



Enacting Moving Images

Film Theory and Experimental Science within a *New Cognitive Media Theory*

Joerg Fingerhut and Katrin Heimann

Abstract: This article highlights ways to relate psychology, neuroscience, and film theory that are underrepresented in the current debate and that could contribute to a *new cognitive media theory*. First, we outline how neuroscientific approaches to moving images could be embedded in the embodied, enactive cognition framework and recent predictive processing theories of the brain. Within this framework, we understand filmic engagement as a specific way of worldmaking, which is co-constituted by formal elements such as framing, camerawork, and editing. Second, we address experimental progress. Here we weigh the promises and perils of neuroscientific studies by discussing the motor neuron account to camera movements as an example. Based on the limitations we identify, we advocate for a multi-method study of film experience that brings cognitive science into dialogue with philosophical accounts and qualitative in-depth explorations of subjective experience.

Keywords: 4E cognition, editing, enactivism, expansive habits, micro-phenomenology, neuromediality, new cognitive media theory, predictive processing

This article identifies what we consider to be underrepresented issues in the (neuro)science of film. On a general theory level, we believe that as central cultural phenomena cinema and moving images should occupy a more central role in philosophy of mind and cognitive science. They constitute pervasive worldmaking media that have to be reckoned with. After suggesting some directions the field could take, we discuss experimental findings in the (neuro)science of film and the critical assessment in film theory. While we partially agree with their criticism, we do not agree with a wholesale dismissal. Instead, we argue for a theoretical and experimental pluralism by pointing out the potential as well as the limitations of different approaches, and by proposing to include micro-phenomenology as a method to capture the experience of film.

Some preliminary remarks might help to navigate these topics. We find that embodied and enactive theories of cognition provide the central tools

to capture how film, pictures, and other media co-constitute domains of engagement and value (Fingerhut 2021; for recent assessments of the mind as embodied, embedded, extended, and enactive, see Gallagher 2017; Newen et al. 2018). First, experience is not something that happens to us; it is an active exploration. This has been the credo of sensorimotor enactivism that identifies experience with the executed practical knowledge regarding how movements produce sensory feedback (O'Regan and Noë, 2001). Such skill-based accounts can be extended to cultural domains and to artifactual habits of engagement that are jointly constituted by our brain-body nexus, properties of the respective media, and the socio-cultural environment. Such habits are *locationally*, *temporally*, and *transformatively expansive* (more on this in a bit). Through our exposure to pictures and moving images, we have integrated new modes of perceptual and affective access into our ways of worldmaking (Fingerhut 2020, 2021). Film constitutes a specific mode of exploration using its own set of means. It entrains us via visual framing, camera and lens movements, editing, etc., with its very scenes and stories.¹

Second, there has been increased interest in phenomenological theorizing, introspective methods, and different ways to assess subjective experience, including neurophenomenology (Petitot et al. 1999; Thompson 2007). The micro-phenomenological method that we introduce later in this article (Petimengin 2006) can be seen as another move in this direction and promises to contribute significantly to an understanding of the experience of different media, as well as playful engagements and our valuing of art (Heimann and Roepstorff 2018; Fingerhut, Grøn, and Heimann, in prep).

The two main issues we are therefore concerned with in this article are how film constitutes an important, culturally mediated way of cognition and the need to re-assess the experimental study of filmic experience. The first issue can be captured by portraying human cognizers as expert performers in different media environments. Here the central task for a cognitive science of film is to unpack how seeing a movie is different to either direct social interactions or to engagements with other cultural artifacts (consider, for example, the activity of exploring a city that is also constrained by properties of cultural objects such as the surrounding architecture, street signs, transportation infrastructure, etc.). A related aim is to capture the brain's contribution pertaining to the dimensions of interacting and valuing film, which we already addressed in this journal by introducing the concept of *neuromediality* (Fingerhut 2020).

Regarding the study of film experience, we will address available methodologies mainly by focusing on how film-specific formal elements might contribute to the conscious experience of a scene or storyline. In philosophy of mind, the relation of subjective experience and exteroceptive ac-

cess to such processes has been a central topic. Whereas standardized questionnaires and physical measurements (including neural activity) can doubtlessly contribute to our knowledge regarding conscious experiences (Pauen and Haynes 2021) they also have severe limitations. They rely on highly controlled stimuli and a rigorous repetitious design that, while necessary for measurement, also distorts the very experiences the experiment aims to measure. This might make them ill suited to capture what we value in film experience. Yet, if one does not want to uncritically rely on individual reports of subjective experiences (either of laymen or film critics and scholars), it becomes necessary to search for methodological ways to assess such experiences more directly. We will argue that micro-phenomenology constitutes such a way that hitherto has not been applied to film and that could be an integral part of a future multi-method approach in the science of film.

Toward a New Cognitive Film Theory

Since its inauguration in the 1980s, cognitive film theory has been expanding beyond classical theories of film by incorporating insights from evolutionary theory, anthropology, sociology, cognitive psychology, and increasingly cognitive neuroscience (Nannicelli and Tauberman 2014). Yet cognitivists have also been critical with respect to the reach of neuroscience and dismissive regarding the relevance of experiments for a philosophy of film. Carroll et al.'s recent (2019) companion on the philosophy of film is a striking example: neuroscientific studies are discussed only once in the book's one thousand pages, the entry on "Cognitive Theory of Moving Images" discusses neuroscientific approaches in only one paragraph referencing no studies at all.

Cognitivism provides an understanding of how our cognitive traits—as they are researched in the sciences—are exploited by filmmakers mostly with respect to character engagement and filmic narrative. This focus on character and plot engagement renders cognitive film theory, as it was originally proposed by Bordwell and Carroll (1996), susceptible to specific kinds of studies—namely, those that provide insights into emotional engagement, memory, the understanding of others, or the intentionality and causality of events. Those insights refer to schemata we first and foremost apply in the extra-filmic realm. Among the few studies on movies that have been discussed favorably in film theory are those that established inter-viewer synchrony of neural activity while watching feature films (Hasson et al. 2008) along with related research on attentional synchrony (Smith and Mital 2013). Those provided some evidence that movies have a power to entrain us into their ways of presenting a world. Such studies therefore also spoke to another central *explanandum* of cognitive film theory—

namely, the fact that cinema engages a broad variety of spectators in a largely unified way.

As said, the main body of (neuro-)psychological research referenced by cognitive film theory has been generated to explain cognitive engagements outside media domains, such as in vision science, emotion theory, and social cognition. Moreover, the cognitive capacities under study were addressed as rather generic features of the human organism. Carroll and Seeley's theory of movies as "attentional engines" (Carroll and Seeley 2013; Seeley 2020), therefore also mostly references studies pertaining to general cognitive features, such as natural perceptual-recognition capabilities that they see as universal. Even when more specific filmic means are discussed (such as the so-called "diagnostic features" that we rely on to follow a filmic narrative) we often find a recourse to the domain general nature of the cognitive capacities that we bring to the movies: "Variable framing and continuity editing can therefore be used to recapitulate the underlying pacing, cadence, and structure of ordinary vision" (Seeley 2020, 202).

We have argued before that such a focus might be too limited. Additionally, to understanding our perceptual, cognitive, and affective engagements as natural capacities, we should recognize that some cultural artifacts—and among them moving images—are so pervasive that they permeate and re-structure such engagements in systematic ways (Fingerhut 2020). Filmmakers combine multiple perspectives into scenes that therefore differ from anything we could experience in everyday perception. We have nonetheless integrated such explorations into our embodied engagements to an extent that in such cases we employ a "filmic body schema" (Fingerhut and Heimann 2017). Crucially, this includes specific constraints: moving your own body will not bring you to a better perspective with respect to a scene in a movie. Film rather requires us to lend our body (and related perceptual and cognitive capacities) to the camera movements and the editing that in turn drive the exploration of a scene.² Another way to capture this is by acknowledging that filmic explorations constitute their own embodied way of seeing.

This has implications for how to study the experience of film. Motor activity that may underlie our engagement with actions and facial expression of actors (as claims regarding the mirror neuron system would have it) should also be potentially related to camera movements and also couched in the understanding of learned body schemas of cinematic engagement. As we have additionally argued (Fingerhut and Heimann 2017; see also Fingerhut 2018, 2020), to mature further, cognitive studies of film will have to identify more precisely how film affords a specific *twofold* seeing and potentially an interaction of two motor engagements (with the configuration and the content of a scene) that jointly constitute our experience of film.³

Film is therefore not just a stimulus for a natural brain-body nexus. Film engagement is rather constituted by skills and habits that only can be captured by referencing patterns in the medium itself. Instead of following a cognitivism that explains our engagement with film via domain general capacities, we should shift the focus toward a *new cognitive film theory* that focuses on how experiential domains are generated via embodied media habits. Here, embodied and enactive cognition can contribute to an understanding of our habits being *temporally* (they relate us to our history of engagement and integrate over time), *locationally* (they are structured by neuronal and bodily processes but crucially also by the media technologies and formal features of film itself), and *transformatively* expansive (they incorporate recurring patterns in the cultural environment into cognitive routines; Fingerhut 2021).

Such theorizing borders on ontological claims (What are the constitutive elements of a mental state? Do they compromise operations of the media apparatus itself?). Yet, it mainly has central explanatory consequences for cognitive theories of media. To justify that such enactive habits should occupy a more central role in cognitive film theory it would be important (1) to demonstrate the ways film engagement differs from the everyday instantiations of perceptual, cognitive, and affective processes; and (2) to show how elements beyond the brain, such as physiological changes and bodily movements during watching (think of saccades, body sways, etc.), as well as patterns of editing or certain camera movements (associated to, say, emotional tones or levels of immersion) can be experimentally tackled and included in explanations of filmic mental processes.

Scientific studies on fundamental perceptual and experiential differences between filmic and non-filmic engagement are sparse. A striking example is the work by Ildirar and colleagues on film-inexperienced viewers that demonstrated their difficulty in perceiving edited film clips as pertaining to one and the same scene (Ildirar and Schwan 2015). More recent studies have also addressed the lack of transfer of affective elements across cuts in such viewers (Ildirar and Ewing 2018) that for others seem to be the norm and are often treated under the label of the Kuleshov effect (Calbi et al. 2019). Ildirar's studies seem to establish that filmic patterns belong to an experiential repertoire that is not present in media-inexperienced subjects. Standard formal filmic elements are not simply and naturally given in perception but have become integrated over time and via our exposure to filmic contents (e.g., by following a narrative across shots and scenes) and film form (e.g., to camera work and editing).⁴ Some studies have established expertise effects of other media usages, such as videogaming (including neuroplastic changes; see the review by Pappas and Drigas 2009), but only few studies address film competences in systematic ways.

Some Remarks on Film Stimuli in Neuroscience

The lack of neuroscientific studies on media expertise relates to a more general issue: most neuroimaging studies that use film clips operate under the assumption that filmic stimuli do not require a specific kind of competence and can function as a stand-in for reality impinging on a biological brain. Film stimuli are not chosen for their formal features and the ways those contribute to the exploration of a scene. They are rather employed as part of a *naturalistic* neuroscience that aims to include “more socially valid life-events” (Kauttonen et al. 2015, 136; see also Aliko et al. 2020). Despite this focus on realism, Kauttonen et al. also annotated stimuli with respect to formal features such as framing, camera angle and movement, etc. and found correlations between cinematic elements and the independent components (ICs) of their fMRI data. They aimed to understand neural activity with respect to a plurality of factors (and their article is a proof of methods for a data-driven approach). Yet, questions such as how a specific camera movement (e.g., a dolly shot) engages us differently with a scene, compared to, say, a similar feature such as a lens movement (e.g., a zoom) are not within the purview of such studies. They also do not access subjective ratings or reports to correlate them with their neural data.

Such studies only barely engage in filmic interpretations of their correlational data (providing thus far mostly “mere implementation” stories) and mainly interpret neural activity with reference to previously established, extra-filmic functions. Their impact for a theory of film therefore remains limited. This can, exemplarily, be seen with respect to the discussion of camera movements and their correlates: “IC33’s pronounced selectivity to camera movements, e.g. panning, suggests involvement in the processing of optic flow, which is the property of MST subarea of MT+” (Kauttonen et al. 2015, 145). The translations of such a neuroscience of film into a proper scientific study of film would require further hypothesizing and interdisciplinary interpretations, aiming at least to identify further elements such as the interactions of MST with neural activity related to motor engagement, kinesthesia, and proprioception (that normally would co-activate with optic flow because in most real-life settings we would have to move to experience optic flow) but might not in the filmic case (where we mostly sit still). Such further analysis would ideally be within the purview of the very scientist conducting the original studies, because of their command over the complex data sets that they have invested significant time to collect.

Our focus on expansive habits of film viewing additionally demands studies with conditions that explicitly contrast real-life situations with filmic scenes (see for an animal model of this: Caggiano et al. 2011), or that contrast different filmic means with each other (more on this below). Taking the mediated character of mental states seriously entails moving beyond

both cognitive film theory's focus on domain general cognitive features and neuroscience's relative lack of interest in the specifics of filmic explorations (and their use of the latter as a stand-in for reality). Such a switch in thinking is central for a *new cognitive film theory*. Such a theory, by building on an embodied and enactive cognitive science, addresses more directly how we have *integrated* moving images into our expansive, artifactual habits of understanding the world (Fingerhut 2021) and captures how certain norms of engagement and valuing have evolved over time. Once we get a grasp on those fundamental ways of enacting scenes and stories (or world models) within a medium, we can focus on more specific questions—namely, how we experientially *interact* with an individual work and how certain movies or filmic elements move or challenge us (in ways that might render the filmic world models beautiful or interesting). *Integration* and *interaction* are therefore two separate but related issues that a philosophy and cognitive science of our filmic mind needs to tackle, theoretically and experimentally.

Embodied Predictive Processing and Designer Environments

Recent predictive processing accounts (Clark 2015, 2016) that build on a unified theory of the brain under the free energy principle (Friston 2010) could be employed to explain both the *integration* of recurring filmic patterns into our cognitive routines (e.g., by capturing how a predictive brain dovetails into specific media ecologies; Fingerhut 2021), and the more fine-grained *interaction* with a specific filmic work (e.g., with respect to the reward we experience when we reinstate predictability within a filmic narrative), following what could be considered the *erotetic* (question and answer, see Seeley 2020) narrative strategies of feature films (for a comparable predictive processing account of literature see: Kukkonen 2020).

The brain is now seen as a prediction machine operating on hypotheses or predictions regarding the state of the world (*priors*) that are part of a multilayered, hierarchical system of probabilistic generative models. Within such a neural infrastructure, input is processed as deviations from predictions (as *prediction error*), with each layer mainly predicting the input from an adjacent lower layer in the hierarchy. Such a view of the brain has proven to be explanatory powerful (as is demonstrated for a wide range of psychological and neurobiological phenomena in Clark 2016). It also provides a quite different view compared to more traditional accounts of the brain that saw its main function in building percepts in a bottom-up fashion. The brain's central operating principle is now the reduction of prediction error. This can be done more locally (updating the generative models on lower levels) or more globally by switching strategies and employing different models at higher levels of the hierarchy, which becomes necessary

once a series of predictions does systematically not accommodate the respective input anymore.

Filmic engagement as opposed to a real-life situation might be just such a case. As mentioned above, movements toward the screen won't provide more information about what can be seen. This already requires a switch to a motor-perceptual-cognitive regime that differs from a real-life situation. Clark's embodied predictive processing highlights both the action orientation of the neuronal engagement machine, on the one hand (he also discusses proprioceptive feedback, interoception, etc., in their roles to sustain environmental engagement), and the role of context for conscious perceptual experience, on the other. All this, according to Clark, can work nicely towards explanations of how we operate in so-called "designer environments" that address us in specific ways and therefore enable novel cognitive activities (Clark 2016). Moving images and media ecologies are just such designer environments that are geared to engage us in specific cognitive moves. The point we want to make here is that the generative models we employ in our multilayered, rich, and flexible inner economy have to be described in their dovetailing with the explorations film itself provides (think again of the transition of images via camerawork and editing). The visual-perceptual vocabulary of film is designed, so to say, for affective, aesthetic, or generally cognitive engagement in comparable ways to how language "is 'designed' for communication, and various forms of self-stimulation" (Clark 2015, 22).⁵

While the general theory has accumulated significant empirical support (from predictive coding in vision science to general modeling approaches based on Friston's work), there are to date no accounts that explain film within this framework. However, there is a tangible shift in the field toward cognitive and cultural niches in free energy paradigms and initial proofs of concept such as simulation studies that explore the effect of material culture on culturally mediated attention styles (Constant et al. 2021).

There are already predictive processing accounts with respect to why we value the arts in other domains besides film, such as for visual artworks (focusing on the gain in predictability with respect to ambiguous stimuli; Van De Cruys and Wagemans 2011) and literature (focusing on how literary works structure our expectations and engagement; Kukkonen 2020). Our current article focuses on film form (framing, camera work, and editing) and the way it engages us. In film, self-initiated exteroceptive engagement is replaced by attention that is manipulated and driven by the medium itself. The lack of self-initiation might also free cognitive resources for interoceptive emotional explorations we otherwise would not engage in (Fingerhut [2021] and Kukkonen [2020] discuss such tradeoffs between extero- and interoceptive engagement).

These interpretations are highly speculative at this moment but might demonstrate how neuroscientific theorizing, beyond single case studies, could spur new ideas. We believe that a general reflection regarding the relationship between film and neuroscience should also include predictive accounts: they could stimulate debates in film theory and potentially propel our understanding of filmic engagement within a larger cognitive media theory.

Embodied Simulation and Film Engagement

The discussion so far has also some bearing on experimental designs, especially those isolating different filmic means such as framing, camera work, or editing. Here, studies could provide a window into differential engagement that such means afford to film viewers. Examples of this are studies that explore the effect of the positioning of a close-up or affective content within a film sequence and its effect on affective and cognitive processing of a narrative (Bálint and Rooney 2019; Calbi et al. 2019). Another example is studies on motor processing of formal elements of film such as editing (Heimann et al. 2017) and camera and lens movements (e.g., Heimann et al. 2014, 2019).

Strangely enough, both proponents of more classical film theory and cognitivists alike (see Turvey 2020; Bordwell 2020a, 2020b) have been largely critical of such neuroscientific attempts in general and their extensions into an embodied simulation theory of filmic empathy in particular (see for the latter: Gallese and Guerra 2020). This, again, is a testimony to the rather limited role cognitivists ascribed to neuroscience and its impact on film theory. Carroll, for example, rejects the idea that “film theory is a science, or that it can be or should be transformed into one, though [he does] think that there may be certain questions of film theory—perhaps concerning perception—that may be pursued scientifically” (Bordwell and Carroll 1996, 59).

Yet, embodied simulation theorists such as Gallese and Guerra (2020) would not content themselves with contributing to only perceptual issues; they aim for a neuroscience of film proper. The latter is also what aforementioned cognitivists and other critics such as Malcom Turvey (2020) object to. In the following, we want to focus on two elements of his critique: (a) the lack of contextualization of filmic elements; and (b) the threat of interpretative bloat (i.e., that very little experimental evidence is used to explain too much). On both issues we tend to agree; yet we do not agree on the consequences drawn and the separation of film theory and (neuro)scientific experiments that Turvey advocates. We will encourage *more* engagement between the two fields and we suggest experimental studies beyond the neurosciences to counter some of the limitations Turvey might worry about.

Let's focus for the moment on the example of camera work. Here, anthropomorphic camera movement has been identified as a key stylistic mean in the psychology of film. Gibson's (1979) account of direct perception, which highly influenced psychological theorizing on film experience (Tan 2018), puts it center stage: camera movements "provide a first-order guide for the composing of film. The moving camera, not just the movement in the picture, is the reason for the empathy that grips us in the cinema." We can find even stronger claims in Gallese and Guerra (2020, 91): "the involvement of the average spectator is directly proportional to the intensity of camera movements." They additionally identify an underlying mechanism, namely mirror neuron system (MNS) activity in the premotor cortex, as evidence for the psychological mechanism of simulation of camera movements and the ensuing involvement.

This is where Turvey and others (Bordwell 2020a) take exception. First, they argue that we need a differentiation between involvement, immersion, empathy, etc., that is sometimes lacking in the embodied simulation accounts. Second, the claim that involvement is "directly proportional" to the intensity of camera movements seems to be plain wrong. As Turvey argues, Gallese and Guerra cherry-pick their examples. There are also films that engage us more intensely by not using camera movements (or editing, as we will see shortly) and there might be camera movements that rather detract from such engagement (as Gallese and Guerra [2020] acknowledge with respect to failed attempts to use enduring point-of-view shots). This is why contextualization is central (although this is not the term Turvey uses): it is not the isolated camera movement itself that mediates involvement (or even spatiotemporal immersion) but its use in the context of other formal filmic means used before and after the shot (e.g., when camera does not move for some minutes and then picks up such movement again) and with respect to whether the movement fits what is depicted in the shot (for some discussion of this see also Fingerhut forthcoming).

Gallese and Guerra base their theory on experiments that use simple scenes with specific camera or lens movements (e.g., Steadicam vs. a zoom) and established enhanced motor activity in the perception of an actress grasping something or even for an approach to an empty room when a Steadicam is used (Heimann et al. 2014, 2019). To ensure the necessary statistical power in the EEG data, those same length stimuli (that differ only in their camera movements) need to be shown multiple times. Any context is eliminated. What the results therefore establish is that, *ceteris paribus*, a Steadicam seems to engage motor areas more than a zoom or a fixed camera. That such differences become statistically salient speaks to some robustness of the phenomenon, but interestingly no statistical differences in the post-EEG subjective ratings of involvement were found (Heimann et al.

2014). It could indeed be considered interpretative bloat to claim that such studies directly support an understanding of spatiotemporal immersion (i.e., that we feel transported by a camera, as Gallese and Guerra discuss with respect to some famous scenes from movie history), and there would be more steps needed (and should be taken) to arrive at the interpretation of the cinematic examples Gallese and Guerra put forward. Yet what speaks in their favor is that they at least aim at an interpretation of such findings and take a stab at what a theoretical framework (of filmic empathy) could look like. They therefore provide something other film scholars can quibble with. Again, they might be wrong in claiming that certain elements of filmic form are in themselves sufficient for filmic immersion; yet it is a heuristic advantage of Gallese's and Guerra's work that they at least try to weave those findings into explanations of our film engagement (and much more so than the correlation studies we quoted above).

It might ironically be the latter interdisciplinary transgression between neuroscience and film studies that has set off film theorists who fear the impending "imperialism" of cognitive neuroscience over the humanities (Slugan 2020). They see progress for film *theory* exclusively in a deeper scholarly engagement with filmic traditions, categories, specific works, or formal elements in cinema (to stay with our topic of film form) and consider film neuroscience and related experiments to only contributing tangentially to the field, if at all. Our take is different: we propose to widen the experimental praxis by using a broader plurality of methods to capture filmic experience. Those methods encompass crucially the humanities and film studies, yet also neuroscience with a focus on *neuromedial* elements (i.e., differences in neural responses to media compared to everyday interactions), as well as further experimental methods. Experimental design and interpretations of findings will then encompass inter- and transdisciplinary collaborations (i.e., including several disciplines as well as filmmakers and audiences; see Fingerhut forthcoming).

Let's for a moment stay with neuroscientific methods. Since the beginning of the psychology of film in Münsterberg's 1916 *The Photoplay*, one central *explanandum* has been the specific awareness film generates, the *what it is likeness* of filmic experience (Tan 2018). Here, it seems doubtful whether a focus on isolated components (such as the motor involvement with respect to artificial camera stimuli) will provide us with a satisfying account. Film experience seems to be severely truncated in such experiments (due the countless repetitions and the artificiality of the stimuli), and one might even argue that it is almost an impossible task for the participant to rate "how much involved" one is under such conditions. Moreover, many tasks and scales remain opaque for participants that also might interpret concepts such as engagement, complexity, interestingness quite differently.

This also holds for more ecologically valid viewing conditions and studies that include continuous rating paradigms in the neuroscience of film (Hasson et al. 2008; Isik and Vessel 2019). The neural component measured in such studies might still have some bearing on an overall understanding of the processing of film; yet they are but one contribution. Substantial work is still needed to correlate such findings to actual experiences of engagement and to integrate them into a theory that aims to explain why we aesthetically value certain films over others. And we might still need studies that, as we argue in our conclusion, employ quite different methodologies.

From Introspection to Micro-Phenomenology

How else could we scientifically explore how we enact film form? In closing we want to discuss this with respect to a case study on editing (or the relative lack thereof) in documentaries. As argued above, through editing we are guided in a specific exploration of the scene depicted and the meaning brought forth is partially constituted by this formal element. Phenomenological descriptions already help us to understand better how such editing provides rhythm and order on the configurational level of a filmic experience (Pearlman 2012).⁶ In the short remainder of this article we focus on differences between long-takes and cross-cutting between scenes as assessed by a micro-phenomenological design. By this, we aim to demonstrate the contribution of a more qualitative experimental approach to an understanding of film and its potential integration into neuropsychological studies as well. We believe that this can amend and in future improve the setup and interpretation of psychological and neuroscientific findings. This research is still underway, so instead of presenting results we use it as a case study to explore alternative methodologies to study film experience.

As seen, experiential engagement with moving images is a central *explanandum* of film psychology. It has also been with respect to our filmic experiences that film scholars have explicitly expressed their doubts whether science could appropriately capture them or be informative about them. Yet, to uncritically rely on experiential self-reports (of laymen, scholars, and critics) also has severe limitations. While some reports of scholars and critics might guide us toward new experiences (once we, e.g., follow the hints of a critic toward exceptional camera work, or clever editing), it might also be the case that those experiences were not shared by a broader audience in their exposure to the cinematic work. Some more reflective, scholarly claims regarding the structure of filmic experience might additionally be *post hoc* confabulations that are in line with other theoretical commitments but do not match the original structure of audience experiences. The risk is then to end up with a “lone-wolf auto-phenomenol-

ogy" (Dennett 2003, 23) and with nothing but anecdotal evidence of film experiences pertaining to specific movies.⁷ We think that the method of micro-phenomenology (Petimengin 2006) promises to provide some steps in the right direction.

Our example is a study on *Manakamana* (Stephanie Spray and Pacho Velez, 2013).⁸ This documentary received universal acclaim when it was released. A review in the *The New Yorker* (Brody 2014) was quick to point out the challenging way it engages its spectators and how it deviates from standard filmic norms. The film shows a series of ten minute cable car rides with different passengers between a village and the Manakamana temple on a mountain top, by having a fixed camera in a (moving) cable car without any visible cuts or changes in perspective: "The filmmakers follow the narrow limits of a self-imposed rule, and their obstinacy courts cinematic disaster" (Brody 2014). Yet, so the critic claims, the film nonetheless engages its audience in intense ways.

The experiment worked with two conditions: one showing two scenes with only one cut—which is hidden, because it happens in the darkness of the cable car station—between them (*long take condition*, very close to the actual movie, with the cut hidden in the darkness of the cable car station) and a second that cuts back and forth between the two scenes (*cross-cut condition*, deviating from the original movie). The participants upon viewing engaged in a micro-phenomenological interview that aimed to map the experiences they had while watching the movie in detail and chronological order. Following the data-driven approach of the micro-phenomenological analysis, the experimenters identified several differences between the two conditions concerning the timeline and synchronic dimension of the experience. For example, the interviews on the *long take condition* indicated a much higher level of associated viewing: Far more often than in the *cross-cut condition*, viewers reported that the movie made them feel as if they were sitting by themselves in the cabin, watching the protagonists from a short distance just like when using public transport. They sometimes even struggled morally with the situation ("Is it okay to watch them like this?" Or, when the protagonists seemed upset or sad: "What could I do to help them?"). On the other hand, in the *cross-cut condition* participants were more occupied with understanding the story that seemed to be indicated by the cutting (What do the different protagonists have in common? Will they meet? What were the intentions of the filmmaker?).

Interestingly, participants also differed widely with respect to their preference for one or the other condition. Some enjoyed the intense engagement with the protagonists and the situation in the *long take condition*. Others experienced this exposure as rather unsettling and felt more entertained by figuring out the edit-related filmic story in the *cross-cut condition*.

We believe that such comments inform us about the different ways of being involved in a movie experience. Similar things would have been hard to capture with a standardized questionnaire or even physiological measures. These kinds of micro-phenomenological studies on media and the arts are quite new (see also Fingerhut, Grøn, and Heimann in prep); yet they promise to provide in-depth insights into both the phenomenology of film experience and into methodological limitations of standard (neuro)psychological experiments (related to presentation formats, such as order effects, the washing out of meaningful experiences after multiple repetitions, etc.). For one, they might unveil certain dimensions of engagement (feeling of unease because of spatiotemporal immersion; fascination of having to solve a puzzle posed by a certain editing technique) that might have been overlooked and could inform future assessments of filmic involvement in the cognitive neurosciences and the aesthetic emotions related to that.

While neuroscientific experiments (such as those revealing motor differences with respect to camera and lens movements) might show something quite robust (that even survives the hostile conditions of large repetition cycles of an EEG experiments), they can only provide a small contribution to the study of our experiential engagement with film. This might be one reason why film theorists expect advances in film theory rather to come from the humanities and not from bloated interpretations of neuroscientific or psychological studies. We believe that this might itself be an over-generalization. We agree that single studies should not be over-emphasized with respect to the generality of their effects. Yet what we nonetheless see as a promising alley for progress is a growing plurality of approaches, experiments, and interdisciplinary endeavors. It might be an ensuing comparative methodology that will contribute to a better understanding of filmic ways of worldmaking and the impact of films or filmic scenes on their viewers.

Conclusion

By spanning such a wide net of topics regarding the neuroscience of film, we aim to defend a rather capacious research program. First, we introduced an enactive understanding of film form that could constitute a basis for an interdisciplinary cognitive science of moving images. Only when we understand neural activity as embedded in a larger, embodied and culturally mediated praxis of engaging film (or as *neuromedial* activity) can we make headway toward an integrative *new cognitive film theory*. Movies are more than naturalistic stimuli for the socio-biological brain. They are part of a pervasive media ecology, and we have to capture the artifactual habits that engage them and therefore the specific ways of world-making that film affords. Second, we discussed issues regarding experimental approaches to film experience (such as “involvement”) and identified some reasons to

take issue with the way neuroscientific experiments address such experiences. Here we concluded with a call for plurality in methods and measures and introduced a new and promising methodology to study the subjective experience of film.

Acknowledgments

We thank Laura Di Summa, Vittorio Gallese, Michele Guerra, Jesse J. Prinz, and Andreas Roepstorff for discussions of the topics addressed in this paper as well as two anonymous reviewers. Work on this paper was funded by the European Union's *Horizon 2020 Research and Innovation Program* under grant agreement no. 870827: ARTIS and the *Croatian Science Foundation* under the project UIP-2020-02-1309.

Joerg Fingerhut is currently Professor for Philosophy of Mind at the University of Munich (LMU). He is also research group leader at the *Berlin School of Mind and Brain/Department of Philosophy* at Humboldt-Universität zu Berlin. He holds a PhD in philosophy and works in an empirically engaged way on cultural artifacts (such as architecture, film, visual art) and the models of the world they embody. Email: joerg.fingerhut@hu-berlin.de
<https://orcid.org/0000-0002-9021-9617>

Katrin Heimann is a postdoc at the *Interacting Minds Centre* at Aarhus University and a researcher at the Max-Planck-Institute of Empirical Aesthetics, Frankfurt. She is educated in philosophy (MA) and in cognitive neuroscience (MSc and PhD), and after years of neuroscientific praxis has specialized in experimental approaches to explore subjective experiences, such as micro-phenomenology. Email: katrinheimann@cas.au.dk
<https://orcid.org/0000-0001-6524-7408>.

Notes

¹ Film crucially also entrains us by character and story engagement, yet the focus of the present paper will be on film form and the access it affords.

² Body schema in this context refers to the systematic way in which an organism links motor engagements and proprioceptive changes to navigate its environment. We claim that humans entertain different body schemas that all contribute in *prenoetic* ways to perception and cognition (Gallagher 2005) and that our film experience supervenes upon such a schema into which patterns and regularities of the medium have become integrated.

³ This twofoldness of picture perception is not the main topic of the present paper, though it will be touched upon several times. Our focus on camerawork and editing (film form, more generally) naturally sways us toward discussions of the contribution of the configurational layer to the twofold experience. Yet the concept of twofoldness entails

that the specific kind of world models that cinema, TV, and other moving image media contain are present in a blend of configurational and recognitional layers—in other words, the two layers are only theoretically separate.

⁴ The standard case comprises the integration of the configurational (form) and recognitional (content) layers of film into one experience within an artifactual habit. Yet, there might also be filmic effects that not only contribute to driving a narrative, but might be also used to portray a single idea, or distort the viewers, etc. Any interpretation of those also has to be presented against the backdrop of a learned schema of film engagement.

⁵ It might heuristically also be valuable to describe certain media works as generative models themselves (albeit at quite different timescales than those realized at neural hierarchical processing), as probabilistic machines that produce specific sets of output that are valued (or not) by their viewers whose responses they predict. Algorithms underlying novel digital or enactive media (Kaipainen et al. 2011) implement such an idea more directly by using physiological data of the viewers to “predict” the image (or other stimuli) that will be presented next (see also Fingerhut 2021 for some discussion).

⁶ The concept of a configurational layer references back to the concept of the two-foldness of our film experience (Fingerhut 2020) as well as the related idea of a two-seriesness of film (Terrone 2018) that enters our appreciation of film.

⁷ We believe that, here, both are needed; an accompanying form of heterophenomenology (as it is employed in the cognitive sciences) as well as more in-depth second person or dialogical phenomenology (as it might characterize the interview method and analysis technique of micro-phenomenology). The authors of this article differ from one another with respect to what extent heterophenomenology is compatible to either phenomenology (for a negative assessment of this, see Zahavi 2007) or the micro-phenomenological method. One would claim rather more (JF) the other rather less (KH).

⁸ The analysis of the study is still in progress; the working title of the forthcoming paper is Heimann et al.: “The Experience of the Long Take in Ethnographic Film. An Empirical Study.”

References

- Aliko, Sarah, Jiawen Huang, Florin Gheorghiu, Stefanie Meliss, and Jeremy I. Skipper. 2020. “A Naturalistic Neuroimaging Database for Understanding the Brain Using Ecological Stimuli.” *Scientific Data* 7 (347). <https://doi.org/10.1038/s41597-020-00680-2>.
- Bálint, Katalin E., and Brendan Rooney. 2019. “Narrative Sequence Position of Close-Ups Influences Cognitive and Affective Processing and Facilitates Theory of Mind.” *Art & Perception* 7 (1): 27–51. <https://doi.org/10.1163/22134913-20191095>.
- Bordwell, David. 2020a. “Brains, Bodies, and Movies: Ways of Thinking about the Psychology of Cinema.” *Observations on Film Art*, 29 April. <http://www.davidbordwell.net/blog/2020/04/29/brains-bodies-and-movies-ways-of-thinking-about-the-psychology-of-cinema/>.
- Bordwell, David. 2020b. “Can the Science of Mirror Neurons Explain the Power of Camera Movement? A Guest Post by Malcolm Turvey.” *Observations on Film Art*, 3 May. <http://www.davidbordwell.net/blog/2020/05/03/can-the-science-of-mirror-neurons-explain-the-power-of-camera-movement-a-guest-post-by-malcolm-turvey/>.
- Bordwell, Davin, and Noël Carroll, eds. 1996. *Post-Theory: Reconstructing Film Studies*. Madison, WI: University of Wisconsin Press.
-

- Brody, Richard. 2014. "A Documentary Set Entirely Within the Capsule of a Cable Car." *The New Yorker*, 16 April. <https://www.newyorker.com/culture/richard-brody/a-documentary-set-entirely-within-the-capsule-of-a-cable-car>.
- Caggiano, Vittorio, Leonardo Fogassi, Giacomo Rizzolatti, Joern K. Pomper, Peter Thier, Martin A. Giese, and Antonino Casile. 2011. "View-Based Encoding of Actions in Mirror Neurons of Area F5 in Macaque Premotor Cortex." *Current Biology* 21 (2): 144–148. <https://doi.org/doi:10.1016/j.cub.2010.12.022>.
- Calbi, Marta, Francesca Siri, Katrin Heimann, Daniel Barratt, Vittorio Gallese, Anna Koleznikov, and Maria A. Umiltà. 2019. "How Context Influences the Interpretation of Facial Expressions: A Source Localization High-Density EEG Study on the 'Kuleshov Effect'." *Scientific Reports* 9 (1): 1–16. <https://doi.org/10.1038/s41598-018-37786-y>.
- Carroll, Noël, Laura T. Di Summa, and Shawn Loht, eds. 2019. *The Palgrave Handbook of the Philosophy of Film and Motion Pictures*. Cham, Switzerland: Palgrave Macmillan.
- Carroll, Noël, and William P. Seeley. 2013. "Cognitivism, Psychology, and Neuroscience: Movies as Attentional Engines." In *Psychocinematics: Exploring Cognition at the Movies*, ed. Arthur P. Shimamura, 53–75. New York: Oxford University Press.
- Clark, Andy. 2015. "Radical Predictive Processing." *The Southern Journal of Philosophy* 53 (1): 3–27. <https://doi.org/10.1111/sjp.12120>.
- Clark, Andy. 2016. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. New York: Oxford University Press.
- Constant, Axel, Alexander Daniel Dunsmoir Tschantz, Beren Millidge, Felipe Criado-Boado, Luis M. Martinez, Johannes Müller, and Andy Clark. "The Acquisition of Culturally Patterned Attention Styles under Active Inference." 2021. *Frontiers in Neurobotics* 15: 1–13. <https://doi.org/10.3389/fnbot.2021.72>.
- Dennett, Daniel. 2003. "Who's On First? Heterophenomenology Explained." *Journal of Consciousness Studies* 10 (9–10): 19–30.
- Fingerhut, Joerg. 2018. "Enactive Aesthetics and Neuroaesthetics." *Phenomenology and Mind* 14: 80–97. https://doi.org/10.13128/Phe_Mi-23627.
- Fingerhut, Joerg. 2020. "Twofoldness in Moving Images: The Philosophy and Neuroscience of Filmic Experience." *Projections* 14 (3): 1–20. <https://doi.org/10.3167/proj.2020.140302>.
- Fingerhut, Joerg. 2021. "Enacting Media. An Embodied Account of Enculturation Between Neuromedia and New Cognitive Media Theory." *Frontiers in Psychology* 12 (635993). <https://doi.org/10.3389/fpsyg.2021.635993>.
- Fingerhut, Joerg. Forthcoming. "The Mediated Brain: A Case Study." In *Worlding the Brain*, eds. Flora Lysen, and Stephan Besser. Leiden: Brill.
- Fingerhut, Joerg, and Katrin Heimann. 2017. "Movies and the Mind: On Our Filmic Body." In *Embodiment, Enaction, and Culture: Investigating the Constitution of the Shared World*, eds. Christoph Durt, Thomas Fuchs, and Christian Tewes, 353–377. Cambridge, MA: MIT Press.
- Fingerhut, Joerg, Karen Grøn, and Katrin Heimann. In prep. "Choosing Your Favorite Artwork. A Micro-Phenomenological Study at the Trapholt Art Museum."
- Friston, Karl. 2010. "The Free-Energy Principle: A Unified Brain Theory?" *Nature Reviews Neuroscience* 11 (2): 127–138. <https://doi.org/10.1038/nrn2787>.
- Gallagher, Shaun. 2005. *How the Body Shapes the Mind*. Oxford: Oxford University Press.
- Gallagher, Shaun. 2017. *Enactivist Interventions: Rethinking the Mind*. Oxford: Oxford University Press.
- Gallese, Vittorio, and Michele Guerra, ed. 2020. *The Empathic Screen: Cinema and Neuroscience*. Oxford: Oxford University Press.

- Gibson, James J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton, Mifflin and Company.
- Hasson, Uri, Ohad Landesman, Barbara Knappmeyer, Igancio Vallines, Nava Rubin, and David Heeger. 2008. "Neurocinematics: The Neuroscience of Film." *Projections* 2 (1): 1–26. <https://doi.org/10.3167/proj.2008.020102>.
- Heimann, Katrin, Federica Cavaletti, Joerg Fingerhut, Mikk Grins, Andreas Steenfath, Christian Suhr. In prep. "The Experience of the Long Take in Ethnographic Film. An Empirical Study."
- Heimann, Katrin, and Andreas Roepstorff. 2018. "How Playfulness Motivates—Putative Looping Effects of Autonomy and Surprise Revealed by Micro-Phenomenological Investigations." *Frontiers in Psychology* 9. <https://doi.org/10.3389/fpsyg.2018.01704>.
- Heimann, Katrin, Sebo Uithol, Marta Calbi, Maria A. Umiltà, Michele Guerra, Joerg Fingerhut, and Vittorio Gallese. 2017. "'Cuts in Action': A High Density EEG Study Investigating the Neural Correlates of Different Editing Techniques in Film". *Cognitive Science* 41 (6): 1555–1588. <https://doi.org/10.1111/cogs.12439>.
- Heimann, Katrin, Sebo Uithol, Marta Calbi, Maria A. Umiltà, Michele Guerra, Joerg Fingerhut, and Vittorio Gallese. 2019. "Embodying the Camera: An EEG Study on the Effect of Camera Movements on Film Spectators' Sensorimotor Cortex Activation." *PLOS ONE* 14 (3). <https://doi.org/10.1371/journal.pone.0211026>.
- Heimann, Katrin, Maria A. Umiltà, Michele Guerra, and Vittorio Gallese. 2014. "Moving Mirrors: A High-Density EEG Study Investigating the Effect of Camera Movements on Motor Cortex Activation During Action Observation." *Journal of Cognitive Neuroscience* 26 (9): 2087–2101. https://doi.org/10.1162/jocn_a_00602.
- Ildirar, Sermin, and Louise Ewing. 2018. "Revisiting the Kuleshov Effect with first-time viewers." *Projections*, 12 (1): 19–38. <https://doi.org/10.3167/proj.2018.120103>.
- Ildirar, Sermin, and Stephan Schwan. 2015. "First-time viewers' comprehension of films: Bridging shot transitions." *British Journal of Psychology*, 106 (1): 133–151. <https://doi.org/10.1111/bjop.12069>.
- Isik, Ayse I., and Edward A. Vessel. 2019. "Continuous Ratings of Movie Watching Reveal Idiosyncratic Dynamics of Aesthetic Enjoyment." *PLOS ONE* 14 (10). <https://doi.org/10.1371/journal.pone.0223896>.
- Kaipainen, Mauri, Niklas Ravaja, Pia Tikka, Rasmus Vuori, Roberto Pugliese, Marco Rapino, and Tapio Takala. 2011. "Enactive Systems and Enactive Media: Embodied Human-Machine Coupling Beyond Interfaces." *Leonardo* 44 (5): 433–438. doi: https://doi.org/10.1162/LEON_a_00244.
- Kauttonen, Jeanne, Yevhen Hlushchuk, and Pia Tikka. 2015. "Optimizing Methods for Linking Cinematic Features to fMRI Data." *NeuroImage*, 110: 136–148. <https://doi.org/10.1016/j.neuroimage.2015.01.063>.
- Kukkonen, Karin. 2020. "The Cognitive Work of Form". In *Probability Designs*, ed. Karin Kukkonen, 178–192. Oxford: Oxford University Press.
- Münsterberg, Hugo. 1916. *The Photoplay: A Psychological Study*. New York: D. Appleton and Company.
- Nannicelli, Ted, and Paul Taberham, eds. 2014. *Cognitive Media Theory*. New York: Routledge.
- Newen, Albert, Leon De Bruin, and Shaun Gallagher, ed. 2018. *The Oxford Handbook of 4E Cognition*. Oxford: Oxford University Press.
- O'Regan, J. Kevin, and Alva Noë. 2001. "A Sensorimotor Account of Vision and Visual Consciousness." *Behavioral and Brain Sciences* 24 (5): 939–73. <https://doi.org/10.1017/S0140525X01000115>.

-
- Pappas, Marios A., and Athanasios S. Drigas. 2019. "Computerized Training for Neuroplasticity and Cognitive Improvement". *International Journal of Engineering Pedagogy (IJEP)* 9 (4): 50–62. <https://doi.org/10.3991/ijep.v9i4.10285>.
- Pauen, Michael, and John-Dylan Haynes. 2021. "Measuring the Mental." *Consciousness and Cognition* 90: 103106. <https://doi.org/10.1016/j.concog.2021.103106>.
- Pearlman, Karen. 2012. *Cutting Rhythms: Shaping the Film Edit*. New York: Routledge.
- Petitmengin, Claire. 2006. "Describing One's Subjective Experience in the Second Person: An Interview Method for the Science of Consciousness." *Phenomenology and the Cognitive Sciences* 5 (3–4): 229–269. <https://doi.org/10.1007/s11097-006-9022-2>.
- Petitot, Jean, Francisco J. Varela, Bernard Pachoud, and Jean-Michael Roy, eds. 1999. *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*. Stanford, CA: Stanford University Press.
- Seeley, William P. 2020. *Attentional Engines: A Perceptual Theory of the Arts*. Oxford: Oxford University Press.
- Slugan, Mario. 2020. "Film Studies and the Experimental Method." *NECSUS_European Journal of Media Studies* 9 (2). <https://doi.org/10.25969/MEDIAREP/15317>.
- Smith, Tim J., and Parag K. Mital. 2013. "Attentional Synchrony and the Influence of Viewing Task on Gaze Behavior in Static and Dynamic Scenes." *Journal of Vision* 13 (8). <https://doi.org/10.1167/13.8.16>.
- Tan, Ed S. 2018. "A Psychology of the Film." *Palgrave Communications* 4 (1): 1–20. <https://doi.org/10.1057/s41599-018-0111-y>.
- Terrone, Enrico. 2018. Why to Watch a Film Twice. In *The Pleasure of Pictures*, ed. Jérôme Pelletier, and Alberto Voltolini, 225–246. New York: Routledge.
- Thompson, Evan. 2007. *Mind in Life*. Cambridge, MA: Harvard University Press.
- Turvey, Malcom. 2020. "Mirror Neurons and Film Studies: A Cautionary Tale from a Serious Pessimist." *Projections* 14 (3): 21–46. <https://doi.org/10.3167/proj.2020.140303>.
- Van de Cruys, Sander, and Johan Wagemans, J. 2011. "Putting Reward in Art: A Tentative Prediction Error Account of Visual Art." *I-Perception* 2 (9): 1035–1062. <https://doi.org/10.1068/i0466aap>.
- Zahavi, Dan. 2007. "Killing the Straw Man: Dennett and Phenomenology." *Phenomenology and the Cognitive Sciences* 6 (1–2): 21–43. <https://doi.org/10.1007/s11097-006-9038-7>.
-