, vilket indikerar om en stimulus kräver omedelbar uppmärksamhet eller någon typ av reaktion. Exempelvis kräver en ringande telefon en reaktion i motsats till bakgrundsmusik. Som mina empiriska resultat visar, är det mindre sannolikt att stimuli med ekologisk relevans till den diegetiska världen avleder betraktarens uppmärksamhet än att akuta stimuli gör det.

På grund av interaktion och aktivt kroppsdeltagande möter mobilanvändaren ofta ett fragmenterat, men ändå personligt berättande. Enligt min ”phenonarratology”-modell, påverkar fysiskt engagemang både det tittaren ser och hör. Genom sensoriska modaliteter (huvudsakligen syn, ljud och känsel), kan tittaren interagera med mobilen på ett sätt som ändrar ljudets eller bildens komposition. När det gäller ljud kan anslutning av hörlurar eller justering av inställningar resultera i ändringar i intensiteten eller surround-effekten, vilket ändrar balansen mellan filmens ljud och ljudet från miljön runt omkring. Vad tittaren ser är en kombination av synliga element på skärmen, som inbegriper den koreograferade filmiska mise-en-scènen, och den aktuella mobilskärmens design samt de

visuella intrycken från det omgivande rummet. Denna kombination ändras när man justerar skärmens position, bildstorlek, bildförhållande, färg, luminans eller andra skärmegenskaper.

Mobilanvändarens kropp deltar inte enbart i tittande utan även i innehållsskapande. De nyaste smartphonemodellerna från slutet av 2010-talet är utrustade med kameror och mikrofoner som kan spela in med professionell eller så gott som professionell kvalitet. Detta deltagande märks framför allt hos amatöranvändaren som använder smartphones både reflexivt och induktivt. Medan det reflexiva deltagandet pekar på diskussionsforum och på en ”beundrarkultur,” handlar det induktiva deltagandet om allt från ”streaming” till distribution men framför allt handlar det om inspelning. Användargenererat innehåll gynnar kvantitet och tillgänglighet och inspelningen blir påverkas av hur mobilen positioneras, exempelvis hur mobilen passar i användarens hand. Vertikala sekvenser som spelas in på mobilen är lämpliga exempel på detta.

Den teoretiska analysen av filmupplevelse, det kroppsliga engagemanget samt påverkan från miljön där filmtittandet i mobil sker ledde till ett beteendevetenskapligt experiment. Syftet med experimentet var att jämföra filmupplevelse i smartphones i miljöer som kan leda till störningar med filmupplevelse framför stora filmdukar och biografliknande utrymmen utan störningar. Experimentet genomfördes med icke-mobilanpassad film i ett utrymme med olika stimuli och deltagarna som tittade filmen i mobil fick själva anpassa uppspelandet och skärmens position. Deltagarna mättes i fyra undersökningssituationer: de tittade på ett 9 minuter långt filmklipp antingen på en mobilskärm eller projektorskärm med eller utan ytterligare ljud- och bildstimuli. Experimentet mätte uppmärksamhet, narrativ förståelse och känslan av närvaro genom ögonrörelse- och hudresponsmätningar (elektrodermal aktivitet), självutvärdering samt narrativt förståelsetest.

Experimentresultaten visade att visningssituationen (i mobil eller på stor duk med eller utan distraktioner) påverkar några av variablerna för ögonrörelse, hudrespons, självutvärdering och narrativ förståelse. Att skärmstorleken och ytterligare stimuli inte påverkade alla variabler förklaras med att millenniegenerationen (personer födda från tidigt 1980-tal till slutet av 1990-talet, som representerade experimentdeltagarna) har vant sig väl vid mobilenheter och att den här kunskapen gör att visning på mobila skärmar kan blir lika bekväm, trevlig och effektiv som visning på stora dukar. Däremot påverkade ytterligare ljud- och bildstimuli några andra variabler, nämligen att deltagarna bedömde sin känsla av närvaro och empati med filmens karaktärer lägre och de fick

mindre poäng på förståelseprovet när de hade tittat på filmen med ytterligare stimuli.

De övergripande resultaten av den här avhandlingen är att användare kan kompensera smartphonens lilla skärm genom skärmpositionering och tillfälliga anpassningar för en oavbruten visningsupplevelse. Men smartphone-filmtittare är ändå mottagliga för distraktioner beroende på distraktionernas egenskaper. Avhandlingen har dessutom belyst det snabbt föränderliga medielandskapet som upptar en ökande del av människors vardag och har erbjudit en grund för film och medieforskning för att undersöka dess förändringar och konsekvenser i mer detalj framöver. Nya trender i filmkonsumtionen speglar förändringar i innehållet samt i den tid som avsätts framför skärmar, vilket påverkar narrativ förståelse, kognitiva processer, känslomässigt engagemang och vilket kan ha långsiktiga effekter på inlärning, fantasi och kognitiv utveckling som ständigt måste utforskas.

References

Books and Articles

Agar, J. (2003). *Constant touch: A global history of the mobile phone*. Cambridge: Icon Books.

Alexander, N. (2017). Speed watching, efficiency, and the new temporalities of digital spectatorship. In P. Hesselberth & M. Poulaki (Eds.), *Compact cinematics: The moving image in the age of bit-sized media* (pp. 103–112). New York, London: Bloomsbury Academic.

Altman, R. (1992a). Material heterogeneity of recorded sound. In R. Altman (Ed.), *Sound theory, sound practice* (pp. 15–31). New York: Routledge.

Altman, R. (1992b). Sound space. In R. Altman (Ed.), *Sound theory, sound practice* (pp. 46–64). New York: Routledge.

Auer, K., Vitouch, O., Koreimann, S., Pesjak, G., Leitner, G., & Hitz, M. (2012). *When music drives vision: Influences of film music on viewers' eye movements.* Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, Thessaloniki, Greece.

Augé, M. (1995). *Non-places: Introduction to an anthropology of supermodernity* (J. Howe, Trans.). London: Verso.

Bálint, K., & Tan, E. S. (2015). "It feels like there are hooks inside my chest": The construction of narrative absorption experiences using image schemata. *Projections, 9*(2), 63–88.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*(8), 617–629.

Barker, J. M. (2009). *The tactile eye: Touch and the cinematic experience*. Berkeley: University of California Press.

Bayne, T., & Montague, M. (Eds.). (2014). *Cognitive phenomenology*. Oxford: Oxford University Press.

Bazin, A. (1967). *What is cinema?* (Vol. 1). Berkley: University of California Press.

Benjamin, W. (1997). *Charles Baudelaire: A lyric poet in the era of high capitalism*. London: Verso.

Bergson, H. (2002). Concerning the nature of time. In K. Ansell-Pearson & J. Mullarkey (Eds.), *Key writings* (pp. 205–219). New York: Continuum.

Berry, D. S., Kean, K. J., Misovich, S. J., & Baron, R. M. (1991). Quantized displays of human movement: A methodological alternative to the point-light display. *Journal of Nonverbal Behavior, 15*(2), 81–97.

Berry, M., & Schleser, M. (Eds.). (2014). *Mobile media making in an age of smartphones*. New York: Palgrave Macmillan.

Bezdek, M. A., & Gerrig, R. J. (2017). When narrative transportation narrows attention: Changes in attentional focus during suspenseful film viewing. *Media Psychology, 20*(1), 60–89.

Bezdek, M. A., Gerrig, R. J., Wenzel, W. G., Shin, J., Pirog Revill, K., & Schumacher, E. H. (2015). Neural evidence that suspense narrows attentional focus. *Neuroscience, 303*, 338–345.

Bickhard, M. H. (2009). Interactivism: A manifesto. *New Ideas in Psychology, 27*(1), 85–89.

Bickhard, M. H. (2017). *Representing is (for) what?* Paper presented at the 13th SweCog conference, Uppsala, Sweden.

Bolter, J. D., & Grusin, R. (1999). *Remediation: Understanding new media*. Cambridge: MIT Press.

Bonath, B., Noesselt, T., Martinez, A., Mishra, J., Schwiecker, K., Heinze, H.-J., & Hillyard, S. A. (2007). Neural basis of the ventriloquist illusion. *Current Biology, 17*(19), 1697–1703.

Bordwell, D. (1985). *Narration in the fiction film*. London: Methuen.

Bordwell, D. (1986). Classical Hollywood cinema: Narrational principles and procedures. In P. Rosen (Ed.), *Narrative, apparatus, ideology* (pp. 17–34). New York: Columbia University Press.

Bordwell, D. (1989). A Case for cognitivism. *Iris*(9), 11–40.

Bordwell, D. (2002). Intensified continuity visual style in contemporary American film. *Film Quarterly, 55*(3), 16–28.

Bordwell, D. (2006). *The way Hollywood tells it: Story and style in modern movies*. Berkeley: University of California Press.

Bordwell, D., & Thompson, K. (2001). *Film art: An introduction* (6. ed.). New York: McGraw Hill.

Bordwell, D., Thompson, K., & Staiger, J. (1985). *The Classical Hollywood cinema: Film style & mode of production to 1960*. New York: Columbia University Press.

Boucsein, W. (2012). *Electrodermal activity*. Boston: Springer.

Branigan, E. (1992). *Narrative comprehension and film*. London: Routledge.

Broeren, J. (2009). Digital attractions: Reloading early cinema in online video collections. In P. Snickars & P. Vonderau (Eds.), *The YouTube reader* (pp. 154–165). Stockholm: National Library of Sweden.

Brown, B. (2004). Thing theory. In B. Brown (Ed.), *Things* (pp. 1–22). Chicago, London: University of Chicago Press.

Bruno, G. (2014). *Surface: Matters of aesthetics, materiality, and media*. Chicago, London: University of Chicago Press.

Bruns, A. (2009). *From prosumer to produser: Understanding user-led content creation*. Paper presented at Transforming Audiences, London, United Kingdom.

Buckland, W. (2000). *The cognitive semiotics of film*. Cambridge: Cambridge University Press.

Burch, N. (1979). *To the distant observer: Form and meaning in the Japanese cinema*. Berkeley: Univeristy of California Press.

Burch, N. (1982). Narrative/Diegesis—thresholds, limits. *Screen, 23*(2), 16–33.

Calleja, G. (2011). *In-game: From immersion to incorporation*. Cambridge: MIT Press.

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

Carroll, N. (2003). *Engaging the moving image*. New Haven, London: Yale University Press.

Casetti, F. (2008a). *Eye of the century: Film, experience, modernity*. New York: Columbia University Press.

Casetti, F. (2008b). The last supper in Piazza della Scala. *Cinéma & Cie, 11*(2), 7–14.

Casetti, F. (2011a). Back to the motherland: The film theatre in the postmedia age. *Screen, 52*(1), 1–12.

Casetti, F. (2011b). Cinema lost and found: Trajectories of relocation. *Screening The Past,* (32). http://www.screeningthepast.com/2011/11/cinema-lost-and-found-trajectories-of-relocation/

Casetti, F. (2012). The relocation of cinema. *NECSUS,* (2). http://necsus-ejms.org/the-relocation-of-cinema/#top

Casetti, F., & Sampietro, S. (2012). With eyes, with hands: The relocation of cinema into the iPhone. In P. Snickars & P. Vonderau (Eds.), *Moving data: The iPhone and the future of media* (pp. 19–32). New York: Columbia University Press.

Cass, J. L. (1930). The illusion of sound and picture. *JSMPE*(14), 323–326.

Chabris, C. F., & Simons, D. J. (2010). *The invisible gorilla: And other ways our intuitions deceive us*. New York: Crown.

Chateau, D., & Moure, J. (Eds.). (2016). *Screens*. Amsterdam: Amsterdam University Press.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America, 25*(5), 975–979.

Chion, M. (1983). *Guide des objets sonores: Pierre Schaeffer et la recherche musicale*. Paris: Buchet/Chastel.

Chion, M. (1994). *Audio-vision: sound on screen*. New York: Columbia University Press.

Christie, I. (2016). The stuff of screens. In D. Chateau & J. Moure (Eds.), *Screens* (pp. 70–79). Amsterdam: Amsterdam University Press.

Coëgnarts, M. (2017). Cinema and the embodied mind: Metaphor and simulation in understanding meaning in films. *Palgrave Communications, 3*. http://dx.doi.org/10.1057/palcomms.2017.67

Cohen, A.-L., Shavalian, E., & Rube, M. (2015). The power of the picture: How narrative film captures attention and disrupts goal pursuit. *Plos One, 10*(12). http://dx.doi.org/10.1371/journal.pone.0144493

Cohen, A. J. (2009). Music as a source of emotion in film. In P. N. Juslin (Ed.), *Handbook of music and emotion: Theory, research, applications* (pp. 879–908). Oxford: Oxford University Press.

Colman, A. M. (2015). Oxford Dictionary of Psychology (4th ed.): Oxford University Press.

Cooley, H. R. (2004). It's all about the fit: The hand, the mobile screenic device and tactile vision. *Journal of Visual Culture, 3*(2), 133-155.

Coutrot, A., Guyader, N., Ionescu, G., & Caplier, A. (2012). Influence of soundtrack on eye movements during video exploration. *Journal of Eye Movement Research, 5*(4), 1–10.

Currie, G. (1995). *Image and mind: Film, philosophy, and cognitive sciences*. Cambridge: Cambridge University Press.

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*(6), 569–576.

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

D'Aloia, A. (2012a). Cinematic empathy: Spectator involvement in the film experience. In D. Reynolds & M. Reason (Eds.), *Kinesthetic empathy in creative and cultural practices* (pp. 91–108). Bristol: Intellect.

D'Aloia, A. (2012b). Upside-down cinema: (Dis)simulation of the body in the film experience. *Cinema: Journal of Philosophy and the Moving Image, 3*, 155–182.

Doane, M. A. (2002). *The emergence of cinematic time: modernity, contingency, the archive*. Cambridge: Harvard University Press.

Dockney, J., & Tomaselli, K. G. (2009). Fit for the small(er) screen: Films, mobile TV and the new individual television experience. *Journal of African Cinema, 1*(1), 126–132.

Dockney, J., Tomaselli, K. G., & Hart, T. B. (2010). Cellphilms, mobile platforms and prodsumers: Hyper-individuality and film. In N. Hyde-Clarke (Ed.), *The citizen in communication: Re-visiting traditional, new and community media practices in South Africa* (pp. 93–115). Cape Town: Juta.

Drucker, J. (2011). Humanities approaches to interface theory. *Culture Machine, 12*, 1–20.

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice* (2nd ed.). London: Springer.

Eisenstein, S. (1975). *The film sense* (J. Leyda, Trans.). San Diego: Harcourt Brace & Company.

Eisenstein, S., & Leyda, J. (1982). *Film essays and a lecture*. Princeton: Princeton University Press.

Ellis, J. (1982). *Visible fictions: Cinema, television, video*. London: Routledge & Kegan Paul.

Engberg, M., & Bolter, J. D. (2017). Mobile cinematics. In P. Hesselberth & M. Poulaki (Eds.), *Compact cinematics: The moving image in the age of bit-sized media* (pp. 165–175). New York, London: Bloomsbury Academic.

Escera, C., Alho, K., Winkler, I., & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change in the acoustic environment. *Journal of Cognitive Neuroscience, 10*(5), 590–604.

Espinosa, J. G. (1983). For an imperfect cinema. In M. Chanan (Ed.), *Twenty-five years of the new Latin American cinema* (pp. 28–33). London: BFI Books.

Fahlenbrach, K. (2005). The emotional design of music videos: Approaches to audio-visual metaphors. *Journal of Moving Image Studies, 3*(1), 22–28.

Fahlenbrach, K. (2006). Aesthetics and audiovisual metaphors in media perception. *Formamente. Rivista Internazionale di Ricerca sul Futuro Digitale. International Research Journal on Digital Future, 1*(2), 63–77.

Fahlenbrach, K. (2007). Embodied spaces: Film spaces as leading audiovisual metaphors. In J. D. Anderson & B. Fisher Anderson (Eds.), *Narration and spectatorship in moving images* (pp. 105–124). Cambridge: Cambridge Scholar Press.

Fahlenbrach, K. (2008). Emotions in sound: Audiovisual metaphors in the sound design of narrative films. *Projections, 2*(2), 85–103.

Fanchi, M. (2005). *Spettatore*. Milan: Il Castoro.

Foucault, M. (1986). Of Other Spaces. [Des Espace Autres]. *Diacritics, 16*(1), 22–27.

Fowler, C., & Voci, P. (2011). Brief encounters: Theorizing screen attachments outside the movie theatre. *Screening The Past,* (32).

Friedberg, A. (1993). *Window shopping: Cinema and the postmodern*. Berkeley: University of California Press.

Friedberg, A. (2006). *The virtual window: from Alberti to Microsoft*. Cambridge: MIT Press.

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

Gallese, V. (2003). The manifold nature of interpersonal relations: The quest for a common mechanism. *Philosophical Transactions of the Royal Society B: Biological Sciences, 358*(1431), 517–528.

Gallese, V. (2005). Embodied simulation: From neurons to phenomenal experience. *Phenomenology and the Cognitive Sciences, 4*(1), 23–48.

Gallese, V. (2014). Bodily selves in relation: Embodied simulation as second-person perspective on intersubjectivity. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1644). http://dx.doi.org/10.1098/rstb.2013.0177

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*(12), 493–501.

Gallese, V., & Guerra, M. (2012). Embodying movies: Embodied simulation and film studies. *Cinema, 3*, 183–210.

Gallese, V., & Sinigaglia, C. (2011). What is so special about embodied simulation? *Trends in Cognitive Sciences, 15*(11), 512–519.

Gauvain, M. (2001). *The social context of cognitive development*. New York: Guilford Press.

Gaver, W. (1991). *Technology affordances.* Proceedings of SIGCHI Conference on Human Factors in Computing Systems, New Orleans, United States.

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

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology* (pp. 67–82). Hoboken: John Wiley & Sons.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gilbert, A. (2008). *What the nose knows: The science of scent in everyday life*. New York: Crown.

Goffman, E. (1986). *Frame analysis: an essay on the organization of experience*. Boston: Northeastern University Press.

Goggin, G. (2006). *Cell phone culture: Mobile technology in everyday life*. London, New York: Routledge.

Goggin, G. (2008). *Mobile phone cultures*. London: Routledge.

Goggin, G., & Hjorth, L. (2009). *Mobile technologies: From telecommunications to media*. New York: Routledge.

Goldman-Rakic, P. S. (1996). The prefrontal landscape: Implications of functional architecture for understanding human mentation and the central executive. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*(351), 1445–1453.

Gombrich, E. H. (2000). *Art and illusion: A study in the psychology of pictorial Representation* (2 ed.). London: Phaidon.

Grabowski, M. (2015). Neuromediation: An ecological model of mediated communication. In M. Grabowski (Ed.), *Neuroscience and media: New understandings and representations* (pp. 3–14). New York, London: Routledge.

Grodal, T. (1997). *Moving pictures: a new theory of film genres, feelings and cognition*. Oxford: Clarendon Press.

Grodal, T. (1999). Emotions, cognitions, and narrative patterns in film. In C. Plantinga & G. M. Smith (Eds.), *Passionate views: Film, cognition, and emotion* (pp. 127–145). Baltimore: The John Hopkins University Press.

Grodal, T. (2006). The PECMA flow: A general model of visual aesthetics. *Film Studies, 8*(1), 1–11.

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

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition and Emotion, 9*(1), 87–108.

Grusin, R. (2004). Premediation. *Criticism, 46*(1), 17–39.

Grusin, R. (2009). YouTube at the end of new media. In P. Snickars & P. Vonderau (Eds.), *The YouTube reader* (pp. 60–67). Stockholm: National Library of Sweden.

Grusin, R. (2016). DVDs, video games, and the cinema of interactions. In S. Denson & J. Leyda (Eds.), *Post-cinema: Theorizing 21st-century film* (Online ed., pp. 65–87). Falmer: Reframe Books.

Gunning, T. (1986). The cinema of attractions: Early film, its spectator and the avant-garde. *Wide Angle, 8*(3–4), 63–70.

Gunning, T. (1995). An aesthetic of astonishment: Early film and the (in)credulous spectator. In L. Williams (Ed.), *Viewing positions: Ways of seeing film* (pp. 114–133). New Brunswick: Rutgers University Press.

Hales, C. (2005). Cinematic interaction: From Kinoautomat to cause and effect. *Digital Creativity, 16*(1), 54-–64.

Hall, E. T. (1990). *The hidden dimension*. New York: Random House, Anchor Books.

Hanich, J. (2018). *The audience effect: On the collective cinema experience*. Edinburgh: Edinburgh University Press.

Hansen, M. (1995). Early cinema, late cinema: Transformations of the public sphere. In L. Williams (Ed.), *Viewing positions: Ways of seeing film* (pp. 134–154). New Brunswick: Rutgers University Press.

Harries, D. (2002). Watching the internet. In D. Harries (Ed.), *The new media book* (pp. 171–183). London: British Film Institute.

Harris, R. J., & Cook, L. (2010). How content and co-viewers elicit emotional discomfort in moviegoing experiences: Where does the discomfort come from and how is it handled? *Applied Cognitive Psychology, 25*(6), 850–861.

Hartson, H. R. (2003). Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour and Information Technology, 22*(5), 315–338.

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.

Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., & Heeger, D. J. (2008). Neurocinematics: The neuroscience of film. *Projections, 2*(1), 1–26.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science, 303*(5664), 1634–1640.

Head, H. (1920). *Studies in neurology* (Vol. 2). London: Oxford University Press.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology, 57*(2), 243–259.

Hines, P. J. (1997). Unconscious odors. *Science, 278*(5335), 79.

Hirose, N., & Nishio, A. (2001). The process of adaptation to perceiving new action capabilities. *Ecological Psychology, 13*(1), 49–69.

Hjorth, L., & Goggin, G. (2014). *The Routledge companion to mobile media*. New York: Routledge.

Hochberg, J., & Brooks, V. (2017). Film cutting and visual momentum. In J. D. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and the higher psychological functions* (pp. 293–314). New York, Oxon: Routledge.

Holmes, N. P., & Spence, C. (2004). The body schema and the multisensory representation(s) of peripersonal space. *Cognitive Processing, 5*(2), 94–105.

Horst, H. A., & Miller, D. (2006). *The cell phone: An anthropology of communication*. Oxford, New York: Berg.

Horton, D., & Wohl, R. R. (1956). Mass communication and para-social interaction: Observations on intimacy at a distance. *Psychiatry, 19*(3), 215–229.

Howe, N., & Strauss, W. (2000). *Millennials rising: The next great generation*. New York: Vintage.

Huhtamo, E. (2007). Twin-touch-test-redux: Media archaeological approach to art, interactivity and tactility. In O. Grau (Ed.), *MediaArtHistories* (pp. 71–101). Cambridge, London: MIT Press.

Huhtamo, E. (2016). The four practices? Challenges for an archaeology of the screen. In D. Chateau & J. Moure (Eds.), *Screens* (pp. 116–124). Amsterdam: Amsterdam University Press.

Hutto, D. D., & McGivern, P. (2015). How embodied is cognition? *The Philosophers' Magazine*(68), 77–83.

IJsselsteijn, W. A., de Ridder, H., Freeman, J., & Avons, S. E. (2000). *Presence: Concept, determinants and measurement.* Proceedings of Human Vision and Electronic Imaging V, San José, United States.

Ionescu, V. (2014). What do you see? The phenomenological model of image analysis: Fiedler, Husserl, Imdahl. *Image and Narrative, 15*(3), 93–110.

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

Jenkins, H. (2008). *Convergence culture: Where old and new media collide*. New York: New York University Press.

Johnson, L. C., & Landon, M. M. (1965). Eccrine sweat gland activity and racial differences in resting skin conductance. *Psychophysiology, 1*(4), 322–329.

Johnson, N. J. (2006). The mobile revolution: The making of mobile services worldwide. *Choice: Current Reviews for Academic Libraries, 43*(8), 1450–1450.

Kane, B. (2014). *Sound unseen: Acousmatic sound in theory and practice*. New York: Oxford University Press.

Kaplan, E. A. (1987). *Rocking around the clock: Music television, postmodernism, and consumer culture*. New York: Methuen.

Katz, J. E. (2003). *Machines that become us: The social context of personal communication technology*. New Brunswick: Transaction Publishers.

Katz, J. E., & Aakhus, M. A. (2002). *Perpetual contact: Mobile communication, private talk, public performance*. Cambridge, New York: Cambridge University Press.

Kiss, M., & Willemsen, S. (2017). *Impossible puzzle films: a cognitive approach to contemporary complex cinema*. Edinburgh: Edinburgh University Press.

Kittler, F. A. (1999). *Gramophone, film, typewriter* (G. Winthrop-Young & M. Wutz, Trans.). Stanford: Stanford University Press.

Koskinen, I. K. (2007). *Mobile multimedia in action*. New Brunswick: Transaction Publishers.

Kosonogov, V. (2011). Listening to action-related sentences impairs postural control. *Journal of Electromyography and Kinesiology, 21*(5), 742–745.

Kracauer, S. (1968). *Theory of film: The redemption of physical reality*. New York: Oxford University Press.

Kracauer, S. (1987). Cult of distraction: On Berlin's picture palaces. *New German Critique*(40), 91–96.

Kraft, R. N., Cantor, P., & Gottdiener, C. (1991). The Coherence of Visual Narratives. *Communication Research, 18*(5), 601–616.

Kuhl, J. (1984). Volitional aspects of achievement motivation and learned helplessness: Toward a comprehensive theory of action control. In B. A. Maher & W. B. Maher (Eds.), *Progress in experimental personality research* (pp. 99–171). New York: Academic Press.

Langkjær, B. (2000). *Den lyttende tilskuer: Perception af lyd og musik i film*. Copenhagen: Museum Tusculanums Forlag, Københavns Universitet.

Lauwereyns, J. (1998). Exogenous/endogenous control of space-based/object-based attention: Four types of visual selection? *European Journal of Cognitive Psychology, 10*(1), 41–74.

Lederman, S. J., & Klatzky, R. L. (2009). Haptic perception: A tutorial. *Attention, Perception, and Psychophysics, 71*(7), 1439–1459.

Lefebvre, S. (2016). The disappearance of the surface. In D. Chateau & J. Moure (Eds.), *Screens* (pp. 97–106). Amsterdam: Amsterdam University Press.

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin and Review, 4*(4), 501–506.

Levin, D. T., & Simons, D. J. (2000). Perceiving stability in a changing world: Combining shots and integrating views in motion pictures and the real world. *Media Psychology, 2*(4), 357–380.

Levin, D. T., & Varakin, D. A. (2004). No pause for a brief disruption: Failures of visual awareness during ongoing events. *Consciousness and Cognition, 13*(2), 363–372.

Lewis, E., & Lloyd, D. M. (2010). Embodied experience: A first-person investigation of the rubber hand illusion. *Phenomenology and the Cognitive Sciences, 9*(3), 317–339.

Ling, R. S. (2004). *The mobile connection: The cell phone's impact on society*. San Francisco: Morgan Kaufmann.

Ling, R. S. (2012). *Taken for grantedness: The embedding of mobile communication into society*. Cambridge: MIT Press.

Loertscher, M., Weibel, D., Spiegel, S., Flueckiger, B., Mennel, P., Mast, F., & Iseli, C. (2016). As film goes byte: The change from analog to digital film perception. *Psychology of Aesthetics Creativity and the Arts, 10*(4), 458–471.

Lombard, M. (1995). Direct responses to people on the screen: Television and personal space. *Communication Research, 22*(3), 288–324.

Lombard, M., & Ditton, T. (1997). At the heart of It all: The concept of presence. *Journal of Computer-Mediated Communication, 3*(2). http://dx.doi.org/10.1111/j.1083-6101.1997.tb00072.x

Lombard, M., & Ditton, T. B. (2000). *Measuring presence: A literature-based approach to the development of a standardized paper-and-pencil instrument*. Paper presented at Presence 2000: The Third International Workshop on Presence Delft, The Netherlands.

MacEntee, K., Burkholder, C., & Schwab-Cartas, J. (Eds.). (2016). *What's a cellphilm?* Rotterdam: Sense Publishers.

Manovich, L. (2001). *The language of new media*. Cambridge: MIT Press.

Manovich, L. (2007). Information as an aesthetic event. http://manovich.net/content/04-projects/056-information-as-an-aesthetic-event/53_article_2007.pdf

Maravita, A., & Iriki, A. (2004). Tools for the body (schema). *Trends in Cognitive Sciences, 8*(2), 79–86.

Marks, L. U. (2000). *The skin of the film: Intercultural cinema, embodiment, and the senses*. Durham: Duke University Press.

Marks, L. U. (2002). *Touch: Sensuous theory and multisensory media*. Minneapolis: University of Minnesota Press.

Mateas, M. (2001). A preliminary poetics for interactive drama and games. *Digital Creativity, 12*(3), 140–152.

May, J. G., Kennedy, R. S., Williams, M. C., Dunlap, W. P., & Brannan, J. R. (1990). Eye movement indices of mental workload. *Acta Psychologica, 75*(1), 75–89.

McAuley, G. (1999). *Space in performance: Making meaning in the theatre*. Ann Arbor: University of Michigan Press.

McGowan, T. (2015). Atemporality amid Lumière temporality. *Empedocles: European Journal for the Philosophy of Communication, 5*(1–2), 59–64.

McGrenere, J., & Ho, W. (May, 2000). *Affordances: Clarifying and evolving a concept.* Proceedings of Graphics interface, Montréal, Canada.

Merleau-Ponty, M. (1962). *Phenomenology of perception*. London: Routledge.

Merleau-Ponty, M. (1968). *The visible and the invisible* (A. Lingis, Trans.). Evanston: Northwestern University Press

Messaris, P. (1994). *Visual "literacy": Image, mind, and reality*. Boulder: Westview Press.

Mestre, D. R. (2005). Immersion and presence. http://www.ism.univmed.fr/mestre/projects/virtual%20reality/Pres_2005.pdf

Meyrowitz, J. (1985). *No sense of place: The impact of electronic media on social behavior*. New York: Oxford University Press.

Michotte van den Berck, A. (1953). La participation émotionelle du spectateur à l'action représentée à l'écran. *Revue internationale de filmologie, 4*(13), 87–96.

Minsky, M. (1980). Telepresence. *OMNI, 2*(9), 44–52.

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.

Mitry, J. (1997). *The aesthetics and psychology of the cinema*. Bloomington: Indiana University Press.

Mitry, J. (2000). *Semiotics and the analysis of film* (C. T. King, Trans.). Bloomington: Indiana University Press.

Mobbs, D., Weiskopf, N., Lau, H. C., Featherstone, E., Dolan, R. J., & Frith, C. D. (2006). The Kuleshov effect: The influence of contextual framing on emotional attributions. *Social Cognitive and Affective Neuroscience, 1*(2), 95–106.

Mullarkey, J. (2009). *Refractions of reality: Philosophy and the moving image*. Hampshire: Palgrave Macmillan.

Murch, W. (2001). *In the blink of an eye: A perspective on film editing* (2. ed.). Los Angeles: Silman-James Press.

Münsterberg, H. (2014). *The photoplay: A psychological study*. New York: Appleton/Createspace.

Nannicelli, T., & Taberham, P. (2014). *Cognitive media theory*. New York, Oxon: Routledge.

Nielsen, P., & Fjuk, A. (2010). The reality beyond the hype: Mobile internet is primarily an extension of PC-based internet. *The Information Society, 26*(5), 375–382.

Nissen, M. J. (1977). Stimulus intensity and information processing. *Attention, Perception, and Psychophysics, 22*(4), 338–352.

Noë, A. (2004). *Action in perception*. Cambridge: MIT Press.

Norman, D. A. (1999). Affordance, conventions, and design. *Interactions, 6*(3), 38–43.

Norman, D. A. (2013). *The design of everyday things*. New York: Basic Books.

Noton, D., & Stark, L. (1971a). Scanpaths in eye movements during pattern perception. *Science, 171*(968), 308–311.

Noton, D., & Stark, L. (1971b). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research, 11*(9), 929–932.

Oakley, I., Brewster, S., & Gray, P. (2001). *Solving multi-target haptic problems in menu interaction*. Paper presented at the Conference on Human Factors in Computing Systems, Seattle, USA.

Odin, R. (1995). For a semio-pragmatics of film. In W. Buckland (Ed.), *The film spectator: From sign to mind* (pp. 213–226). Amsterdam: Amsterdam University Press.

Odin, R. (2012). Spectator, film and the mobile phone. In I. Christie (Ed.), *Audiences* (pp. 155–169). Amsterdam: Amsterdam University Press.

Odin, R., & Casetti, F. (1990). De la paléo- à la néo-télévision. *Communications, 51*, 9–26.

Ogden, C. K., & Richards, I. A. (1923). *The meaning of meaning: A study of the influence of language upon thought and of the science of symbolism* (1st ed.). New York: Harcourt, Brace and World.

Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing, 16*(1), 105–114.

Parkhurst, D. J., & Niebur, E. (2003). Scene content selected by active vision. *Spatial Vision, 16*(2), 125–154.

Pham, P., & Wang, J. (2017). *Understanding emotional responses to mobile video advertisements via physiological signal sensing and facial expression analysis*. Paper presented at the 22nd ACM Conference on Intelligent User Interfaces, Limassol, Cyprus.

Postman, N. (1970). The reformed English curriculum. In A. C. Eurich (Ed.), *High school 1980: The shape of the future in American secondary education* (pp. 160–168). New York: Pitman.

Postman, N. (1979). *Teaching as a conserving activity*. New York: Delacorte Press.

Potter, S. (2014). *Naked cinema: Working with actors*. London: Faber & Faber.

Prince, S. (1996). True Lies: Perceptual realism, digital images, and film theory. *Film Quarterly, 49*(3), 27–37.

Prothero, J. D., Parker, D. E., Furness, T. A., & Wells, M. J. (1995). *Towards a robust, quantitative measure for presence.* Proceedings of Proceedings of the Conference on Experimental Analysis and Measurement of Situation Awareness.

Qin, H., Rau, P.-L. P., & Salvendy, G. (2009). Measuring player immersion in the computer game narrative. *International Journal of Human–Computer Interaction, 25*(2), 107–133.

Rambusch, J., & Susi, T. (2008). The challenge of managing affordances in computer game play. *Human IT, 9*(3), 83–109.

Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., . . . Thompson, R. F. (2009). Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory, 92*(2), 135–138.

Ravaja, N. (2004). Contributions of psychophysiology to media research: Review and recommendations. *Media Psychology, 6*(2), 193–235.

Rescorla, M. (2017). The computational theory of mind (Revised). In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Stanford: Metaphysics Research Lab, Stanford University.

Richardson, I. (2007). Pocket technospaces: The bodily incorporation of mobile media. *Continuum: Journal of Media and Cultural Studies, 21*(2), 205–215.

Rutherford, M. D., & Kuhlmeier, V. A. (Eds.). (2013). *Social perception: Detection and interpretation of animacy, agency, and intention*. Cambridge: MIT Press.

Schwan, S., & Ildirar, S. (2010). Watching film for the first time: How adult viewers interpret perceptual discontinuities in film. *Psychological Science, 21*(7), 970–976.

See-To, E. W. K., Papagiannidis, S., & Cho, V. (2012). User experience on mobile video appreciation: How to engross users and to enhance their enjoyment in watching mobile video clips. *Technological Forecasting and Social Change, 79*(8), 1484–1494.

Shaviro, S. (1993). *The cinematic body*. Minneapolis: University of Minnesota Press.

Shaviro, S. (2013). Accelerationist aesthetics: Necessary inefficiency in times of real subsumption. *e-flux journal, 46*. http://www.e-flux.com/journal/46/60070/accelerationist-aesthetics-necessary-inefficiency-in-times-of-real-subsumption/

Shimamura, A. P. (2000). The role of the prefrontal cortex in dynamic filtering. *Psychobiology, 28*(2), 207–218.

Shimamura, A. P. (2013). *Psychocinematics: Exploring cognition at the movies*. New York: Oxford University Press.

Shimamura, A. P., Cohn-Sheehy, B. I., Pogue, B. L., & Shimamura, T. A. (2015). How attention is driven by film edits: A multimodal experience. *Psychology of Aesthetics, Creativity, and the Arts, 9*(4), 417–422.

Simion, F., Bardi, L., Mascalzoni, E., & Regolin, L. (2013). From motion cues to social perception: Innate predispositions. In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social perception: Detection and interpretation of animacy, agency, and intention* (pp. 37–60). Cambridge, Massachusetts: MIT Press.

Simons, D. J. (2000). Attentional capture and inattentional blindness. *Trends in Cognitive Sciences, 4*(4), 147–155.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28*(9), 1059–1074.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin and Review, 5*(4), 644–649.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 6*(6), 603–616.

Slater, M. D., & Rouner, D. (2002). Entertainment education and elaboration likelihood: Understanding the processing of narrative persuasion. *Communication Theory, 12*(2), 173–191.

Smalley, D. (2007). Space-form and the acousmatic image. *Organized Sound, 12*(1), 35–58.

Smith, M. (2017). *Film, art, and the third culture: A naturalized aesthetics of film*. Oxford: Oxford University Press.

Smith, T. J. (2006). *An attentional theory of continuity editing.* (Doctoral dissertation), University of Edinburgh, Edinburgh.

Smith, T. J. (2012). The attentional theory of cinematic continuity. *Projections, 6*(1), 1–27.

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. 165–192). New York: Oxford University Press.

Smith, T. J., & Henderson, J. (2008). Attentional synchrony in static and dynamic scenes. *Journal of Vision, 8*(6), 773. http://journalofvision.org/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.

Smith, T. J., & Mital, P. K. (2013). Attentional synchrony and the influence of viewing task on gaze behavior in static and dynamic scenes. *Journal of Vision, 13*(8), 1–24.

Snickars, P., & Vonderau, P. (Eds.). (2012). *Moving data: The iPhone and the future of media*. New York: Columbia University Press.

Sobchack, V. (1992). *The address of the eye: a phenomenology of film experience*. Princeton: Princeton University Press.

Sobchack, V. (2004). *Carnal thoughts: Embodiment and moving image culture*. Berkeley: University of California Press.

Sobchack, V. (2014). Comprehending Screens: A Meditation *in Medias Res*. *Rivista di estetica,* (55), 87-101. http://estetica.revues.org/959?lang=en#quotation

Sobchack, V. (2016). The scene of the screen: Envisioning photographic, cinematic, and electronic "presence". In S. Denson & J. Leyda (Eds.), *Post-cinema: Theorizing 21st-century film* (pp. 88–128). Falmer: Reframe Books.

Spiegel, A. (1976). *Fiction and the camera eye: Visual consciousness in film and the modern novel*. Charlottesville: University Press of Virginia.

Steinbock, D. (2005). *Mobile revolution*. London: Kogan Page.

Steyerl, H. (2009). In defense of the poor image. *e-flux journal, 10*. http://www.e-flux.com/journal/10/61362/in-defense-of-the-poor-image/

Stiegler, B. (2009). The Carnival of the New Screen: From Hegemony to Isonomy. In P. Snickars & P. Vonderau (Eds.), *The YouTube Reader* (pp. 40-59). Stockholm: National Library of Sweden.

Strauss, B. (2002). Social facilitation in motor tasks: A review of research and theory. *Psychology of Sport and Exercise, 3*(3), 237–256.

Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge: Cambridge University Press.

Sundström, P. (2005). *Exploring the affective loop.* (Licentiate thesis), Stockholm University, Stockholm.

Tan, D. S. (2004). *Exploiting the cognitive and social benefits of physically large displays.* (Doctoral dissertation), Carnegie Mellon University, Pittsburgh.

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

Theeuwes, J. (1994). Endogenous and exogenous control of visual selection. *Perception, 23*(4), 429–440.

Toddi, E. (2016). Rectangle-film [25x19]. In D. Chateau & J. Moure (Eds.), *Screens* (pp. 25–28). Amsterdam: Amsterdam University Press.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*(4), 766–786.

Treisman, A. M. (1964). Verbal cues, language, and meaning in selective attention. *The American Journal of Psychology, 77*(2), 206–219.

Troscianko, T., Meese, T. S., & Hinde, S. (2012). Perception while watching movies: Effects of physical screen size and scene type. *I-Perception, 3*(7), 414–425.

Tseng, P.-H., Carmi, R., Cameron, I., Munoz, D., & Itti, L. (2009). Quantifying center bias of observers in free viewing of dynamic natural scenes. *Journal of Vision, 9*(7), 1–16.

van Deursen, A. J. A. M., Helsper, E. J., & Eynon, R. (2015). Development and validation of the Internet Skills Scale (ISS). *Information, Communication & Society, 19*(6), 804–823.

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.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge: MIT Press.

Vera, A. H., & Simon, H. A. (1993). Situated action: A symbolic interpretation. *Cognitive Science, 17*(1), 7–48.

Verhoeff, N. (2009). Theoretical consoles: Concepts for gadget analysis. *Journal of Visual Culture, 8*(3), 279–298.

Verhoeff, N. (2012). *Mobile screens: The visual regime of navigation*. Amsterdam: Amsterdam University Press.

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

Vittadini, N., Siibak, A., Carpentier Reifová, I., & Bilandzic, H. (2014). Generations and media: The social construction of generational identity and differences. In N. Carpentier, K. C. Schrøder, & L. Hallett (Eds.), *Audience transformations: Shifting audience positions in late modernity* (pp. 65–81). New York, Abingdon: Routledge.

Voci, P. (2010). *China on video: Smaller-screen realities*. London and New York: Routledge.

Walden, K. L. (2017). Archaeology of mobile film: Blink, bluevend, and the pocket shorts. In P. Hesselberth & M. Poulaki (Eds.), *Compact cinematics: The moving image in the age of bit-sized media* (pp. 133–142). New York, London: Bloomsbury Academic.

Wallengren, A.-K., & Strukelj, A. (2015). Film music and visual attention: A pilot experiment using eye-tracking. *Music and the Moving Image, 8*(2), 69–80.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review, 9*(4), 625–636.

Wilson, R. A., & Foglia, L. (2017). Embodied cognition. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Stanford: Metaphysics Research Lab, Stanford University.

Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., . . . Jäncke, P. (2007). A process model of the formation of spatial presence experiences. *Media Psychology, 9*(3), 493–525.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*(3), 225–240.

Young, R. M., & Cardona-Rivera, R. E. (2011). *Approaching a player model of game story comprehension through affordance in interactive narrative*. Paper presented at

Intelligent Narrative Technologies IV., 2011 AIIDE Workshop, Palo Alto, United States.

Zacks, J. M. (2015). *Flicker: Your brain on movies*. New York: Oxford University Press.

Zacks, J. M. (April, 2010). How we organize our experience into events. *Psychological Science Agenda*. Retrieved from https://www.apa.org/science/about/psa/2010/04/sci-brief

Zacks, J. M., & Magliano, J. P. (2011). Film, narrative, and cognitive neuroscience. In F. Bacci & D. Melcher (Eds.), *Art and the senses* (pp. 435–454). Oxford: Oxford University Press.

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. https://doi.org/10.3389/fnhum.2010.00168

Žuravljev, B. V. (2006). Useful adaptive effect as a physiological process in functional systems of the organism. *Zbornik Matice Srpske za Prirodne Nauke*(110), 5–14.

Films, Videos, and Artworks

Alvey Jr., G. H. (Director), H. Thomas (Producer) (1956). *The Door In The Wall* [Motion picture]. United Kingdom: British Film Institute.

Baker, S. (Director), S. Baker, K. Cox, M. Cox, D. Dean, & S.-C. Tsou (Producers) (2015). *Tangerine* [Motion picture]. United States: Duplass Brothers Productions, Magnolia Pictures.

Bass, J. (Director), D. Horler & S. Markovitz (Producers) (2017). *High Fantasy* [Motion picture]. South Africa, Luxembourg: Big World Cinema.

Blinkoff, S. (Director), T. Farnsworth-Smith (Producer) (2015). *The Numberlys* [Motion picture]. United States: Amazon Studios.

Brookner, A. (Director), P. Vaccaro (Producer) (2011). *The Silver Goat* [Motion picture]. United Kingdom, Argentina: Pinball London.

Burton, T. (Director), J. Roth, J. Todd, S. Todd, & R. D. Zanuck (Producers) (2010). *Alice in Wonderland* [Motion picture]. United States: Disney.

Činčera, R. (Director), L. Kalas (Producer) (1967). *One Man and His World (One Man and His House)* [Motion picture]. Czechoslovakia.

Cronenberg, D. (Director), D. Cronenberg, A. Hamori, & R. Lantos (Producers) (1999). *eXistenZ* [Motion picture]. Canada: Dimension Films.

Cuarón, A. (Director), N. Celis, A. Cuarón, & G. Rodriguez (Producers) (2018). *Roma* [Motion picture]. Mexico, United States: Netflix.

Egoyan, A. (Director), S. Brauer (Producer) (2007). *Artaud Double Bill, Chacun son cinéma: Une déclaration d'amour au grand écran* [Motion picture]. France: Cannes Film Festival, Elzévir Films.

Fleming, V. (Director), M. LeRoy (Producer) (1939). *The Wizard of Oz* [Motion picture]. United States: Warner Bros.

Hitchcock, A. (Director), J. Harrison (Producer) (1961). *Alfred Hitchcock Presents: Bang! You're Dead* [TV series]. United States: Alfred J. Hitchcock Productions.

Lumière, A., & Lumière, L. (Directors), A. Lumière & L. Lumière (Producers) (1895). *The Arrival of a Train at La Ciotat Station* [Motion picture]. France: Lumière.

Lynch, D. (Director), D. Lynch & M. Sweeney (Producers) (2007). *David Lynch on iPhone: video material for Inland Empire* [DVD]. United States: Absurda.

Lyne, C. (2017). Frames and containers. *[in]Transition: A MediaCommons; Cinema Journal Project.* [Video essay] Retrieved from http://mediacommons.org/intransition/2017/05/31/frames-and-containers-0

Maigret, N., Renaud, C., Roszkowska, M., & Qu, Y. (2015–2018). Shanzhai Archeology.

Mobile Motion Film Festival Trailer—Shot entirely on smartphones! (April, 2017). *YouTube.* [Video] Retrieved from https://www.youtube.com/watch?v=z-cqItXhBnk&ab_channel=MobileMotionFilmFestival

. *Nashville*, C. Khouri (Producer) (2012–2018) [TV series]. United States: ABC Studios.

Park, C.-K., & Park, C.-W. (Directors), C.-K. Park, C.-W. Park, & W. Jeong (Producers) (2011). *Paranmanjang – Night Fishing* [Motion picture]. South Korea: Moho Film.

Paul, R. W. (Director), (1901). *The Countryman and the Cinematograph (The Countryman's First Sight of the Animated Pictures)* [Motion picture]. United Kingdom: Paul's Animatograph Works.

Porter, E. S. (Director), (1902). *Uncle Josh at the Moving Picture Show* [Motion picture]. United States: Edison Manufacturing Company.

Potter, S. (Director), A. Fierberg & C. Sheppard (Producers) (2009). *Rage* [Motion picture]. United States, United Kingdom: Adventure Pictures.

Press Conference on Son of Saul. (March, 2016). *Index*. [Video] Retrieved from https://index.hu/video/2016/03/04/saul_fia_sajtotajekoztato/#

Rodriguez, R., & Moraes, T. (Directors), M. Ewing (Producer) (2013). *Two Scoops* [Motion picture]. United States: Troublemaker Studios.

Scott, B. (Director), (2017). *Sleep Has Her House* [Motion picture]. United Kingdom: Ether Films.

. *Sherlock*, M. Gatiss & S. Moffat (Producers) (2010–2017) [TV series]. United Kingdom, United States: BBC.

Spielberg, S. (Director), G. Eckstein (Producer) (1971). *Duel* [Motion picture]. United States: Universal Television.

Steve Jobs Keynote Macworld 2007 [video]. (December, 2011). *YouTube*. [Video] Retrieved from https://www.youtube.com/watch?v=P-a_R6ewrmM&ab_channel=EvgenyZ

Tornatore, G. (Director), F. Cristaldi & G. Romagnoli (Producers) (1988). *Cinema Paradiso* [Motion picture]. Italy, France: Cristaldifilm.

Winding Refn, N. (Director), M. Litvak, J. Palermo, M. Platt, & G. Pritzker (Producers) (2011). *Drive* [Motion picture]. United States: FilmDistrict.

Zemeckis, R. (Director), J. Rapke, S. Starkey, & R. Zemeckis (Producers) (2015). *The Walk* [Motion picture]. United States: Sony Pictures Entertainment.

Web-Based Sources and Magazine Articles

Allen, I. (2000). Screen size: The impact on picture & sound. *Dolby Laboratories, San Francisco*. Retrieved from https://www.dolby.com/us/en/technologies/screen-size-the-impact-on-picture-and-sound-ioan-allen.pdf

Bordwell, D. (August, 2009). Paolo Gioli's vertical cinema [Blog entry]. *Observations on Film Art.* Retrieved from http://www.davidbordwell.net/essays/gioli.php

Byford, S. (March, 2017). Sorry, David Lynch: Watching movies on phones is pretty good now. *The Verge*. Retrieved from https://www.theverge.com/2017/3/30/15120510/movies-on-phones-david-lynch-lg-samsung

Chen, B. X. (February, 2010). Steve Jobs' 6 sneakiest statements. *Wired*. Retrieved from https://www.wired.com/2010/02/steve-jobs/

Donadio, R. (May, 2017). Netflix defends strategy at Cannes: 'The culture is changing'. *The New York Times*. Retrieved from https://www.nytimes.com/2017/05/19/movies/cannes-okja-netflix.html?_r=0

Google. (2019a). Consumer Barometer: Mobile Multimedia Usage among people born between 1984–1993. *Google Consumer Barometer.* Retrieved from https://www.consumerbarometer.com/en/

Google. (2019b). Consumer Barometer: Video consumtion on smartphones. *Google Consumer Barometer.* Retrieved from https://www.consumerbarometer.com/en/

Huhtamo, E. (September, 2017). Erkki Huhtamo on tactile media… [Blog entry]. *Haptic Media Studies.* Retrieved from https://hapticmediastudies.wordpress.com/2017/09/15/erkki-huhtamo-on-tactile-media/

International Telecommunications Union. (2018). Global ICT developments, 2001–2018. Retrieved from https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx

Jiang, J. (May, 2018). Millennials stand out for their technology use, but older generations also embrace digital life. *Pew Research Center*. Retrieved from http://www.pewresearch.org/fact-tank/2018/05/02/millennials-stand-out-for-their-technology-use-but-older-generations-also-embrace-digital-life/

Naimark, M. (May, 1998). Interval trip report: World's first interactive filmmaker, Prague. Retrieved from http://www.naimark.net/writing/trips/praguetrip.html

Nielsen, J. (May, 2010). iPad usability: First findings from user testing. *Nielsen Norman Group*. Retrieved from https://www.nngroup.com/articles/ipad-usability-first-findings/

Nielsen, J. (May, 2011). iPad usability: Year one. *Nielsen Norman Group*. Retrieved from https://www.nngroup.com/articles/ipad-usability-year-one/

Norman, D. A. (November, 2008). Affordances and design [Blog entry]. *Don Norman: Designing for People.* Retrieved from https://jnd.org/affordances_and_design/

Ooyala. (June, 2018). Video consumption on mobile devices stabilizes in Q1 2018 at nearly three of every five videos watched. Retrieved from https://www.ooyala.com/resources/news/press-releases/video-consumption-mobile-devices-stabilizes-q1-2018-nearly-three-every

Pinball. (n.d.). The Silver Goat. Retrieved from https://pinballonline.co.uk/productions/details/1610/The-Silver-Goat

Richter, F. (September, 2013). 305 billion people to use mobile phones by 2017. *Statista*. Retrieved from https://www.statista.com/chart/1517/worldwide-mobile-phone-users/

ScreeningRoom. (n.d.). Retrieved from https://www.screeningroom.org/about

Silent Cinema Hire. (n.d.). Retrieved from https://www.silentnoizeevents.com/silent-cinema-hire/

Silent Cinema–Silent Disco Direct. (n.d.). Retrieved from http://www.silentdiscodirect.co.uk/silent-cinema/

Silent Summer Screenings. (n.d.). Retrieved from https://www.silentsummerscreenings.com/

Smith, T. J. (December, 2014). Laptop vs. IMAX: An eyetracking experiment [Blog entry]. *Continuity Boy.* Retrieved from http://continuityboy.blogspot.co.uk/2014/12/laptop-vs-imax-eyetracking-experiment.html

Storaro, V. (1998). *Univision 2:1*. Retrieved from http://www.cinematography.net/Files/univision.pdf

Tagholm, R. (June, 2012). Pinball London produces first feature film to debut on iPad,. *Publishing Perspectives*. Retrieved from https://publishingperspectives.com/2012/06/pinball-london-produces-first-feature-film-to-debut-on-ipad/

Talwalkar, P. (November, 2011). What's the best place to sit in a movie theater? [Blog entry]. *Mind Your Decisions.* Retrieved from https://mindyourdecisions.com/blog/2011/11/04/whats-the-best-seat-in-a-movie-theater/

Third City. (n.d.). Our quirky premiere launched the UK's first iPad movie, making film history. Retrieved from http://thirdcity.co.uk/silver-goat/

Thompson, C. (September, 2017). Phones are changing how people shoot and watch video. *Wired*. Retrieved from https://www.wired.com/story/thompson-smartphone-video?mbid=social_fb

Vertical Cinema. (n.d.). Retrieved from http://www.verticalcinema.org/