Università degli Studi di Parma
Dottorato di Ricerca in Neuroscienze
Ciclo XXVII

‘How movies move us just the right way’
Exploring the role of camera movement and montage in human film perception

First steps on a joint venture of 4EA approaches to cognition and empirical Neuroscience

Coordinatore:
Chiar.mo Prof. Vittorio Gallese

Tutor:
Chiar.mo Prof. Vittorio Gallese

Dottoranda:
Katrin Heimann

TABLE OF CONTENTS

0. Outline

Film is an omnipresent medium in today's world, playing a crucial role in a massive amusement industry, in advertisement as well as in information distribution, education, sciences and the arts. What and how we perceive the world when we perceive it via film thus seems an important question to answer. Still, the principles of film perception are barely understood. This thesis presents some of the newest steps in an interdisciplinary approach to exploring human creation and perception of film. Specifically, it reports about two high density EEG studies investigating the role of camera movement and montage in live-action edited moving images while being informed by modern theoretical approaches to cognition as will be outlined in the following. Live-action edited moving images comprise any kind of film or video, consisting of shots recorded by means of a camera filming live events (in contrast to animations), with these shots being then edited together to form a perceptually continuous stream of images.[1] This thesis argues that contemporary concepts of how the mind works, known as 4EA (embedded, embodied, enactive, extended and affective) approaches to cognition, can be used to develop a notion of the human perception of life-action edited movies which can crucially illuminate the principles of movie making and movie experience. Especially, as this thesis will outline, they can help to gain an understanding of the role of camera movement and montage in procuring deep involvement of the spectators in the film's fictional world. This thesis will show how such theoretical conceptualisations can be used to derive precise hypotheses testable in an experimental setup of cognitive neuroscience, enabling us to enhance our knowledge about how movies can move us the way they do. The following paragraphs will shortly outline the main aspects of this project.

From the perspective of 4EA approaches to cognition camera movement and montages can be described as rendering a film experience into an interaction between a fictional world and an embodied inhabitant of this world, the spectator. Specifically, they suggest that camera movements and montages create visual experiences that recruit and sustain the complete embodied skills of the perceiver, that is, his habits and experiences from his daily life as an embodied being actively exploring the real world. In consequence, the illusion of a fictional environment gets created, in interaction with which humans can deeply immerse in the movie, an experience known to all of us. Moreover, it is proposed that moving images closely

---

[1] From here on the terms "film" or "movie" will refer to this specific subclass of moving images, unless stated otherwise.

[2] "Massive Open Online Courses" used by millions of learners of all nationalities, ages and educational backgrounds are nowadays revolutionising the educational field. See http://ideas.ted.com/2014/01/29/moocs-by-the-numbers-where-are-

resembling vision in real life take over a substantiating role in this relationship, while also allowing the occurrence of violations, such as innovative perspectives and moves or temporal or spatial jumps, for which movies are equally famous. Lastly, it seems to be derivable that the embodied perceptual activity of a spectator, guided by the movies, over time, can even learn new habits in perceiving the world. Such process might turn effects that were previously experienced as strong violations into only slight deviations or even known perceptual circumstances.

To follow, test and exploit such a concept of film perception, it seems to be necessary to better understand, what it means to perceive familiarity or violation of a perceptual habit in a film. The following thesis will show how the above sketched notion of film perception and recent theoretical and empirical work on related topics can be used to formulate three precise hypotheses, able to guide two neuroscientific experiments that can precisely illuminate this point. In the experiments, we assessed the spectators' brain activity and subjective experience during a movie by a high density EEG recording and a related rating task, while comparing responses to moving images that, due to the application of camera movement or montage, either more or less closely resemble vision as experienced in real life. The hypotheses that guided the analysis (as will be developed in the thesis) concerned 1) the detection of a violation of visual habits, 2) the postperceptual processing of a violation of visual habits, and 3) the effect of familiarity with and violations of perceptual habits on activation of motor cortex areas during action observation (mirror mechanism). The first study focused mainly on the last point, while looking for differences in motor cortex activations of spectators watching short action movie clips produced by applying different types of camera movements or uses (steadicam, dollycam and zoom). The second study tested all three hypotheses by investigating differences in ERP and frequency data of spectators watching short action movie clips produced by applying different types of montages (continuity editing versus reverse shot). The results are discussed within the chosen framework while underlining the new findings made and the ideas developed for future research due to the specific perspective taken.

The thesis is structured as follows:

The introduction answers the most fundamental question justifying this work, that is, why film should be a topic of neuroscientific research to begin with. It does so by pointing out the strong presence of film and the important role it plays for a wide range of sectors in our daily life. It then illustrates the explicit differences between a world presented by a movie and the world that surrounds us in real life, while stressing our common neglect of these differences, crucially enabling the strong impact of films on us. This finally allows this chapter to present filmmakers as 'specialists of human cognition' and their products as an object worth studying also from a neuroscientific perspective.

The second chapter introduces 4EA approaches to cognition by pointedly contrasting them with competing perspectives within a theory of mind and illustrating the effect of the different stances on experimental works on cognition. It then reports about the particularly difficult standing of 4EA approaches in the history of media and art research, while at the end highlighting promising outcomes of most recent experimental works that use this framework. In the following part, it presents the specificity of film in comparison to other media and describes how this special character can be conceptualized using 4EA approaches to cognition. Lastly, it formulates the hypotheses regarding brain activations during film perception that can be derived from such a reconceptualization, and that served as the basis for the neuroscientific experiments reported on in Chapter III.

The third chapter describes the two experimental studies on film conducted during my PhD, comprising detailed presentation of design, analyses, findings and interpretations.

Lastly, the fourth chapter summarizes the results made in both studies, while conclusively underlining their implications for our understanding of why movies move us the way they do. An outlook to further experiments and future projects at the end of this chapter will conclude the thesis.

I. Introduction: Film as a topic for psychological and neuroscientific research

Taking into account how much there is still left to do in neuroscientific research about human perception and cognition in real life settings, some might argue that exploring mechanisms of film perception might be a project of minor importance. This view however dismisses that film perception is everything but a side phenomenon of today's life. As a matter of fact, film is omnipresent in today's world. Recent surveys reported that average people for as much as about 1/5 of their lifetime are confronted with edited moving image material, a big part of which consists of live-action edited movies in the above defined sense (U.S. Bureau of Statistics, 2014; see also Casetti, 2015). This is easy to believe considering that these kind of films are one of the main products of a massive amusement industry with an enormous impact on economics, amounting in the United States alone to 7-8 billion cinema admissions and about 25 billion US$ income due to video/DVD sales every year. Films also play a major role in all contexts focusing on the distribution of information, ranging from public media to the steadily growing sector of MOOCs[2]. In behavioral, cognitive and neurological scienes, films often find application as stimuli that allow the controlled assessment of behavioral and physiological responses to real world stimuli. Furthermore, films represent a specific medium of artistic production, one of the most impressive capacities of mankind, fascinating scholars of various disciplines. Lastly, in all of these contexts, films may implicitly or explicitly be used and are used for purposes of decision and opinion making, which alone makes them a topic worth studying from different angles.

One might hypothesize that the popularity of live-action movie usage, illustrated above, is due to its close similarity to normal perception. Thanks to today's recording techniques, film has the capacity to present the world visually and auditorily in real time, thus, in comparison to other media, to best simulate the sensual richness, directness and vividness of the real world. This view however pays too little attention to the differences between the world presented in film and the experiences we entertain in real life. Even if leaving out animation and special effects made possible by today's digital capacities, it is easy to point out how the two most traditional narrative devices of film, that is the use of camera movements and montage, create a representation of the world extremely divergent from the way our normal daily life and

---

[2] "Massive Open Online Courses" used by millions of learners of all nationalities, ages and educational backgrounds are nowadays revolutionising the educational field. See http://ideas.ted.com/2014/01/29/moocs-by-the-numbers-where-are-we-now.

bodily being allows it us to see and experience. Shots from high above the ceilings or from low above the ground, wide angles and extreme close-ups, zoom ins and zoom outs, looks over the shoulder of the murderer or even from the eyes of the victim – the camera literally multiplies our capacities of being and moving. Moreover, montages editing shots sometimes only seconds in length but connecting scenes far apart in time and space, allow us to switch from yesterday to today or the far future, leaping from earth to the moon or jumping from buying a gun to burying a person in a split second. Furthermore, they even enable the quasi-sensory experiences of parallel, circular or reverse action, phenomena so contrary to our capacities as real living creatures that it is indeed hard to imagine how we could have imagined them without media and arts.

Remarkably, at least when referring to the tradition of narrative cinema and movies, little of these manipulations in film catch our direct attention or challenge us intellectually to a degree that throws us out of our deep involvement with the movie's plot. On the contrary, previous research has shown, that the capacity to easily follow movies is neither culture or education nor specifically age dependent[3], with neither small kids nor media inexperienced adult spectators of non-western origin showing a problem in understanding the narratives presented by edited moving images of the type described above (Schwan & Ildirar, 2010; Comuntzis, 1987, 1991[4]). Moreover, further research showed, that the almost transparent nature of the devices reported about (that is camera use and montage) is able even to resist us when we are explicitly focusing on their application. To give an example, Zacks, Speer & Reynolds (2009) showed, that even if participants were explicitly asked to press a button every time they spotted a cut in a movie, some of the cuts were missed.[5] The apparent perception-deluding power of movies demonstrated by these findings did from the very start of film production stimulate researchers to think of filmmakers as implicit specialists of perception, and of their products as objects obviously worth being studied by all sciences determined to illuminate the principles and capacities of human cognition.[6]

---

[3] Excluded are of course infants or age-related demented people, who are not able to follow even real life events in basic terms usual for a normal functioning adult.

[4] However, for an interesting study that indicates some culture-specific differences, see Worth & Adair, 1972, also critically discussed in Smith, 2005.

[5] For further details about this study see chapter III.2.1 of this thesis.

[6] Indeed, already in 1916 the psychologist Hugo Münsterberg dedicated a whole book to the topic of film perception, that is especially to the mental performances he assumed to underlie this special kind of experience. Even if film experience as it will be described in this thesis is not ultimately captured by his assumptions, his observations in total are of highly illuminating character and the book strongly recommended. See Münsterberg (1916): *The Photoplay*. Antonio Damasio stated, that those filmmakers who made film what it is and can do today must have explicitly or implicitly taken into account the functioning of the human brain (Damasio, 2008).

Interestingly, though this view has also been promoted by scholars of the neuroscientific field (see for example Dudai, 2012), neuroscientific studies explicitly dedicated to this issue are still of limited number. A crucial reason for this might be the immense richness of movies not fitting to the general principle of experimental research, which must use controlled stimuli only. To get interpretable results it is considered highly necessary to create stimuli and tasks allowing that all participants actually engage in the same cognitive processes. Thinking about the complex interactions of film plot, film style and the individual background of the perceiver, especially over the rather long duration of a movie, indeed raises the question if there can be anything shared at all between spectators - concerning the processing as well as the related subjective experience of the film. Likely, the most famous study dedicated to the exploration of this question was conducted by Hasson et al. (Hasson, Furman, Clark, Dudai & Davachi, 2008; but see also for a publication in a journal of film sciences: Hasson, Landesman, Knappmeyer, Vallines, Rubin, & Heeger, 2008). Using fMRI recording of participants watching actual commercial movies, the authors demonstrated that for some movies and during certain time windows BOLD responses and gaze behavior of participants showed a significant correlation. This finding was interpreted as indicating that movies are able to “exert considerable control over brain activity and eye movements” (Hasson et al., 2008b, p.1). Nevertheless, the authors also stressed that the level of correlation between participants’ responses significantly varied as a function of movie content, editing and directing style. Interestingly, after this, only very few other studies used neuroscientific methods to explore these observations on a smaller scale, showing only short sequences of film and focusing on one problem at a time, for instance the effects of specific filmic devices such as montage and camera movement on spectators’ brain responses (see for an exception Magliano & Zacks (2011) further explained in Chapter III of this thesis).

The studies presented in this thesis were designed to start filling this gap, while being guided by what is called the “4EA approaches to cognition”. The following chapter will introduce this notion by contrasting it against competing perspectives, as well as by illustrating how in previous decades it has influenced experimental research about cognition in general and art perception specifically. Finally it will be argued how 4EA approaches considerably affect our understanding of film perception and how this can be exploited by cognitive neuroscience exploring film perception.

II. Theoretical Background: Why introduce 4EA approaches to the study of film?

II.1 Enacted, embodied, embedded, extended and affective cognition as a new concept in the philosophy of mind and in empirical research about cognition

Cognitive Sciences historically developed in demarcation to behaviorism as the dominant paradigm in life sciences at the beginning of the 20th century. As explained by B.F. Skinner in his book *Verbal behavior*, in order to understand human performance in bodily as well as mental tasks, behaviorism promotes focusing on functional relations between stimulus and response only, without positing internal representations of an actually unapproachable mind (Skinner, 1957). In the 70s and 80s of the last century, a huge part of research separated from this prevalent perspective in order to move on to directly exploring the organization and processing mechanisms of human perception and cognition. One strong school in this new wave especially focused on computer models and language to model cognition.[7] Consequently, in these newly arising theories, perception has been conceptualized as being based on unfiltered sensory input from an outside world passively registered by our senses. Internally then, that is mainly in the brain, sharply distinct and highly specified neural mechanisms, shaped by natural selection and encoded in genetic structures, process and combine the information received. This at the end leads to discrete modal as well as combined amodal representations of what is assumed to be the outside world, on the basis of which future behavior can be planned and anticipated (see for a pointed description also Noë, 2004, p.35). Within this framework, fundamental research of the last decades produced crucial insights into the principles of human cognition leading to enormous progress in sciences as well in applied fields such as IT, medicine or education.

Nevertheless, this idea of the mind has been challenged by embedded, embodied, extended, enacted and affective perspectives on cognition (referred to as 4EA approaches, see for example Ward & Stapleton, 2010) commonly stressing the highly dynamical character of cognition described as arising always anew in the interaction between animate bodies and their enlivened world. Shortly referring to the different aspects involved, these approaches can be described as characterizing the mind as being scaffolded by environmental structures (→ embedded cognition, see for example Donald, 1991), that are available to the mind due to the fact that it is realized not just in the brain, but rather in a body with its specific sensory

---

[7] See Horst (2011).

motor skills (→ embodied cognition, see for example O'Regan and Noë, 2001; Gallese[8], 2008). These skills enable the embodied mind to even integrate elements of the world in its cognitive functioning that are not part of the body as such (→ extended cognition, see for example Clark & Chalmers, 1998). Furthermore, such a complementation is considered to be realized not once and for ever, but again and again in a self-refining continuous interaction of this body with the world, including other social agents (→enacted cognition, see Varela, Thompson and Rosch, 1991; de Jaeger & di Paolo, 2007). Finally, it is taken into account that the body, in its interaction with the world, approaches and responds to it as a body in all its facets, including viscerality and affects (→ affective cognition, see Forgas, 2000; Colombetti 2014).[9] Taken together: for an embedded, embodied, extended, enactive and affective mind situated in a world "at his grasp", objects are not just passively perceived with regard to their physical steady features, but actively experienced and approached as affording interactions. The precise nature of these affordances on the one side depends on the capacities and current states of the embodied mind (that are not only given by his physical and sensual possibilities and actual feelings and needs but also by his prior experiences, see also Gibson, 1979), on the other side by possibilities only just arising from the interactions. Consequently, the realization of certain affordances might, depending on body and objects, create new experiences that, given time and space, give rise to new affordances and so on. In a nutshell, cognition is a work in progress, happening and evolving in the dynamic interaction between animate bodies and a world including objects and other social agents.

Some scholars argue that, given this concept, one can abandon notions such as "internal representations" altogether (see Hutto & Myin, 2012), while others only claim a major influence of habits of body-world interactions on all cognitive functions (still involving mental representations, although in a "bodily format", see Gallese and Sinigaglia, 2010, 2011).

Independent of this question and the answer chosen, the 4EA approaches have been enormously influential over the last decades in the humanities as well as in the natural sciences. Especially for the latter, this is mostly due to the fact that they allow to reformulate old and to

---

[8] [8] As Gallese argued: "Our vision of the world can be characterized as dynamic, relational and intentional. As Merleau-Ponty writes in *The Phenomenology of Perception* (1962),'The identity of the thing through perceptual experience is only another aspect of the identity of one's body throughout exploratory movements; thus they are the same in kind as each other'. I posit that this dynamic account of vision depends on agency. In this respect vision shares with other dynamic senses (for example,touch) the same logic of operation (2000, p. 34).

[9] Of course these inclusions and reconceptualizations were not done all at once. It was rather a slow widening of perspective on what should be considered to be crucial for cognition. From a traditional view describing the brain and our nervous system in their genetically prewired conditions as the only important player, positions step by step opened up to other possible contributors to cognition, that is culture and situational context, the lived body, and an inanimate and animate surrounding world with its capacities and helping facilities. For further explanations see back in the text.

pose new questions and hypotheses regarding long-known problems, often leading to surprising discoveries and answers in empirical research. In the following, I will illustrate this with respect to two questions, which are of apparent importance also for the perception of live-action edited moving images: firstly, the question of how we know about the inanimate environment, and secondly, the question of how we understand animate others.

To keep a clear structure, I will talk about the two questions separately, starting with our access to the inanimate world. Furthermore, I will always first sketch how the cognitivist approach to cognition did look at a certain problem, then outline the problematic aspects of the related models. Following this, I will describe some reconceptualizations suggested by scholars defending 4EA approaches to cognition, while lastly giving examples of experimental explorations of these latter approaches. Results will be summarized shortly at the end.

II.1.1 Question 1: How do we know about the inanimate world?

II.1.1.1 From passive perception to active exploration

As already sketched above, traditional cognitivist theories of mind and related scientific models describe perception as a process in which the brain uses a set of evolutionary preselected, highly specialized algorithms to (re-)construct a detailed representation of the outside world, starting from the "imprints" of this world registered by our senses (for scientific models describing vision in this way, see for example Marr, 1992; Poggio, Torre & Koch, 1985). In the following, I will explicitly focus on visual perception to illustrate problems and new solutions regarding this view.

One of the biggest difficulties, especially for visual sciences, has been the question of "how the brain bridges the gap between what is given to the visual system and what is actually experienced by the perceiver" (Noë et al., 2000, p.96). As a matter of fact, growing knowledge about the actual anatomical and physiological circumstances of perception, such as the constant movements of our eyes and the blind spot of the retina, made more and more of a miracle the fact that what we apparently perceive is not rambling patches of a fragmentary world, but rather smooth and continuous impressions of a stable environment. Traditional cognitivist takes on vision accounted for this by assuming and trying to model complex innate mechanisms compensating for the apparent flaws of the sensory information registered. This has been described as equivalent to attributing to the brain the task to produce what has actually to be called an "illusion" of the world, a world given to us as a complete whole only in form of an inner representation relying on sense-substituting postperceptual processing (for a review and critical discussion, see Bridgeman, Van der Hejiden, & Velichkovsky, 1994; Pessoa, Thompson, & Noë, 1998).

One way to expose the limitations of such a view is by taking into account the phenomena of change and inattentional blindness. Change blindness refers to the intriguing fact that under certain circumstances our capacity to detect even obvious perceptual changes in our environment can be significantly attenuated, even if we are explicitly told to watch out for these changes. While early investigations suggested that this apparent perceptual failure was caused by eye movements such as saccades or blinks restricting sensory input (Blackmore, Brelstaff, Nelson, & Troscianko, 1995; O'Regan, Rensink, & Clark, 1996; Resink, O'Regan, & Clark, 1997; O'Regan, Deubel, Clark, & Rensink, 2000; O'Regan, Rensink, & Clark, 1997), further exploration of the phenomenon supported the notion that instead the effect is mostly due to the participants' attentional focus only (Rensink et al., 1997). Maybe most spectacularly, this was demonstrated by a study of Simons and Levin (1998, see also 1997) showing that even in a real life scenario and over an extended period of time, people that are occupied with a cognitive task, such as explaining directions or taking a photo of a stranger, often do not detect huge changes such as the substitution of this stranger by another[10].

These results were taken by proponents of 4EA approaches to cognition as evidence that contrary to cognitivist assumptions, in normal viewing conditions only a minimal part of the environment is actually sampled by the brain (Noë, Pessoa, & Thompson, 2000[11]). Yet, it is also a false idea to assume that our knowledge about the outside world relies on a "grand illusion" – that is a complete and detailed mental reconstruction on the basis of disjoint snapshots of minor quality. On the contrary, it is suggested that we do not have to build such an internal illusion because "seeing constitutes an active process of probing the external environment as though it were a continuously available external memory" (O'Regan, 1992, p.484). In other words, we experience the world as being so rich and detailed because we constantly have access to a spatio-temporal continuous environment. And it is on the basis of this experience-based knowledge that we explore the world, focusing on the information supposedly needed in each and every single case.

---

[10] In cases in which participants are on purpose asked to engage in another task, this phenomenon is also described by the term "inattentional blindness". It has meanwhile been demonstrated by a great number of studies, while also being captured on many videos to be found on the Internet. For an example from an American television show see http://www.youtube.com/watch?v=qN7s9E6M4RQ(last visited the 5th of January 2015).

[11] However, for studies finding evidence that some information apparently not available for consious awareness might be still registered see Williams & Simons (2000) and Koivisto & Revonsuo (2003). Both studies report that in trials of change blindness experiments in which the participants did not detect a change occuring, still reaction times were delayed in comparison to real no-change trials, suggesting that "at some level the change was detected although explicitly no change was reported" (Koivisto & Revonsuo, 2003, p. 425).

In a nutshell: 4EA approaches propose that it is not a highly workloaded processing and optimization of a great amount of passively received environmental stimuli that leads to our detection or neglect of details and changes in the world, but our skill and our experience as a perceiving body in this world guiding our explorations and anticipations. This reconceptualization led to specific predictions regarding neuroimaging and physiological data acquired during stimulus perception in the tasks reported on. Specifically, scholars hypothesised to find activations before actual stimulus appearance that would predict detection as well as further processing of this stimulus, while reflecting either conscious awareness or neglect. Indeed, recent studies by Koivisto & Revonsuo, 2003; Pourtois, De Pretto, Hauert, & Vuilleumier, 2006 and others (for a review, see Railo, Koivisto, & Revonsuo, 2011), designed following this suggestion, describe such effects. Specifically, Pourtois et al. (2006), in an ERP study, found that visual awareness of a change was associated with an enhanced occipital P1 response and a sustained frontal activity after the first display of the image, that is before the change occurred. Furthermore, Pourtois et al. (2006) as well as Koivisto & Revonsuo (2003) found that change detection elicited a stronger early negativity and a more pronounced late positivity peaking at parietal sides. I will come back to these findings in Chapter III.2. For the moment, they should be taken as an illustration of how the described change in the theoretical paradigm conceptualizing cognition allowed the development of new hypotheses and experimental designs, leading to new insights into basic principles of cognition also in experimental cognitive neuroscience. The next section will give more examples of that, still referring to our knowledge of the inanimate world while stressing another reconceptualization thanks to 4EA approaches to cognition.

II.1.1.2 From amodal representation to multimodal body knowledge

Based on the idea that what we know about the world is the passive perception of what there is to see (see last section), cognitivist concepts also assume that perception is explicitly designed to equip us with objective information about the world. Senses, according to this hypothesis, are described as entrances to distinct pathways, processing the received information via symbolic transformations independently of current body states or other influences, finally leading to an objective quintessence of the stimulus describable in abstract and amodal symbolic terms (for perspectives embracing such amodal account see Haugeland, 1991; Pylyshyn, 1973; for rejections see Barsalou, 1999; Gallese, 2000; Gibbs, 2006; Wilson, 2002).

Again, this view has been challenged in the last years. Firstly, some studies indicated that even when explicitly asked for a neutral evaluation of physical information of the world, we are often not able to do so. Secondly, studies further investigating how stimuli are processed in those tasks indicated that to derive knowledge about the world we rely on interaction with the world rather than on the rational processing of neutral information. To give a prominent example of the first finding, that is our obvious incompetence to neutrally evaluate the physical information of the world, Bhalla and Proffitt (1999) found that when asking participants to judge the steepness of a hill by visual inspection only, those participants carrying a heavy backpack during this task judged the slope steeper than those of a comparison group not carrying any weight. Schnall, Harber, Stefanucci, & Proffitt (2008) on the other hand found that participants judged hills as less steep when accompanied by or even only thinking about a friend.

Regarding the second finding, that is our apparently interactive processing of information, Shepard & Metzler (1971, see also Kosslyn, 1980) found that, when asking participants to decide as quickly as possible if the difference between two objects shown after each other was only in the orientation of the object or also in its shape, response time was proportional to the degree of rotation distinguishing the two objects. According to scholars supporting 4EA approaches to cognition, these findings indicate that cognition is fundamentally grounded in the needs and embodied capacities of human beings (see Barsalou, 2008). Rather than focusing on the extraction of physically correct information while relying on complex and resource-demanding symbolic processing, cognition proves to be pragmatic, i.e. it uses its bodily capacities and experience to come to an action-oriented estimation of the situation[12]. Remarkably, this even holds for tasks pursued in imagery only.

Again, this new view on cognition led to predictions also in cognitive neuroscience, for example the hypothesis that also processes not involving direct actions, such as pure observation, recall or imagery, on a neural level should as well activate networks usually responsible for real interactions. Indeed, imaging and EEG experiments of recent years collected interesting evidence that there is a crucial overlap of activations measured during visual perception and those recorded during visual imagination (see for a good overview also stressing task dependent differences Bértolo, 2005). Most impressively maybe, as indicating the existence of direct action-perception links even on a single cell level, Murata, Fadiga,

---

[12] To give an example, in a situation in which you have to decide if and in what pace you want to climb a mountain, the weight of a luggage to be carried or the presence of a friend can make a crucial difference. It is remarkable though, that this action oriented reasoning is still applied when the task explicitly demands to solve a geometric problem apparently unrelated to any active decision to be taken in the close future.

Fogassi, Gallese, Raos, & Rizzolatti (1997) described a special kind of neurons in the ventral premotor cortex (area F5) of macaques that fire during grasping movements as well as during the observation of graspable objects.[13] Not surprisingly, since their discovery these "canonical neurons" have been granted enormous attention in further research exploring the hypothesis of 4EA approaches to cognition. The next section, focusing on different answers to the question of how we understand animate others, will further illustrate this by giving the example of another type of neurons showing similar capacities: the mirror neuron.

II.1.2 Question 2: How do we understand animate others?
- From mentalizing to embodied forms of empathy

4EA approaches to cognition also criticized and extended the notion of empathy and the understanding of other minds as conceptualized in traditional cognitivist theories. In these theories, these capacities were almost without exception considered as based on "mind reading". This term commonly refers to the process of rationally inferring the mental states of others via complex second-order representations realized either by means of a domain general theorizing capacity or the maturation of a special organ dedicated to the domain of psycho-logy (see for example Baron-Cohen, Leslie, & Frith, 1985, or Baron-Cohen, 2001, and for a general overview Marraffa, 2011).

Even before the rise of 4EA approaches, the idea of this kind of processes being the only base of action understanding and of empathy has been challenged from developmental and evolutionary psychology. The main criticism consisted in the observation that human infants and even human primates show behavior that would typically be interpreted as demonstrating empathetic understanding, although the ones behaving in such way apparently lack the intellectual capacities that would allow high level operations underlying mindreading. Precisely, Jules Masserman et al. (1964) reported, that rhesus monkeys refused to pull a chain supplying them with food if this act at the same time elicited an electric shock being given to another monkey (Masserman, Wechkin, & Terris, 1964). Also, scholars studying infant behavior showed gaze preference for attacked others in infants as young as 10 months of age (Kanakogi, Okumura, Inoue, Kitazaki & Itakura, 2013). Even stronger however, the develop

---

[13] For findings giving evidence for the existence of the same phenomena in humans see Martin, Wiggs, Ungerleider, & Haxby, 1996; Grafton, Fadiga, Arbib, & Rizzolatti, 1997; Perani, Schnur, Tettamanti, Gorno-Tempini, Cappa, & Fazio, 1999; Chao & Martin, 2000; Gerlach, Law, Gade, & Paulson 2002; Boronat, Buxbaum, Coslett, Tang, Saffran, Kimberg, & Detre, 2005; Ponseti, Bosinski, Wolff, Peller, Jansen, Mehdom, Büchel, & Siebner, 2006; Lewis, 2006; Hulme & Zeki, 2006.

mental psychologist Vasu Reddy convincingly demonstrated that if taking into account the situated embodied nature of cognition and the special importance of this embeddedness in infancy, signs of empathy and action understanding can be found even earlier[14]. Precisely, Reddy argued for looking at infants' behavior not during the observation of interactions between others, but during their own interactions with their caregivers in their daily routines. These studies showed that infants as young as two months of age, when perceiving a caregiver about to lift them, make specific adjustments of their bodies, assisting in the smooth execution of the picking up (Reddy, Markova, & Wallot, 2013). Also, Reddy reported that infants from around 3 months of age show clear signs of coyness (Reddy, 2001) and infants from 9 months of age even start to tease others. That means, they engage adults in interactions in which they offer and withdraw objects, deliberately not comply to commands or disrupt other's actions, while looking intently at the other's face and showing obvious signs of enjoyment as a response to adults' expressions of (faked) indignation (Reddy, 1991, 1998, 2008; Reddy, Liebal, Hicks, Jonnalagadda, & Chintlapuri, 2012). It is hard to imagine, how this kind of behavior can arise and be realized if the cognitive capacities of infants do not allow them to have any idea about feelings, sensations and thoughts of animate others – which according to classic cognitivist theories have to be constructed as higher order represent-tations via complex mentalizing processes.

These observations also fit to a maybe even more fundamental critique of cognitivist concepts of empathy: phenomenological descriptions of the understanding actually discussed, mark this understanding as an immediate, implicit and (embodied) feeling, rather than a linguistically reflected inference of someone else's mental state based on time-consuming conscious interpretations (see Zahavi, 2012a, 2012b, Gallese, 2001, 2003). As for example Dan Zahavi pointed out, long before "knowing" about the "state of another", we feel it "in our gut". (Zahavi, 2012b).

Following this intuition, supporters of 4EA approaches to cognition have brought up different reconceptualizations of empathy and action understanding. In these notions, primary, embodied and non-inferential ways of knowing another's mental states are outlined (e.g. Gallese 2001, 2003, 2005; Goldman and Gallese, 1998; Goldman, 2006; Iacoboni, 2009, Gallagher 2005, 2008, 2011a, b, 2012; De Jaegher, Di Paolo, Gallagher, 2010; Rizzolatti and Sinigaglia 2010; Zahavi 2005, 2007, 2008, 2010, 2011, 2012a, b). One version of these reconceptualizations that has become especially prominent in the field of neuroscience is

---

[14] In fact Reddy stresses how terms related to the old conceptualization, such as "intention" (as refering to internally represented action goals of animate being) also need a crucial reconceptualization in view of this background. See Reddy, V. (in print) as well as Heimann & Uithol (in print).

called (Embodied) Simulation Theory. Scholars defending this concept claim that understanding the behavior and feelings of others fundamentally involves – at least at a basic level – processes of automatic bodily simulation building on the reuse of bodily formatted representations of actions, emotions and sensations rather than rational inference (Gallese 2003, 2005; Gallese & Sinigaglia, 2011). Watching others, so goes the hypothesis, we effortlessly and immediately "put ourselves in their shoes", that is, we literally perform a simulation of their supposed state using the capacities of our body and embodied mind instead of engaging in complex mentalizing processes.

Again, this hypothesis gave rise to new approaches investigating the problem also in cognitive neuroscience. It gained even stronger prominence two decades ago with the discovery of the mirror neurons, disclosing a likely candidate of a neural mechanism underlying such simulative processes. The specific characteristic of mirror neurons, originally discovered using single cell recordings in the ventral premotor cortex (F5) of behaving macaque monkeys (Di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Fogassi, & Gallese, 1996), consists in their mutual activation during the performance of goal-directed action as well as during the observation of another individual performing this action. Further experiments showed that this apparent direct action-perception link not only shows action-goal specificity (with certain neurons only firing for performance and observation of certain actions, see Rizzolati, Fogassi, & Gallese, 2001), but can also be evoked even when the action's outcome cannot be seen but only rather imagined by the monkey (Umiltà, Kohler, Gallese, Fogassi, […], & Rizzolatti, 2001). Furthermore, it was discovered that a particular class of mirror neurons ('audio-visual mirror neurons') are activated not only by action observation and execution but also also by the sound produced by someone else's action (see Kohler, Keysers, Umiltà, Fogassi, Gallese, & Rizzolatti, 2002; Keysers, Kohler, Umiltà, Fogassi, Rizzolatti, & Gallese, 2003). Lastly, it was found, that the activation is enhanced for actions with which the monkeys had prior experiences (see for example Rochat, Caruana, Jezzini, Escola, […], & Umiltà, 2010). From the very beginning, therefore, mirror neurons were hypothezised to fundamentally serve our worldly orientation by underlying the understanding of the actions of others.

Recent years delivered extensive evidence for the existence of similar mirror systems also in humans (see Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Grafton et al., 1996; Rizzolatti et al., 1996; Cochin, Barthelemy, Lejeune, Roux, & Martineau, 1998; Decety, Grèzes, Costes, Perani, Jeannerod, Procyk, Grassi, & Fazio, 1997; Hari, Forss, Avikainen, Kirveskari, Salenius, & Rizzolatti, 1999; Iacoboni, Woods, Brass, Bekkering, Mazziotta, & Rizzolatti, 1999; Buccino, Binkofki, Fink, Fadiga, Fogassi, […], & Freund, 2001; Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried, 2010). Moreover, recent findings suggested that similar links between first and third person experiences can be assumed to originate in shared networks processing sensations and emotions (Gallese, 2003). Specifically, it was found that the experience of touch and the observation of somebody being touched (Keysers, Wickers, Gazzola, Anton, Fogassi, & Gallese, 2004; Keysers & Gazzola, 2006) as well as the experience of emotions and the observations of others showing matching emotional expressions (in case of a match of theses emotions) led to the activation of similar brain areas and cortical networks (see for example. Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003; Wicker, Keysers, Plailly, Royet, […], & Rizzotti, 2003; Keysers & Gazzola, 2006).

Together this has been taken as evidence that human's apparent knowledge about, and anticipation of, other people's feelings and future actions rely on the fundamental biological principle of the brain to employ the very same networks for the realization of one's own motor plans and the observation of others moving, as well as the experience of one's own emotions and feelings and the observation of other's behavior or expression of such feelings (Rizzolatti, 2005; Iacoboni, 2009; Keysers & Gazzola, 2006; Gallese & Sinigaglia, 2011; though for thorough critique, especially regarding the role of mirror neurons for action understanding, see Hutto, 2013; as well as Hickok, 2014). I will specifically get back to this hypothesis in Chapter III.1 of this thesis.

Taken together, the previous sections showed that 4EA approaches to cognition allow a fundamental reconceptualization of the relation between humans and world, changing the picture from a rather mechanical process between a high-tech robot-like human and a material environment behaving according to physical laws, to the dynamic interaction between an animate bodily being and an enlived habitat known and continuously shaped through enactive experience. This new perspective highlights neuroscientific research looking for traces of cognitive processing depending on a) situative embedding (reflected for example in attention and body-state), rather than on stimuli features alone, as well as b) embodied problem solutions relying on processes of (simulated) interactions rather than on symbolic transformations. In the next two chapters, I will illustrate how this turn was reflected in research about visual media and arts in general and film in particular.

II.2 Enacted, embodied, embedded, extended and affective cognition as a concept guiding visual art theory and empirical research about visual art perception

As the introductory chapter of this thesis illustrated, we acquire a crucial part of our world knowledge via media. This leads to the question of how 4EA approaches to cognition have influenced theory and research about media experiences. Film being a relatively young medium, it makes sense to start this investigation by looking at embodied approaches in fields dealing with visual representations of the world in general, i.e. theory and philosophy of images.

Interestingly, as a first remark, there is still a perceivable reluctance to leave traditional cognitivist frameworks when looking at aesthetic evaluations of artworks (see for previous notice of this for example Sauer, 2012; Freedberg & Gallese, 2007; Gallese & Di Dio, 2012). At least three reasons are likely to contribute to this resisting hesitation.

Firstly, it is a plain fact, that images show us a world with which we cannot interact directly, that is in the crucial sense of bodily moving in and exploring it. Being “only” on the image, none of my behaviors will grant me literal access to this world. This difference is not only known to us by convention and experience but is rather present for us in the very moment of perceiving an image – “seeing a world in an image” means to consciously perceive, at the same time, the representation and the represented[15]. This difference from real world perception might have contributed to the belated discovery of 4EA approaches in the field.

Furthermore, secondly, for a long time, that is especially during medieval age, art, creation and perception were indeed guided by strong conventions, such as the symbolic representtation of persons and concepts, and even the use of colors and materials, accessible only via specialized knowledge. In later stages of art history, knowledge of the representational habits and contents, marked by the current understanding of beauty or by politics, remained of tremendous relevance. This accounts also for the background of the individual artist, his positioning in the arts of former and current times and the development of his works over the duration of his career. Lastly, even an individual artwork might refer to a context that needs to be known rather than bodily explored to assess the fundamental dimensions of this work of art. Naturally, interpretations of artworks through almost all periods are strongly marked by this access through knowledge, which sometimes has been characterized as similar to

---

[15] See also for example Fingerhut (2012), referred to later in this thesis.

"deciphering" rather than "perceiving" or "seeing" an image (see for such descriptions for example Sauer, 1999). An additional role in supporting this disembodied approach to arts can also be assigned to traditional practices of art exhibition, which do not only follow the above tradition, but also in a rather embodied sense intend to free artistic judgement from normal perceptual routines. Until today concepts like "the white cube" heavily influence the architecture of museums, leading to a presentation favoring decontextualization and estrangement from the real world rather than embedded cognition derived from this real world. Furthermore, trends like audio-guides that provide background information and point out the most important must-sees again refer to art history rather than to our bodies and perceptual and cognitive skills as the right guide to art exploratory perception[16]. One could indeed assume, that the perceptual habits newly arising in such a interaction-free and world-detached context lean heavily on the application of cultural knowledge and symbolic reasoning best purified from bodily groundings.

Thirdly, it needs to be emphazised that within art and media theory, bodily responses and affects, which undoubtedly stand in the center of 4EA approaches to cognition, carried for a long time the stigma of being highly subjective and context-dependent, and were therefore excluded from discussions that investigated the general principles of aesthetic quality and expression. In fact, in his canonical work *The Principles of Art,* R.H. Collingwood clearly states that, although it is precisely the task of artists to express emotions, they have to do so by liberating these emotions from their bodily, sensous base: "The aesthetic activity is an activity of thought in the form of consciousness, converting into imagination an experience which, apart from being so converted, is sensuous"(Collingwood, 1938, p. 295). Collingwood furthermore indicates that this act needs to convert emotion from something non-contingent as a feeling or a body state, into a conventional symbol, thereby making it accessible and descrybable by all perceivers and art theory: "This activity [the aesthetic activity] is a corporate activity belonging not to any one human being but to a community. It is performed not only by the man whom we individualistically call the artist, but partly by all the other artists […] from whom he borrows, executants whom he employs and audience to whom he speaks." (Collingwood, 1938, p. 295). This stance was carried on via famous art historicians such as E.H. Gombrich and C. Greenberg, in their works about the psychological basis of arts exclusively focusing on cognitive levels. It even seems to have influenced philosophers out of the semiotic field, such as Nelson Goodman or Umberto Eco, who explicitly dedicated their work to the

---

[16] With this I am not referring to an exploration in a sense of a strict scientific experiment, but to perceptual exploration, that is an interaction of the spectators with the artwork allowing them to be guided by and play with its perceptual offers etc.

distinction of language and other representational systems (see Gombrich (1960), Greenberg (1961), Eco (1998, 2000), Goodman (1995)). This is not to say that traditional art history or semiotics did not contribute tremendously to our understanding of non-linguistic representations. Rather, what I would like to stress is, that some topics seem to be explicitly excluded by these approaches. Paradigmatic for semiotics' resistance against bodily groundings of symbolic references might be the following lines by Nelson Goodman writing about the separation of artistic expression of emotion from emotion per se: "An actor's facial expression need neither result from nor result in his feeling the corresponding emotions. A painter or composer does not have to have the emotions he expresses in his work. And obviously works of art themselves do not feel what they express, even when what they express is a feeling." (Goodman, 1995, p.47); and a bit earlier, writing about the conventional nature of "resemblance": "Again, what will deceive me into supposing that an object of a given kind is before me, depends upon what I have noticed about such objects, and this in turn is affected by the way I am used to seeing them depicted. Resemblance and deceptiveness far from being constant and independent sources and criteria of representational practice are in some degree products of it." (Goodman, 1995, p. 39). Even when talking about similarity, so Goodman goes, we in fact talk about conventions. We recognize an emotion when seeing it on the face of another, as we learned to do so, we know that a child imitates a telephone when talking to a banana due to cultural experience with these kind of denotations and we detect vast fields of wheat under troubled skies in Van Gogh's *Wheat fields with crowds* (1890)[17] just as we find a human figure in a stick man: by convention. The reason for this body-denying view of image perception can only be guessed from few remarks of the authors about the importance of body and (bodily being)[18]. Reminiscent to Collingwood's words quoted above it seems to lie in the will to prevent the explorable realm of cultural distinctions, accessible by language, from diving into the subjective darkness of bodily being. As Eco formulates it: "Being underpins all discourses except the one we hold about it (which tells us nothing we did not already know the very moment we began to talk about it)" (Eco, 2000, p. 24)[19].

Obviously, none of these three reasons for not looking at images from a 4EA stance withstand further reflection. Firstly, referring to the objection that the world on an image denies bodily access, Chapter II.1 should have sufficiently illustrated that the real world as well is never

---

[17] Van Gogh: Wheatfield with Crows, 1890. Oil on canvas. 50.2 cm x 103 cm. Van Gogh Museum, Amsterdam.

[18] Admittedly, again, to give such an explanation is also not the intention of the authors in their brilliant and most illuminating writings about non-linguistic representations!

[19] For an elaboration on this apparent fear see also Heimann (unpublished Master thesis).

given to us completely. This is not only true due to selective attention depending on my current task, or my physical dependencies not allowing me to be everywhere at the same time. Indeed, some things we think we know of, such as the state of others, can never be perceived in the exact same way in which we perceive ourselves or an object we handle or feel.

The explanations in Chapter II.1 should have illustrated then how, still, the world itself is experienced via embodied mechanisms.

Furthermore, regarding the other two objections to an embodied concept of image perception, that is, the traditional deciphering techniques and the learned conventions of representations that are suggested as the main and most useful tools to interpret images and artworks, I assume that everyone of us has experiences with such representations that clearly do not seem to be explainable by such knowledge. It seems indeed rather implausible to think of a child's charade or our immediate interpretation of the flickering brushdots in a painting of Cezanne as being guided by learned rules of how to show something. And I would even suggest that it is almost impossible to talk about Goya's etching *Que hay che hacer mas?*[20], showing the brutal genital mutilation of a naked man by several soldiers armed with sabres, without including our almost unavoidable bodily reaction to it in our descriptions. Precisely most spectators might strongly remember responses such as squinting the eyes and slightly moving backwards as if trying to flee the painful empathy evoked by this and similar paintings (see also Freedberg & Gallese, 2007). In the same vein, it does not seem reasonable to talk about Anselm Kiefer's *Die Treppe*[21] without referring to the corporal feeling and behavior evoked in the spectator by the overwhelming size and materiality of the artwork. Indeed, some images cannot only be felt as "forcing us in a distance", or "drawing us close" in our attempts to best grasp (or avoid) their affective(!) content, but also as influencing our mood and whole being explicitly by stimulating specific ways of perceptual explorations: rapid gaze jumps or long fixations, such as those evoked by certain material formations or uses of colors, in their gestural nature don't leave us unaffected (see also Sauer, 2006). In a nutshell: neither resemblance nor expression (as two functions of images and artworks) seem to be capturable with reference to learned conventions only.

As remarked in most recent attempts to (re-)introduce body, emotion and affect into the discussion about the principles of art perception (see for example Freedberg & Gallese, 2007;

---

[20] Goya, Que hay que hacer mas? (What more is there to do?), plate 33 from Los Desastres de la Guerra (Disasters of War), etching, Bibliothèque nationale, Paris, France.

[21] Anselm Kiefer, Die Treppe, 1982/83, straw, emulsion, shellac and burn marks on photography, 330 x 185 cm, Artmuseum Bonn – permanent loan Grothe

Di Dio & Gallese, 2009; Gallese, 2012; Gallese & Di Dio, 2012; Sauer, 2012; Fingerhut, 2012), these insights are not completely new. Indeed, a strong tradition of art historians, starting long before the rise of cognitive sciences, emphazises the crucial importance of the living body and its experiences for images and especially visual art experiences. Without the intention to give a complete review of this literature I would like to illustrate the principal ideas developed over time using some selected examples.[22] To begin with a first strong position, Robert Vischer, as early as in 1874, created the term "Einfühlung" [literally translatable as "feeling-in", K.H.] to underline that even plainly looking at an object, and especially at art-works, can evoke non-contingent, sensual and motoric responses of our body. In fact, Vischer's explanations seem to suggest that the body is used as the main access to the artwork, when he writes "Mein geistig sinnliches Ich transportiert sich in das Innere des Objektes und erfuehlt seinen Charakter von innen heraus." (Vischer (1874), p. 48; "My mental sensual I transports itself in the inside of the object to feel (in the sense of explore) its character from the inside." Translation K.H.). This view is apprehended and further developed, especially in the field of architecture, by Heinrich Wölfflin (1946), who stresses that it is the constitution and organization of the body that determines our comprehension of any kind of material artwork. Almost reminding of Wittgenstein's "Die Grenzen unserer Sprache sind die Grenzen unserer Welt" (Wittgenstein, 1922, 5.6; "The limits of my language mean the limits of my world", translation K.H.), Wölfflin writes "Unsere leibliche Organisation ist die Form unter der wir alles Koerperliche auffassen." (Wölfflin, 1946, p. 21; "Our bodily organization is the form according to which we grasp anything corporal (in the sense of material)." translation K.H.). He expresses further that this bodily organization is nothing completely stable, but is influenced by the transient changes of our situated being in a evolving world, stating "Ein architektonischer Stil gibt die Haltung und Bewegung der Menschen seiner Zeit wieder" (Wölfflin, 1946, p. 44; "An architectonic style mirrors the (bodily) stance and movement of the people of its time", translation K.H.) Nevertheless he also stresses, "die Organisation des Körpers [sei, added by K.H.] der bleibende Nenner bei allem Wechsel" (Wölfflin, 1946, p. 44; "The organization of the body remains the common denominator throughout all changes"). This clearly marks the dominance of bodily nature in comparison to culture and education in the sense-making processes referred to. 50 years later, Aby Warburg and Bernard Berenson famously picked up again this line of thought. While Warburg emphasizes once more, that the outward look of a human figure is enough to

---

[22] For more literature, see for example Stueber, (2006, 2009), especially focusing on empathic feelings in aesthetic appraisal and Curtis & Koch (2009) about the history and presence of "Einfühlung" as an aesthetic concept.

grant us a look in its inside and highlights the immediateness and automaticity of this impression (Warburg, 1999), Berenson specifies the underlying mechanism, that is that the observation of an artwork can raise the spectator's inner sense of the muscles involved in the represented posture (Berenson, 1896). While these descriptions refer apparently to the representation of human figures above all, later scholars elaborated on Wölfflin's seemingly broader claim that the experience of the body is indeed the essence of *all* the world perceived by it. As Merleau-Ponty puts it: "The enigma derives from the fact that my body simultaneously sees and is seen. That which looks at all things can also look at itself and recognize, in what it sees, the "other side" of its power of looking. [...] Visible and mobile, my body is a thing among things; it is one of them. [...] Things are an annex or prolongation of itself; they are incrusted in its flesh, they are part of is full definition; the world is made of the very stuff of the body. [...] Since things and my body are made of the same stuff, vision must somehow come about in them. [...] Quality, light, color, depth, which are there before us, are there only because they awaken an echo in our bodies [...]."(Merleau-Ponty, 1993, p. 124-125). Thirty years later, the developmental psychologist Daniel Stern proposed again that our basic access to the world surrounding us consists in what he called "vitality forms": "activation contours of certain bodily sensations" elicited in response to motions as well as shapes via an immediate and automatic resonance activity of the human mind, bridging the apparent gap between inside and outside (Stern (2000) p.89, for an elaboration see also Grueny (2014 and unpublished manuscript)). While Stern admits that this kind of perception can in later stages of life be substituted by other mechanisms based on linguistic categorizations, etc., he also stresses that it is still available and used, especially in situations that confront us with uncategorizable phenomena – such as the arts (see Stern, 2010, as well as Langer, 1953,1957,1984).

It is with explicit reference to this line of thought that David Freedberg and Vittorio Gallese set forward in 2007 their persisting call to finally leave behind the primacy of (disembodied) cognition in art research while instead focusing on the role of embodied cognition and especially embodied empathetic responses to art. Even more concretely, they formulated two distinct hypotheses regarding the precise functions of these kinds of responses in art perception.

Firstly, they suggested that empathetic responses, implemented via fundamental action and emotion-perception links (e.g. shared neural networks and canonical and mirror neuron mechanisms as introduced in Chapter II.1 and II.2), play a crucial role in the observer's embodied perception of the representational content of the artworks (action, intentions or emotions of actors as well as affordances of objects). This role is to animate the depicted content by animating the perceiver. As the descriptions of prior scholars have demonstrated, it is a curious fact that despite the obvious differences between a world represented by artworks and the world that literally surrounds us, artworks show the clear capacity to bring something to appearance rather than "just depicting" it[23]. Looking at Michelangelo's *Prisoners*[24] or Goya's *Desastres de la Guerra*[25], we do not perceive figures bizarrely frozen in time, but rather experience strongly and emotionally an extended moment of flight, fight and action. Freedberg and Gallese suggest that it might be precisely via the activation of our motor system mediated by the mirror neuron system and embodied simulations triggered by such that this is possible.

Secondly, they propose that empathetic responses via the same mechanisms contribute to the observer's embodied perception of the making of the artworks, that is the specific style of an artwork which depends on the artist's creative gestures. Precisely, they hypothesize that also the visible traces of the artist's creative gestures (such as brushwork etc.) can also activate the observer's motor system, thus influencing the experience as a whole.

This idea led to some criticism and extensions. Joerg Fingerhut (2012) discusses for example Shaun Gallagher's critique of Gallese's and Freedberg's position (Gallagher, 2011c). According to Gallagher, Gallese's and Freedberg's argument either downplays what there is to be said with respect to mirror neurons in social cognition, or misses out on aesthetic experience, because it does not account for the ways in which encounters with artworks also differ from real-world situations. One way to mark this difference is to refer back to the relation of "seeing-in" to them, mentioned above: artistic representation are always perceived as on the one hand showing something, that is the depicted, while, on the other hand, also showing themselves, that is the depiction itself. Fingerhut in contrast, interestingly proposes, that, contrary to Gallagher's doubt, the second hypothesis raised by Gallese & Freedberg, – the spectator's affection also by the artist's gesture itself – might indeed be a way to explain an aesthetic version of this phenomenon of seeing-in. With reference to fundamental works in vision research that delivered evidence of the existence of two different pathways in visual perception (see Ungerleider & Mishkin, 1982; Millner & Goodale, 1995; see Jacob & Jeannerod 2003 for a discussion of the philosophical impact) Fingerhut suggests that in normal viewing

---

[23] See for example Paul Klee's reflections on art in Paul Klee, 1920, p. 28.

[24] A series of unfinished sculptures for the tomb of Pope Julius II, produced 1525-1530 by Michelangelo. See for example: The Awakening Slave, 2.67m marble staturr, Galleria dell'accademia, Florence.

[25] A series of 82 graphics (aquatinta) by Francisco de Goya created 1810-1840, showing the atrocities of Napoleon's soldiers fighting against the revolting people of Spain (and vice versa). See also Paas-Zeidler (1978).

conditions, information regarding the surface of the image and information regarding the depicted objects might be processed mainly separately. Indeed, while research has described a ventral stream that was suggested to predominantly serve our conscious impression of the environment as well as related knowledge aquisition, a dorsal stream has been proposed to play a more important role in action preparation (precisely, it would collect fine grained information relevant for a bodily interaction with objects; see Ungerleider & Mishkin, 1982; Millner & Goodale, 1995; and Jacob & Jeannerod, 2003; however for a critique of making a too strong separation between these different pathways, see Gallese, 2007). According to Fingerhut it is thus likely that while the depicted gets mostly processed via the ventral stream, the material features of the depiction itself, obviously more relevant for a direct object handling, might be rather be processed via the dorstal stream. This leads Fingerhut to the further hypotheses that a) our feeling of presence of the depicted objects might fundamentally rely on processing via ventral pathways (including simulation processes initiated or supported by the mirror response to the depicted figures and objects[26]), while the immediate processing of the visual surface of the image via the dorsal pathway rather elicits the (simultaneous!) awareness of the presence of the representation – the material of the artwork – itself. Furthermore, he suggests that the spectator's perceptual activities such as gaze and head movements, are thereby also strongly influenced by the representation (rather than by what is represented only), which makes them differ from the ones that we would normally exhibit if exploring the depicted objects in reality.[27] This can be seen as the very basis of the spectator's possibility to gain new skills in the interaction with the artwork, as it allows him to see something under new perceptual conditions precisely evoked by his own exploratory movements guided by the representation rather than by the represented.[28] I will come back to these hypotheses later in this thesis.

Remarkably, meanwhile, Gallese and Freedberg's hypotheses have been supported by a number of recent neuroscientific studies. Regarding the first hypothesis – that the action perception links reported might play a role in animating also steady images of actors – it was repea-

---

[26] However, it should be added that mirror mechanisms, as far as we know, belong to and/or are recipient of inputs from the dorsal stream.

[27] This latter interaction between material cues and orientating movements, might lead to medium-specific new perceptual habits, crucially influencing not only how but also what we see. I will come back to this topic later in this thesis.

[28] This propositon might even extend Gallagher's main claim that artworks, by confronting us with unfulfillable affordances, can help us to see other action possibilities, that then might be available for us in later real life. Fingerhut promotes, that the spectator can additionaly profit from the fulfillable aesthetic affordances of the representation that might include the marks of the artists that could lead the spectator to a bodily behavior unknown from the real world.

tedly shown that mirror and canonical neurons can be activaed during the observation not only of life actions or real objects, but also of moving and even static images of actions or objects (see for example Johnson-Frey et al., 2003; Urgesi, Moro, Candidi, & Aglioti, 2006; Proverbio, Riva, & Zani, 2009; Grèzes, Armony, Rowe, & Passingham, 2003). Furthermore, it was demonstrated that it is not necessary for their activation that the representation be naturalistic. It has been found rather that also abstractions, such as films showing robots instead of humans as well as even comics can elicit the crucial responses (Gazzola, Rizzolatti, Wicker, & Keysers, 2007; Iwase, Ouchi, Okada, Yokoyama, […], & Watanabe, 2002).
Furthermore, regarding their second hypothesis – that also traces of artistic creations might be able to elicit the mirror mechanism – most recent studies showed that, next to life actions and abstract static representations of such, even the mere traces of actions, such as handwritten letters on or cuts in paper can elicit the activations of the motor system (see Longcamp, Tanskanen & Hari, 2006; Longcamp, Boucard, Gilhodes, Anton, […], & Velay, 2008; Heimann, Umiltà, & Gallese, 2012; Umiltà, Berchio, Sestito, Freedberg, & Gallese, 2012). Indeed, results also indicated that the responses to such stimuli were modulated by the detailed form of the representations, that is the precise actions needed to produce the stimulus, such as typing, handwriting or brushpainting. For example, Longcamp et al. found that printed letters elicited a lower motor response than hand-written ones. (Longcamp et al., 2008). Furthermore, directly looking at abstract works of arts, Umiltà et al. (2012) found that, in comparison with control stimuli consisting of matching linework produced in Adobe Photoshop, only original artworks by Luciano Fontana consisting of cuts in a canvas elicited a motor response during observation. An ERP study by Sbriscia-Fioretti, Berchio, Freedberg, Gallese, & Umiltà (2013) gave similar results when comparing responses to original paintings of Franz Kline with similar line-art not exposing the dynamic features of the brushstrokes. Importantly also, the later two studies found significant results in tasks assessing the aesthetic subjective experiences of spectators supporting the hypothesis that the stronger activation of the motor system coincided with enhanced aesthetic appraisal of the stimuli.

Taken together these findings clearly support the idea that future research looking at visual arts or visual representations in general should no longer neglect embodied approaches to perception but should rather invest in further exploration of the hypotheses raised. In the next subchapter I will illustrate, why this is especially needed regarding the case of research about life-action moving images.

II.3 Enacted, embodied, embedded, extended and affective cognition as a concept guiding empirical research about the perception of life-action moving images

In her professorial dissertation “Der Leihkoerper” (“The surrogate body”) philosopher of media and the arts Christiane Voss argues that cine film, above all, is a medium of illusion that needs an aesthetics of illusion (Voss, 2013). For people from outside of the aesthetic field it shall be remarked here that, even when disregarding discussion of 4EA approaches to cognition and their rejection of the notion of perception as an illusion, this is a problematic or even provocative claim. At least since modern age the term illusion in discussions of art and aesthetics is highly discredited.[29] Illusion is effectively equated with delusion or better even deception. In the affectively involving and convincing presentation of the represented, the artistic representation has been accused of concealing itself as something actually being made – that is it makes itself transparent. The dreaded result of this self-negating process is unquestioned belief instead of critical consciousness of the political nature of every representation. To avoid this delusion, postmodern art rejects all modes of re-presenttations that aim at transcendence of the medium itself such as naturalism in visual arts or linearity of narration in literature. Instead it strives to show itself as the representation it is by stressing the media used relying on means such as leaving linearity and rationality, yes becoming as bulky and non-immersive as possible. Via this, Warburg scholar Sierek claims, the artwork represents a reflection demanding intervention rather than a self-disguising transparent realization. Exemplifying this dynamical approach to pictures on the works of artists like Rene Magritte or Oleg Kulik he writes: “[Diese Kuenstler verstehen (ergaenzt durch K.H.)] das Bildfenster eher als Eingriff denn als Einsicht und ersetzen Transparenz durch eine andere – uebrigens auch der visuellen Metaphorik entnommene – Eigenschaft: durch Reflexion.” (Sierek, 2007, p. 147). (“[These artists understand (complemented by K.H.)] the picture gate rather as an intervention than a simple insight and substitute transparency by another property – incidentally also stemming from visual metaphores: reflection.”]

Even the most postmodern cine films however, so Voss’ emphasis, only partly fulfill this promise. Of course, also when confronted with film, we don’t completely forget that what we see on the screen is not the reality surrounding us. Also in film, we “see the world in” a medium, at the same time being aware of the representation and the represented (see also

---

[29] Indeed, this criticism can already be found as early as in the texts of Platon. See for example Platon (2007).

Fingerhut's description of this phenomenon in Chapter II.2 (Fingerhut, 2012)). However, with reference to our own experience with cinema it seems hard to defend the claim that illusion or even delusion in fact should not be considered as the major practice of most films. Movies apparently explicitly aim at reaching out of the screen, that is carrying us off in other times, spaces and plots, with the absolute surrender of the spectator to the fictional world being the ultimate goal for some. A goal seemingly not so far from to be reached: It is a common experience to see spectators spellbound in front of the big screen, their gaze and body completely controlled by the stimuli of the fictional world only, while all the other that is "the real" world surrounding them seems neglected.[30] Furthermore, the influence of moving pictures obviously often goes beyond an engagement of our cognitive powers. Movies, even in their most artistic forms, unquestionably affect us bodily, they let us laugh out loud, cry uncontrollably and jump from our seats, they literally move or excite us and they might even frighten or disgust us so much, that we turn to the ultimate method of switching them off or leaving the room to escape their grasp. Most impressively, not even this last act might let us astray from their influence, as movies are also well known for their aftereffects. Not only can a movie affect me in a way that might change my mood and with this my world for the rest of the day; movies can even haunt us in our dreams.[31]

Naturally, this strong involving power of movies has equally fascinated audience as well as film scholars from the very beginning of film history – that is long before the invention of color and sound film. As early as in 1896[32] indeed, the Russian author and journalist Maxim Gorky wrote in the French press about the first movies seen

"Quelle que soit la scène ainsi prise et si grand que soit le nombre des personnage ainsi surpris dans les actes de leur vie, vous le revoyez avec toute l'illusion de la vie réelle. […] Vos nerfs se tendent, votre imagination vous etraîne dans une vie étrange, artificiallement uniforme, privée de couleurs et de sons, mais pleine de mouvement. […] C'est la vie même, c'est le mouvement pris sur le vif." (Gorky, [1896]; "No matter what scene is recorded, no matter how large the amount of characters caught in it in their daily actions, you can see it all again just like in real life. […] Your nerves get tense, your imagination leads you into a strange life,

---

[30] A very interesting paper about the actual irritation of the spectator happening in case that this focus of the screen is disturbed by the outer world forcing itself back into consciousness has been published by Julian Hanich (2013).

[31] A study of Eva Murzyn (2008) even indicated that movies might influence the general nature of our dreams by showing that the exposure to black and white movies correlated with the recall of greyscale dreams. See also Fingerhut (2014).

[32] Remarkably, this is not even one year after the first payed presentation of a movie: "La Sortie des usines Lumière" by Louis Lumière (firstly screened in front of a paying audience on the 8th of December at the Grand Café in Paris.)

artificially uniform, deprived of colours and sounds, but full of movement. […] It is life itself, it is movement caught live, then and there.", translation K.H.)

These words also underline that films manage to engage us in their fictional world in a way especially hard to explain by referring to symbolic processes and cognitive inferences. In his enthusiastic descriptions, Gorky is neither dwelling on cultural conventions nor our knowledge about how film footage is produced - that is the fact of film being a recording of "the real", by this still subliminally carrying the illusion-fostering reputation of a trace and an imprint rather than a depiction[33]. Rather, what he naturally puts his emphasis on is a movement taking over our nerves and body with this catapulting us in another here and now.

An intuition as such might be the reason, why film theory much earlier than theory of visual arts in general opened up for notions of film granting the spectator's body and worldly being a crucial role in the aesthetic experience involved. Of course cognitivist approaches to cinema can still be found in canonical works of scholars such as Goodman (1970[1945]), Bettetini (1973, p, 93-96), Gianetti (1976, p. 110-123) and Mast (1977, p. 177-180) who strictly describe movies as sequences of hundreds and thousands of separate images produced by a camera mechanically moving in a geometric space; images that then by a row of complex cognitive and emotional associations and transformations of our computationally functioning mind get connected and interpreted to finally form a narrative. Mostly, however, even in these descriptions, the immediate and physical impact of film on our bodies is granted at least a notice – if not a significant role. Just as in Chapter II.2 regarding embodied approaches to images, I will shortly sketch some important historical positions stressing the dependence of film perception on the perceiver's experiences off the screen.[34]

One of the earliest works to name here might be "The photoplay" by the psychologist Münsterberg, published as early as 1916. Though Münsterberg still quite cognitivistically suggested, that the "illusive" power of the cinema that make us see a world with depth and motion instead of sequences of flat pictures is due to the intensive "work of our mental mechanism" (Münsterberg, 1916, p.71) he also proposed, that the narrative structures of film function by letting us recall the way we normally perceive, think and imagine. In 1917, elaborating on the nature of this "recall", the doctors Toulouse and Mourgue (1917), investigating physiological responses such as breathing performance of film audience, suggested that during movie watching a phenomenon would take place, close to what happens to us

---

[33] For a discussion of this apparent indexicality especially regarding photography see Barthes, 1980 and Sontag, 1977.

[34] For an excellent review about the development of cinema studies see also Grieveson & Wasson (2008).

during hypnosis. Specifically, they stressed that this ‘hypnosis-like-effect’ would be possible via the “motor suggestibility” of the spectator, who, by the movie, could literally be put into the situation shown on the screen. A few years later in the 20s, the famous director Eisenstein who also published a number of essays about filmmaking recommended several techniques how to use camera and especially montage to elicit emotional responses in the spectator precisely by using, challenging or violating habits of every day perception (see Eisenstein, 2007). In 1935 then, Walter Benjamin described how the new medium of film builds on and at the same time further transforms contemporary perceptual habits (Benjamin, 1935) and in 1960 Siegfried Kracauer expressed his conviction, that it is in its mimetic play with real world perception, that film provides us with a phenomenology of this real world experience. Lastly, Maurice Merleau-Ponty, already in 1964 further explicated this by stressing that film experience can only be understood if we take it to be guided by real world perception, precisely meaning, that we do not perceive seperate images but rather a continuous “form in time” due to our overevaluating their movement as the action of an intentionally directed consiousness (Merleau-Ponty, 1964).

Especially this last position of Merleau Ponty was elaborated by what has been called “the ecological theory of film” with James J. Gibson being one of the founding fathers. The main argument of this position is precisely that film perception is not depending on acquired conventions, but based on nothing but the constitution of human vision. This vision is described precisely as found also in 4EA approaches to cognition: that is as the skilled bodily exploration of an environment already known from experience. As Gibson stresses, indeed there are no incompatibilities between film and reality obstructing the viewers ability to perceive a film in the first place. Rather, the changing visual field presented within a shot is similar enough to what we know from real world perception, to allow perceivers’ immersion in the scene via an activation of the whole body (see Gibson, 1979 as well as more recent positions elaborating on this such as Tan (1995), Anderson (1996), Platinga (1999, 2009), Grodal (1999, 2009), Rutherford (2003), Cutting (2004), and especially Sobchack (1982, 2004) and Voss (2011, 2013)). As for example Christiane Voss pointedly puts it
“It is the spectator’s body, in its mental and sensorial-affective resonance with the events on screen, which “loans” a three-dimensional body to the screen and thus flips the second dimension of the film event over into the third dimension of the sensing body. The spectator thus becomes a temporary “surrogate body” for the screen, and this body is, for its part, a constituent feature of the filmic architecture.” (Voss, 2011, p.146)

What we are talking about might in fact be best described as a potentiation of embodied world perception. Obviously, namely, the similarities between experiences in film and in real world are countered by salient differences, that is for example the lack of corresponding information via our olfactory, gustatory, haptic and proprioceptive senses. Nevertheless human cognition seems to overcome this gap via dwelling on its real experiences with the world. Just like we do not always have to really turn around the dice to know about the number in the back, and just like we do not need to wait until the end of an action started to be sure about its goal, film seems to be able to engage our whole way of being, hence being understood via experience rather than through symbolic interpretation only.

This does get even clearer when looking at how precisely camera movements and montage are described to contribute to this first bodily activation. It is especially in the writings of Voss and Sobchack that we get a chance to do so. Film, according to both of these authors, needs to be conceptualized as in its very fundaments operating with a suggestive stance because camera movements and continuity editing, via certain moving images they create, establish a continuous illusion in spectators' perception. Remarkably, the illusive nature referred to here consists in the fact, that the world experienced in a movie – in contrast to the world in our life off the screen – is indeed not out there "ready for our grasp". This fact however is suggested to be overwritten by the special activation of our embodied vision by the film, which is to a certain degree simulating this vision.

In this sense talking about camera movements, Vivian Sobchack stresses that it is the real bodily movement of the camera (remarkably put in contrast to artificial movements such as the zoom) that is "prereflectively understood as the incarnate movement of an animate body in an intersensory and lived unity within the world" (Sobchack (1982), p. 317) - the world with which the embodied mind interacts in meaningful-directedness. Exactly as much as our own physical movement in this world is experienced as serving nothing but this intentional interaction, Sobchack explains further, the movement of the camera is "disregarded in favor of our attending to that which the camera moves toward or away from" (Sobchack (1982), p. 318). Or also: "At this primary level of the body subject, the viewer intersubjectively and prereflectively recognizes and understands the camera as sharing the manner of his or her own existence, as manifestion of the material and kinetic code of an embodied and intentional consciousness." (Sobchack, 1982, p. 327).

And Voss, specifically referring to the contribution of montage in this interaction writes: “Diese kognitiv wie perzeptiv unaufloesbare Bewegungsillusion, die in dem Eindruck eines fliessenden Übergangs der Einzelbilder in der Zeit besteht, laesst sich als eine anthropologisierende Annaeherung der kinematographischen Visualität an die gängigen Bewegungswahrnehmungen verstehen, denen wir unsere Zuschreibungen von Lebendikeit entnehmen; denn den im vermeintlichen Fluss der Kinobilder zum Vorschein kommenden Handlungen des konstellierten Materials wird durch die an ihnen wahrgenommene Eigenbewegtheit der Charakter einer selbsttätigen Lebendigkeit automatisch verliehen. Dieser automatischen Zuordnung des Praedikats “Lebendigkeit” zu einer wahrnehmbaren Entität, die selbsttätig bewegt erscheint, entlehnen wir unserer Alltagspraxis und Erfahrung mit der empirischen Wirklichkeit.” (Voss, 2013, S. 40; “This cognitively as well as perceptually inextricable illusion of movement that consists in the impression of a fluent transition of the single shots in time can be understood as an anthropologizing assimilation of the cinematographic visuality to our usual habits of movement from which we derive our attribution of vitality; due to the perception of this movement, we automatically assign an intrinsic vitality to the actions appearing in the alleged continuity of shots. This automatic attribution of the predicate “vitality” to a perceivable entity that seems to be intrinsically moved, we derive from our daily practice and experience with the empirical reality.”; translation, K.H.)

In quintessence these descriptions indicate that film, in the first place, “works” by using special ways of camera handling and montage, creating images close enough to what we know from our own bodily explorations of the real world, in order to activate our whole mode of embodied beings exploring the world by moving our body and eyes in it. In an act of “signing over” thus, these moving images become the visions of an animate body actively moving through the world he inhabits, on which he depends and which he shapes. Taken seriously, this conceptualization of movie viewing describes film perception as an interaction between a (fictional) world and an embodied mind, an interaction not radically different from the interaction of the body and his environment in the real world.

This allows for a last hypothesis. Obviously, perception in the real world relies on stable patterns given by the lasting constitution of our body and the world. Any strong violation of such patterns will disturb a perceiver to a degree that might even let him doubt in himself or the world. For example, most of us will be familiar with the already deeply disturbing experience of getting up from the bed, but falling on the floor, because a leg had gone to sleep.

Even more so, we can all imagine, that if right now I would start seeing colors out of the normal visual spectrum, grow a third limb, or experience my friend Anna materializing in this room, without having used the door or even the window, I would probably scream out of shock and, even after a first recovery, ask myself if I might dream or have gone insane. If on the other hand my previously half blind and ever limping grandfather from New York would suddenly enter, walking upright and stable and without any stick or glasses straight through the door, I might just assume he finally got his new hip, a laser operation adjusting his eyes and had taken a plane to visit.

Which is to say that interactions between us and the world are open for transient changes – some of which not even depending on technology. Pregnant women happen to bump into doors with their bellies not being used to their "new shape", but an overweight person does not; patients that lost a limb often suffer from phantom pain, but mostly, after some time and body therapy, they recover. And finally, coming back to our habituation to future technologies involving new media, I am pretty sure that in a few decades I will completely expect Anna habitually materializing in my room – if not in person, than as a speaking and moving hologram substituting Skype video calls. Not to think of what Anna will be able to do in the movies those days.[35]

That however means that, while the base of spectator's embodiment in the fictional film world might be guaranteed by images close to experiences known from vision in real life, time and repetitive experience with this fictional world could integrate even effects initially violating perceptional habits with such force that they risk to interrupt film's illusion. Undoubtedly such an integration has already started, enabling movies to use faster and faster cutting rates and more experimental camera techniques year by year (see for example Cutting et al., 2010). Most likely these new perceptual habits have an effect also on our interaction with the world off the screen, that is in our dreams, in our imagination or even in reality.[36] At this point the importance of explicitly investigating film perception when trying to explore contemporary human cognition has hopefully been revealed in its full dimension.

---

[35] Indeed, the question how artworks do foresee what will happen to us in the far future is one of the points of discussions in the new field of posthuman aesthetics.

[36] First experimental proof for this, as mentioned, is given by Eva Murzyn (2008). In the conclusion of this thesis I will formulate some ideas for further investigations of this crucial point.

II.4 Hypotheses derivable from 4EA approaches to film for neuroscientific experiments

It is one of the declared main goals of this thesis to show how 4EA approaches to cognition as a new theoretical framework can influence empirical work on film perception in the way that they suggest hypotheses for experimental research that other conceptual backgrounds would not have yielded. This is to stress, that the experiments reported on in the next chapter are not designed to decide between the truth of cognitivist or constructivist notions of film and positions more influenced by 4EA approaches to cognition. Rather, they explicitly dwell on ideas developed under the elected influence of 4EA approaches to cognition suggesting that film experience, as we know it especially from Hollywood cinema explicitly aiming at completely involving the spectator in a fictionally world, is most likely not based on symbolic processing linked to learned conventions. Instead the chosen notion of film perception suggests that moving images, in their character similar to how we perceive world in the real, activate embodied vision per se, as a bodily interaction with a fictional world. Such an interaction crucially builds on our experience-based skill to not consciously miss information that is not always completely given to us via direct sensory input in all situations. Consequently it seems to be able to extend its normal habitual patterns to immediately perceive moving images as (motivated by) the vision of a(nother) animate body in a responding world.[37]
This conceptualization gives rise to the following hypotheses testable by experimental cognitive neuroscience

1) Detecting Violations of Visual Habits: New Stimulus or Violation of Action

According to the above view, and drawing back on Merleau Ponty's original phrasing[38], all moving images should be perceived as "forms in time", that is, the results of the perceptual observations of a directed consiousness continuously exploring a constantly accessible world. This includes moving images that, due to the application of special narrative devices of film actually exceed our visual capacities as the embodied beings we are, sometimes even to a degree that we explicitly note these deviations (see also Chapter III.1.1 and III.2.1 of this thesis). Therefore, abrupt visual changes happening in movies, such as caused by editing two

---

[37] Remarkably the brackets are added in this last sentence to stress that the camera not always has to be perceived as the eye of a person, that is "a fictional me", moving through the fictional world (though this might be an adequate description of the effect of some camera and montage work). Rather by partly simulating experience as we know it, film can dwell also on our intuitive understanding of body metaphors etc., as will be further illustrated in the next chapter.

[38] See Merleau-Ponty (1964) and Chapter II.3 of this thesis.

shots, should be detected directly as violations of the depicted action/scene rather than as the plain onset of a new stimulus.

As Chapter III.2 will elaborate, such distinctions (differentiating between a totally new stimulus and the perception of a slight violation of an expectation) can be traced via early ERP components directly following the change of shot. On the base of the above outlined it is thus predictable that cuts in comparison to plain onsets of visual stimuli would elicit early ERP components previously correlated with the detection of the violations of perceived actions rather than the onset of new visual stimuli.

2) Processing Violations of Visual Habits: Repair or Reflexion

According to the above presented view, furthermore, the interaction between spectator and the fictional world of a movie is substantiated by moving images closely resembling visual experiences in the real world. Due to this grounding, so the outlined argument, even violations of visual habits do not completely interrupt our interaction with the fictional world and thus our immersion in the film's story. Nevertheless we all know from experience, that while sometimes we do not grant a thought to the camera or the montage used to produce the moving image we currently perceive, at other times these narrative devices are actually drawn in our conscious awareness. We predict that this experiential difference is due to differences in postperceptual processing of detected violations, in their nature depending on the closeness of the moving image produced via the use of such narrative devices. Again, such processes can be traced by ERP analysis, specifically looking for differences in later phases of processing. We thus hypothesize to find differences in late ERP responses between edits producing images more or less closely resembling human vision in the real. Furthermore, these differences should explain the different subjective experiences connected to the perception of the respective stimuli.

3) Familiar Vision and Action Perception Links

Lastly, recent discussions about the role of embodied cognition in art perception have suggested that action perception links enable our body to be affected by a depicted and fictional world by providing access to an environment not completely open to us.

A mechanism that has been repeatedly claimed to enable that is the mirror mechanism as introduced in Chapter II.2 of this thesis. Precisely the mirror mechanism has been suggested to play a major role not only in simulation processes associated with action understanding (see Gallese & Sinigaglia (2010, 2011)) or action preparation in the real world (see Gallagher

(2001, 2007), Hutto (2013)) but also in the immediate feeling of presence of what is “seen in” an image (see Fingerhut, 2012). Combined with the already presented assumption, that spectators’ immersion in film is crucially substantiated by moving images closely resembling visual experiences known from real world interaction, this gives rise to the hypothesis that this substantiation is driven by a stronger activation of the mirror mechanism, measurable via respective motor cortex activations during action observation. We thus predict that the activation of the motor cortex during the observation of actions perceived via moving images is modulated by the similarity of these moving images with visual experiences produced by the encounter with the real world.

The next chapter will elaborate on these hypotheses while precisifying them for the different stimuli investigated, as well as presenting the results achieved in the experiments created to test them.

III. Experiments and Results

III.1. *Moving Mirrors* – A high density EEG study investigating the effect of camera movements on motor cortex activation during action observation

In her text about the semiotic structure of camera movement in cinema Vivian Sobchack talks about four basic kinds of movements defining motion pictures: the subject movement of living beings and objects within the projected image, the movement between projected images called editing, the optical or visual movement of the camera lens from a fixed position, such as the zoom, and finally the bodily movement of the camera itself, the camera movement. (Sobchack, 1982). The first empirical study presented in this paper was explicitly dedicated to the latter two kinds of movements. Specifically, it investigated how the application of different camera and lens movements in film affected spectators' brain activity when watching such stimuli. The analysis focused on differences in the activation of the motor system during action observation. Before reporting about experimental setup and results, the following section will clarify the precise hypotheses behind the investigation by more elaborately illustrating the idea of camera movements being a fundamental mean of the creation of a fictional world by activating and extending the spectator's body.

III.1.1 Introduction: Different camera movements, their narrative functions and their grounding in embodied perception

Literature on film making distinguishes around eleven basic ways of using a camera, five not involving a real displacement of the camera to another coordinate in space, and six others that involve such displacement (see for example Mercado, 2011; Bettman, 2013). To start with the first five: a scene can be recorded from a fixed position, mostly realized by help of a tripod holding the camera. This stable setup also allows what is commonly refered to by the terms "tilting" and "panning" of the camera. "Tilting" refers to moving the camera's lens up or down while keeping its horizontal axis constant. Remarkably, in filmmaking-guides this is often compared to the effect of "looking down" or "looking up", while only moving the eyeballs, not the whole head. "Panning", likewise, refers to moving the camera's lens to the right or to the left while keeping its vertical axis constant. Again, the effect of this movement is often illustrated by referring to "glancing" to your left or "glancing" to your right via sliding the eyeballs only. Stronger versions of these last two movements can be achieved via pedestal camera movements. That is, instead of moving the lens only, a tripod enables also to move the whole camera horizontally or vertically. A vertical pedestal movement is often compared with the effect of moving your chin to your chest to look down or raising your chin towards the sky to look up instead of only sliding the eyeballs up or down. A horizontal pedestal movement, on the other hand, is described as looking to the right or the left while turning the whole head instead of only sliding the eyeballs to one side. Lastly, lens technique allows to magnify the recorded scene without moving the camera in space. This is called "zooming". Remarkably, zooming cannot be illustrated by referring to a movement of the human body.

Furthermore there are six continuous camera movements comprising a real spatial displacement of the camera. To start with the first two, a camera can be moved by mounting it on a small wagon gliding on tracks either forward and backward, called dollying, or sideward, called trucking. Furthermore, there is the possibility of plain handheld shooting, that is the cameraman moving through the room while holding the camera in his hand. Since 1971, there exists also a technically stabilized version of this latter movement, achieved by use of a "floating cam" or "steadicam". These terms refer to a construction, originally invented by the camera man Garrett Brown, in its simplest versions consisting of a harness worn by the cameraman, an iso-elastic arm attached to this harness and a steadicam armature, with the camera mounted at one end and a counterbalance weight on the other, connected to the iso-elastic arm via a multiaxis and ultra-low friction gimbal. In effect, the camera can be directly carried by the cameraman through the room while avoiding rough and shaky pictures as produced by handmovements directly holding the camera. Lastly jibs, cranes or drones can be used to lift the camera off the ground and let it "fly", "swift" or "flow" through the room. Obviously, combinations of all of these techniques are possible and often applied.

In film, the reason for these application, at least in most cases, is of course not a technical demonstration. Rather, camera movements, being one of the two most traditional filmic narrative devices, serve distinct narrative functions. In his doctoral dissertation "Camera Movement in Narrative Cinema – a Taxonomy of Functions", Jacob Isak Nielson distinguishes six of these basic functions of camera movements (Nielson, 2008). The following section will briefly report about these while pointing out few selected examples illustrating how all of them can be considered as building upon a common underlying process - that is the use of camera movement to simulate and extend embodied being and perception.

The six basic functions of camera movement for film narration, according to Nielson, are orientation, pacing, inflection, focalization, reflexion and abstraction. Firstly, thus camera movement can orientate the viewer regarding space or action. One possibility for this is done by the camera articulating depth, that is rendering a 2D environment hard to interpret in a 3D environment as known from normal perception. To illustrate this, Nielson quotes filmmaker Vilos Zsigmond, cameraman and steadicam inventor Garrett Brown as well as director George Miller explaining the use of camera movement as a strategy in film-making:

“We decided to create a third dimension on a two-dimensional film not only with lighting and composition but also with constantly moving the camera. Since the objects change positions differently in different planes, the camera move reveals the right perspective of the objects in space, thus creating the missing third dimension.” (Zsigmond to the camera in a documentary by Carson, 2002)”

“When the camera begins to move, we are suddenly given the missing information as to shape and layout and size. The two-dimensional image acquires the illusion of three-dimensionality and we are carried across the divide of the screen, deeper and deeper into a world that is not contiguous to our own.” (Brown in his documentary about the Moving Camera part I, 2003)

“It’s a compulsion of mine to move the camera, and I now know why. It enhances three-dimensionality. It puts you in the space, and if you move the camera the audience becomes aware of the space. (Miller quoted in Bordwell & Thompson 2004, p. 269)”

Interestingly, these descriptions are reminiscent of conceptual reformulations of the basic mechanisms underlying depth perception within 4EA approaches to cognition. Indeed, understanding depth perception has kept the cognitive sciences busy, as attempts trying to model 3D perception via algorithms extracting depth information out of singular 2D pictures (as captured on the retina), faced serious computational challenges. With the upcoming of 4EA approaches to the problem it was suggested that such costly operations might be significantly simplified by the embodied mind using its ability to move in space to orientate and avoid ambiguous information. (O’Regan and Noë, 2001). A moving camera as described by the filmmakers above seems to simulate this normal strategy of orientation, hence simulating our normal embodied world consciousness, that is the awareness of a 3D space despite the restricted actual sensual information (see Chapter II.1).[39]

Something similar can be said with respect to a second way by means of which the camera can

---

[39] Obviously, there are also camera movements that do not produce images close to what we know from our normal perception. I will refer to this fact later.

orient the spectator, that is by directing spectator's attention. As argued by Nielson, the camera can stress the importance of an event, object or character by insistingly focusing on the essential cue, as it is for example done by long lateral tracking shots, following only one person. Again, to illustrate the principle behind this effect, Nielson adds a quote of cameraman Michael Chapman, precisely talking about this task: "The job of the cinematographer is to tell the audience where to look." (Chapman in a documentary by Glassman, 1992)
Tellingly, again, this citation describes the perception of the manipulation –the camera movement – as a simulation of active vision, rather than a series of images to be interpreted by cognitive strategies. Indeed, as the examples of Change Blindness in Chapter II.1 should have demonstrated, as bodily beings exploring a world "at our grasp" we are used to actually perceiving only a part of what there is to see as a product of our own selective behavior, guided by the objects of our current attention. The quote thus indicates, that a camera focusing on certain contents, in a way similar to what happens during normal vision, might exactly substantiate vision – that is the active exploration of a world known from experience by an embodied being.

Now, it obviously is one of the intriguing features of film, that it is not only able to imitate our normal perception. Rather it seems to have the crucial capacity to extend our vision – and with that our bodily being. To begin with, camera movement, according to Nielson, can also minimize cues of depth and volume – for instance by swish-panning or otherwise making a blur or smear out of the image – in this way creating insecureness, disorientation or feelings of estrangeness in the spectator. Also, it might be used to distract our attention from cues actually relevant to the narration, such as a character's actual identity etc. Both effects, so the rather intuitive suggestion of Nielson, can play a crucial role in creating arcs of suspenses, as known from e.g. criminal stories.

Indeed, I would say that the other five functions of camera movements that Nielson describes, can equally be presented as cases of such apparent fluctuations between simulation and deviations from orientation in the normal world. Nielson for example, talking about "pacing", refers to the "the determination of the velocity at which visual information transpires" by means of the moving camera (Nielson, 2008, p. 230). Clearly, again, this velocity, achieved by the camera movement, can correspond to "human vision in the wild", that is the flow of images during perception in the real world, varying according to our situation. Vision differs during a relaxed Sunday walk and a rushed flight or fierce fight. Nielson assigns to the category “sustained pace” films imitating these different kind of flows and thereby most likely helping us in feeling immersed in the presented scene. Moreover however, directors often use the technical means of today to accelerate or retard the camera movement to a degree that clearly supersizes the capacities of our eyes and bodies. As Nielson seems to suggest, this possibly enables the experience of other kinds of visions and bodies: The camera in “Amour” of Haneke (2012), moving less and less the further the physical and psychological decay of the main actress goes and sometimes getting stopped in weird positions not allowing to visually follow the auditorily ongoing scene, literally puts us in the stuck body of a dying person. On the other hand, a fast camera drone flight over a vast area, might let us directly perceive the bodily and thus perceptual allmightyness of a superhero. Under the term “inflection” and “focalization”, Nielson summarizes similar functions of the camera, this time imbuing the scene by manipulating the mode of its movement rather than its velocity. While illustrating “inflective” uses, Nielson reports how camera movements might crucially enhance physical appearance and behavior of some represented characters by for example ”following a stout elderly man by means of robust, weighty and slow dolly movements, or alternately by following a nimble gazelle by means of a weightlessly drifting steadicam camera” (Nielson, 2008). In cases of “focalizative” uses, such effects are not applied to further characterize the figures or object in the frame, but rather to demonstrate a certain subjective viewpoint, that is the vision of one of the characters on their world. An example of such a camera, granting us a look from within a protagonist on the world as perceived by him, can be given by the movie “Dallas Buyers Club” by Vallée (2013), which tells the story of the life and suffering of an HIV patient. One of the marked features of this film is the use of a steadicam, which indeed often manages to make the spectator see through the eyes of the protagonist. Very impressively, via the unexpected sagging, trundling or spinning of this camera, the spectator is able to even anticipate the protagonist fainting by being granted a direct experience of the actual physical syndromes preceding such a breakdown. Nielson also states that it is via such means that cinema has the capacity to let us experience even a glimpse through another “mind’s eye”, that is not the perceptual experiences of another person or creature in real life, but also the dreams, imaginations or hallucinations of somebody else. Precisely he writes: “*Der letzte Mann* (Murnau, 1924) contains an impressive long take where the porter (Emil Jannings) in a drunken stupor imagines himself walking into Hotel Atlantic carrying a huge suitcase and tossing it up in the air as if it were light as a feather. These events are presented in distorted images and by means of a camera mounted on Karl Freund’s chest following Jannings’ character through the revolving door” (Nielson, 2008, p.244).

Similar descriptions, stressing camera movement simulating and deviating from the spectator's experiences with the real world, could be made for the functions of reflective and abstract uses of camera movements. This however goes beyond the scope of this thesis.[40]

The examples given should be enough though to illustrate a basic claim. This is that camera movements, via imitating experiences from normal vision, are able to immerse the spectator's body in a fictional environment. On this base then, the same body can be extended by deviations from this embodied vision in the real world. First engaged in an interaction with the fictional world, the spectator in his body is able to explore another kind of being.

III.1.2.1 Experimental idea and hypothesis

It is the main hypothesis of the following experiment, that this special relation between spectator and movie should be reflected in the brain activity of the spectator while watching movies. Specifically, the hypothesis is, that despite the possibility of different kinds of visions due to film, our body should react stronger to moving images that better resemble normal visual experiences. This should be especially the case with regard to activations possibly serving the understanding of or the response to what can not be (better) perceived by physical contact or movement – such as a film. Such activations, so goes the further hypothesis, might in fact help the spectator's immersion in the movie, thus most probably correlating with the feeling of being involved in the scene.

The experiment reported on in the following sections thus specifically tested if movie clips that more closely resemble usual vision in real life would stronger activate action perception links suggested to serve the understanding of and interaction with the real world, seen from a distance, and furthermore, if such activations correlate to a subjective experience of

---

[40] A very interesting case though is described for the abstract use of camera movement, that is the representation of abstract ideas by this narrative device. As an example, Nielson refers to a famous scene out of Wim Wender's movie "Der Himmel ueber Berlin" in which the transfer of a soul from a dying body to the realm of god by the help of a sent angel ("Damiel") seems to be understood via a pendulous movement of the camera. Nielson quotes the director himself stating:
"I thought that going back and forth sort of showed more what Damiel was actually doing, in the way that he is - as the man is dying - that he's taking him over, so to speak. In a way, in a strange way, this "action" as well as the pain that came with it, were in that camera movement. [...] as it was somehow about a transition between life and death, it did translate something: not so much his p.o.v., more a mental attitude. Damiel's tenderness and his care for the man were in that back-and-forth movement" (Wenders in Raskin, 1999, p 7; see also Nielson, 2008, p. 258). This example explicitly stresses the impact of bodily dimensions even on something that might best be described as a "filmic metaphor", reminding of Lakoff and Johnsons work "Metaphors we live by" (see Lakoff & Johnson, 1980) in which they emphasize the crucial influence of our bodies on something as much depending on convention as language.

involvement in the scene. Stimuli were clips of the same scene filmed once with a fixed camera, once while applying a zoom, once while approaching the scene with the camera on a dolly track and once while approaching the scene with the camera carried by a camera man and stabilized by a steadicam construction. Importantly, these different uses of the camera can be expected to correspond to different degrees of similarities of the produced moving images with perception as known from real life. Exactly, while we can hold the head still like a camera on a tripod or walk through the room like a steadicam fixed on the body of a cameraman, we more seldomly move on tracks (although, sometimes we might do, when watching out of the head window of a train), and we always need technical support to "zoom in" on things.

As the further paragraphs will describe in detail, the experiment consisted of a high density EEG recording, registering brain activity of the spectators during the watching of the different movie clips. Furthermore, we added a rating task to assess spectators' conscious experiences with the movies. Our hypothesis was that videos produced by camera movements most likely to elicit experiences close to spectators' vision in real life, should not only be judged as the most naturalistic and involving, but also most strongly activate what has previously been described as the mirror mechanism – an action-perception link likely to serve understanding of and response preparation to actions of others as well as our feeling of presence of the surrounding world (see also Chapter II.4) .

An established marker of the mirror neuron-related activity in the human brain is the desynchronization of the "mu rhythm" in EEG (see for instance Muthukumaraswamy, Johnson, & McNair, 2004; Pineda, 2005; Ramachandran & Oberman, 2006). The term "mu-rhythm" refers to an EEG rhythm with one component in the alpha band (8–14 Hz) and one component in the lower beta band (14-20 Hz), recorded from central electrodes. This rhythm has been shown to desynchronize not only during the execution of active movements, but also during their observation (Cochin et al., 1998, 1999; Altshuler, Vaknov, Wang, Ramachandran, & Pineda, 1997; Altschuler, Vankov, Hubbard, Roberts, […], & Pineda, 2000); Hari et al., 1998, Derambure, Defebvre, Dujardin, Bourriez, Jacquesson, Destee, Guieu, 1993; Leocani, Toro, Manganotti, Zhuang, & Hallet, 1997; Pfurtscheller & Lopes Da Silva, 1999; Pfurtscheller & Aranibar, 1979; Pfurtscheller & Berghold, 1989; Stancak and Pfurtscheller, 1996; Toro, Deutschl, Thatcher, Sato, […], & Hallet, 1994). It has therefore been suggested that its suppression (also termed event-related desynchronization (ERD)) is due to a neurophysiological mechanism of motor resonance, the mirror mechanism, caused by similar neurons to the ones originally found in the maquaque's brain (Muthukumaraswamy & Johnson, 2004; Pineda, 2005; Ramachandran & Oberman, 2006). Our precise hypothesis therefore consisted in the prediction that event related desynchronization (ERD) of the mu-rhythm would be modulated by the degree of similarity with vision in the real world achieved by the camera movement used for producing the movie watched by participants.

The following paragraphs will report the details of setup, analysis and results of this experiment.

III.1.2.2 Participants, materials and methods

19 healthy volunteers, recruited by public announcement, participated in the experiment. Two participants were subsequently excluded from analysis due to artifacts in the EEG data or lack of typical ERD pattern (see Section EEG recording and analysis). Among the remaining participants, 7 were males, 10 females, the mean age was 22.8, and all were right handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). All participants were paid 25 euros as reimbursement. Before the experiment, they received written and oral experimental instructions. After the experiment each participant was debriefed. Written informed consent was obtained from all participants before entering the study. The study was approved by the local Ethical Committee.

Stimuli consisted of short video clips of 3 seconds length showing an agent (one female, one male) grasping an object from a table placed in front of the agent. The background consisted of a black wall in front of which we placed a panel with a black and white geometrical pattern to enforce the 3D perception of the room. Grasped objects included a marble, a battery, an eraser, a roll of scotch tape, an espresso cup, a plastic mug, a plastic ball of 8 cm radius, and a packet of tissues (See **Figure 1,** next page).

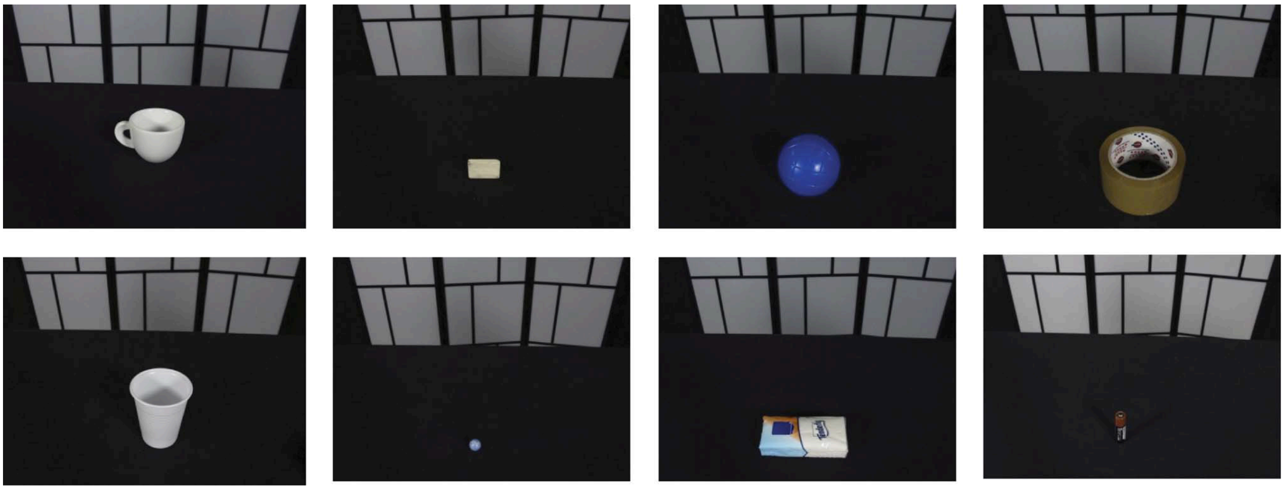


Figure 1. Objects presented in the video clips and in the attentional task. Figure taken from Heimann et al., 2014.

Video clips were recorded in a professional film studio, enabling us to film the same scene 4 times under highly controlled conditions. The camera starting position was always 260 cm far from the object, the end position (in case of movement) was 80 cm from the object. The camera movement speed as well as its height from the ground were kept constant across the three different movement conditions, so that the only difference among them consisted in the type of movement and approach to the scene: In 25% of the videos, the camera stayed fixed on a tripod, 260 cm away from the object. In other 25%, a zoom was applied, resulting in a magnification of the object corresponding to its presentation in a distance of 80 cm. In 25%, the camera during recording was moved towards the scene by means of an automatic transport on dolly tracks (programmed, steady pace) stopping in a distance of 80 cm from the object. And finally, in the last 25%, the camera was carried towards the scene by means of a steadicam construction carried by a professional camera man. **Figure 2** (on the right) shows 4 still frames taken from a video clip filmed with such a steadicam, including start and end positions of the camera.

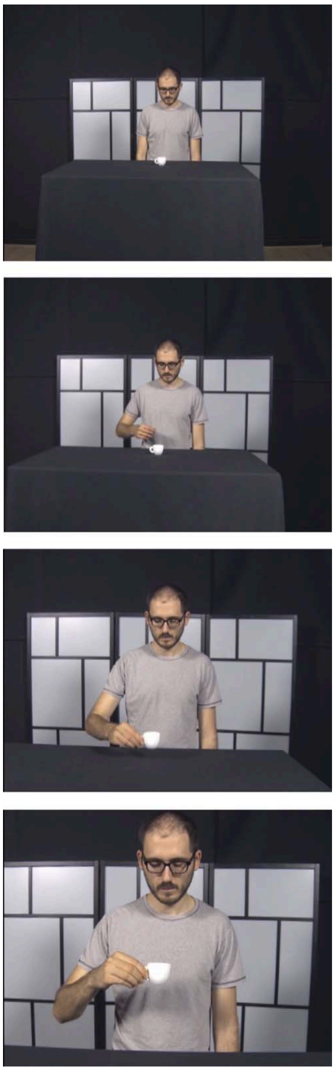


Figure 2. Single frames extracted from one video clip in which the camera (steadicam) approaches the agent. Figure taken from Heimann et al., 2014.

The experiment consisted of two different parts comprising:
1) a 50 minutes EEG recording session;
2) a 10 minutes rating task

1) EEG

The EEG was recorded during five blocks of about 10 minutes length each. After each block participants were given the possibility of a short break. Each block consisted of 80 trials (16 per condition – still, zoom, dolly, steadicam and action execution, for details see below). Participants were seated in an isolated EEG lab in front of a computer screen placed on a table at a distance of 50 cm. Participants were then instructed about the experimental procedure: Each trial began with a fixation cross of 200ms, followed by a video (of 3 sec. length each, presented in random order). In 80% of the trials, after stimulus presentation a grey screen was displayed for 5 seconds (guaranteeing the return of brain activity to baseline). Participants were asked to blink only in the second half of the grey screen period to minimize movement artifacts also in the resynchronization phase. In the remaining 20% of the trials (action execution condition), after stimulus presentation and before the grey screen, in addition, the photo of an object (see above) appeared, and participants were asked to tell whether the object was the same they had seen being grasped in the video displayed just before. The answer had to be given by clicking the mouse with their right index finger. The mouse was positioned on the table in front of participants at a distance of 15 cm from their right hand. When the object photo appeared, participants were asked to move their hand from start position to the mouse and click either the left button of the mouse, when seeing the same object as in the video just displayed before, or the right button of the mouse, when seeing a different object. They were furthermore asked to move back the hand to start position right after the click. If participants gave a wrong answer or did not answer within 3 seconds, they were told the trial was incorrect or the answer given too slow, and the trial was repeated. This action execution condition served both as control for attention and to record participants' ERD during action execution (for experimental paradigm, see **Figure 3** next page).

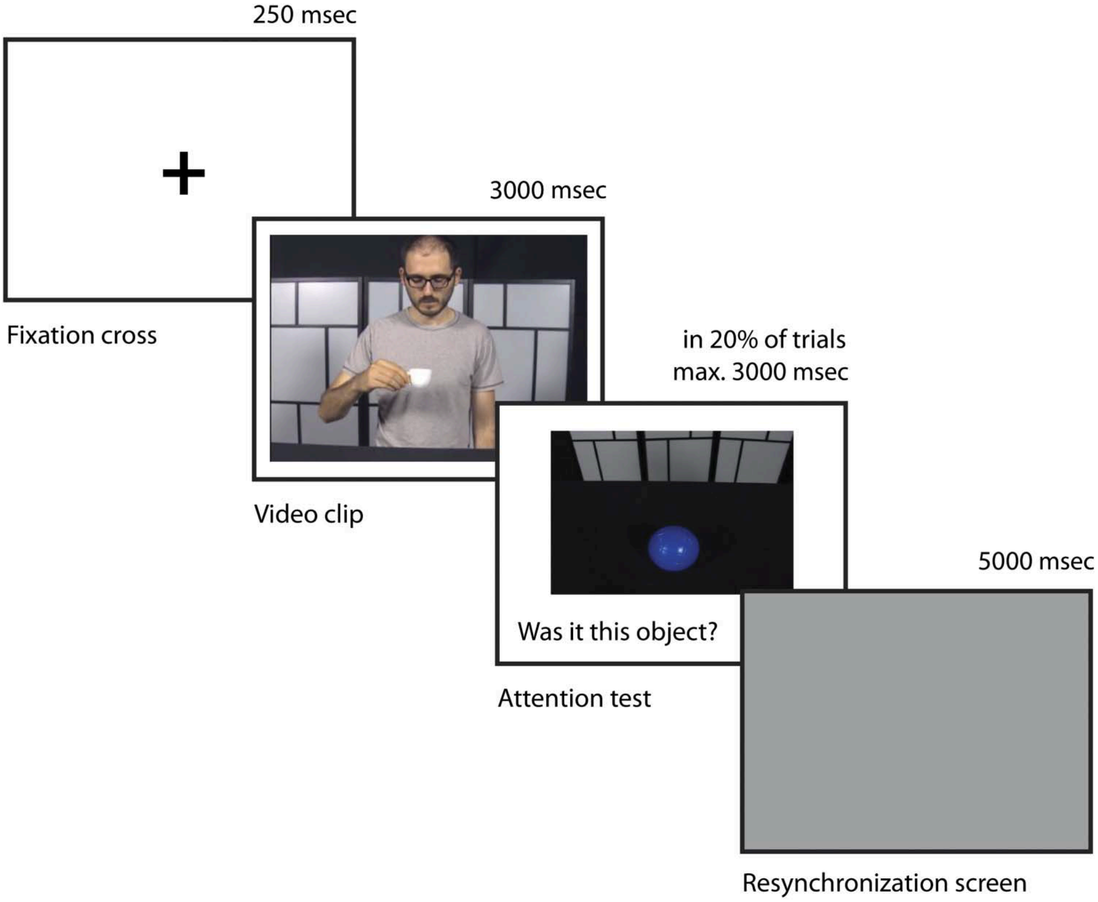


Figure 3. Experimental paradigm employed during EEG recording. Figure taken from Heimann et al., 2014.

2) Rating task

The second part of the experiment consisted of a Rating task. For this task, out of the previous 64 video clips, twelve were chosen again (three per condition). Participants, still sitting in front of the screen as during the EEG recording session, were again shown these video clips and for each were asked six different questions (in 6 separate blocks (one for each question) always conducted in the same order):

1) How much did you feel involved in the scene?

2) How much did you feel like the actor?

3) How much did you feel as if you yourself would approach the scene?[41]

4) How comfortable did you feel watching the scene?

5) How realistic did you find the camera movement?

6) How much did you feel the camera movement resembled a person's movement when approaching the scene?

---

[41] Questions 3, 5 and 6 were not asked for still camera video clips.

The first three questions were designed to investigate participants' potential feeling of involvement with the observed scene either in terms of empathy with the actor or in terms of embodiment with the camera. The fourth question was designed to explore how at ease participants were with the different ways of filming the scene. The last two questions were designed to measure participants' estimation of the ecological plausibility of the different types of camera movements with respect to those of a real observer approaching the scene. The rating was given by using the mouse to place a cursor on a 0-100 rating scale below each stimulus.

EEG data was acquired by a 128-channel Sensor Net (Electrical Geodesic, Eugene, USA) and recorded within standard EGI package Net Station 4.3.1. EEG was sampled at 250 Hz and band-pass filtered at 0.3–100 Hz, electrodes impedance was kept less than 50 kΩ (controlled after each block). The raw EEG data was recorded with the vertex (Cz) as the online reference and re-referenced off-line to the common average (Muthukumaraswamy, Johnson, & McNair, 2004). Stimuli were presented with E-Prime 2.0, and at the beginning of each trial, all event markers were sent to Net Station. Participants' motion was monitored by the experimenter and video-recorded for off-line analysis; if participants moved during the observation or rest conditions, the trial was excluded from further data analysis. EEG data were filtered off-line with band-pass filter 0.3–30 Hz and segmented into specific time epochs. From observation trials the whole three seconds of stimulus presentation plus the first two seconds of grey screen (resynchronization phase, see below) was analyzed. As baseline we used 1000 ms of grey screen ending 1 second before the start of the new trial (the appearance of the fixation cross) in the observation trials. From the action execution trials, segments of 1000 ms were cut, starting 500 ms before the motor response (button press) and ending 500 ms after it. Only the trials in which participants responded correctly were analyzed. The trials in which participants produced eye-blinks and movement artifacts were rejected on the basis of the artifacts detection tool supplied by Net Station and on the basis of a subsequent careful visual inspection of each segment. A minimum number of 50 trials for each condition was kept (fulfilled by all but one participant, who was consequently excluded from further analysis).

The time–frequency analysis was performed by continuous Morlet wavelet transformation in 0.5 Hz intervals in the frequency range from 1 to 30 Hz. Frequency-power coefficients were calculated by taking the average across trials. The wavelet transformation was calculated separately for each participant in all 128 channels for each condition. It was scaled by division using 250ms out of the preceding fixation cross period.

Statistical analysis was performed on a selected cluster of 8 electrodes in each hemisphere located around standard C3 and C4 sites (Electrodes 30, 31, 36, 37, 41, 42, 53, 54 left and 79, 80, 86, 87, 93, 103, 104, 105 right, as used in prior studies, see Streltsowa et al., 2010; Muthukumaraswamy & Johnson 2004; Muthukumaraswamy et al. 2004 ; Bernier, Dawson, Webb, & Murias, 2007).

Data were analyzed with regard to frequency power changes of the different components of the rolandic mu-rhythm, reported to be modulated by voluntary action execution as well as action observation (Derambure et. al., 1993; Leocani et al., 1997; Pfurtscheller & Aranibar, 1979; Pfurtscheller & Berghold, 1989; Stancak & Pfurtscheller, 1996; Toro et al., 1994; Pfurtscheller & Lopez da Silva, 1999). As it was shown that also the ERS following ERD of the rolandic mu-rhythm shows a specific pattern that can be modulated by contextual conditions (Muthukumaraswamy et al., 2004; Pfurtscheller, Neuper, Brunner, & Lopez Da Silva, 2005, Heimann et al., 2013), we analyzed four different consequent time windows representing early ERD, beginning of ERS, rebound (over-synchronization) and return-to-baseline stages (see statistical analysis section for details).

For each participant, specific alpha-frequency bands were selected in the range of 8–14 Hz following the procedure described in previous studies (Oberman et al., 2007a,b; Babiloni et al., 2009). The individual peak (F) of attenuated frequency was determined by calculating the ratio between the frequency power in action execution trials and during baseline in the six following sub-frequency bands: 8–9, 9–10, 10–11, 11–12, 12–13, 13–14 Hz. Each value was then transformed into a log-ratio, and the frequency that corresponded to the log-ratio with the most negative value was taken as F. A 3 Hz range frequency band was chosen for each participant (F - 1; F + 1). For the following statistical analyses, the frequency power in this 3 Hz range was extracted in all conditions (number of participants selected per range: 7-10 Hz: 0; 8-11 Hz: 6; 9-12 Hz: 2; 10-13 Hz: 8; 11-14 Hz: 1; 12-15 Hz:0). The data of a second participant for which no significant difference between baseline and action execution condition was found in any of the checked frequency bands, was excluded from further analysis.

Since the central alpha frequency band (8–14 Hz) overlaps with the posterior alpha band, it is possible that recordings in central areas might be affected by this posterior activity. In order to check whether the alpha recorded in central areas was affected by posterior alpha, for the alpha-range selected for each participant we performed an additional analysis in 4 electrodes per hemisphere in occipital areas (electrodes 69, 70, 73, 74 in left occipital lobe & electrodes 82, 83, 88, 89 in right occipital lobe) using the same frequency bands as previously described.

Furthermore, in every participant three beta-frequency ranges were analyzed (a lower band of 14-20 Hz, a middle band of 18-24 Hz and an upper band of 24-30 Hz) using the same central electrode-cluster as for the alpha-range (regarding range selection see Avanzini, Fabbri-Destro, Dalla Volta, Daprati, […], & Cantalupo, G., 2012).
On the results of the statistical extraction we applied an outlier detection (+/- 2 standard deviations from mean) revealing outlier values in zoom condition of one participant. Such outlier values were replaced with the corresponding average values of the remaining participants.

In order to assess central alpha and beta desynchronization in sensory-motor areas during different observation and action execution trials, we compared the frequency power extracted from wavelet transformation for the different conditions using several ANOVAs.

1) In order to generally assess central alpha and beta ERD in sensory-motor areas, for every band range considered (alpha, low beta, middle beta, high beta) we compared the frequency power extracted from wavelet transformation during baseline (average of 1000ms of period of grey screen ending 1000ms before fixation cross) with its value during observation conditions (average of 3000ms of video presentation) and action execution condition (average of 1000ms: 500ms before and 500ms after button press) using a repeated measures 2x6 ANOVA with two levels of Hemisphere (right vs. left) and six levels of Condition (baseline, four observation conditions (still, zoom, dolly and steadicam) and action execution).
2) In order to assess the time course of central alpha and beta ERD/ERS in sensory-motor areas during the four observation conditions, for every band range considered (alpha and low beta – middle beta and high beta were left out, due to results of analysis I, see below) we firstly analyzed frequency power in 20 separate epochs of 250ms length each. Due to the typical ERD/ERS/rebound pattern seen in the descriptive analysis of this analysis, we then chose 4 time windows to assess statistical differences among conditions in the different stages of the event-related modulation of the mu rhythm. The selected four windows were: ERD window, consisting of the first second of stimulus presentation; ERS window, consisting of the second and third second of stimulus presentation; Rebound window, consisting of the first second after stimulus offset; Return to baseline window, consisting of the second second after stimulus offset. With the values of the power analysis for these 4 windows we performed a repeated measures 2 x 4 x 4 ANOVA with 2 levels of Hemisphere (right vs. left), 4 levels

of Condition (still, zoom, dolly, steadicam), and 4 levels of Time (see above). To keep the relation to baseline in the picture, values used for this ANOVA were the log values of the condition/baseline division.

3) To control for effects in occipital electrodes for the alpha band, both ANOVAs described above were repeated for the occipital electrodes in the same frequency ranges.

Results of the Rating task were analyzed using a Repeated-measures-ANOVA for each question with the single main factor of Condition.
In all performed ANOVAs (of EEG analysis and Rating task) we applied Duncan post-hoc tests to further explore significant factors and interactions. Reported results are automatically adjusted for multiple comparisons. Error bars in all the graphs represent standard errors. In figure 6 and 7, conditions are represented next to each other for the ease of visualization.

III.1.4 Results

III.1.4.1 Results - EEG

In order to generally assess central *alpha* ERD in sensory-motor areas, we compared the frequency power for selected alpha frequency ranges extracted from wavelet during baseline with observation conditions (still, zoom, dolly, steadicam) and action execution condition. Descriptive analysis (see **Figure 4**) showed that, compared to baseline, ERD was present in all four observation conditions as well as during action execution (button press), with a maximum in the latter condition. A 2x6 ANOVA (Hemisphere x Condition) showed only a significant main effect for Condition (F(5,80)=17.882, p<0.001).

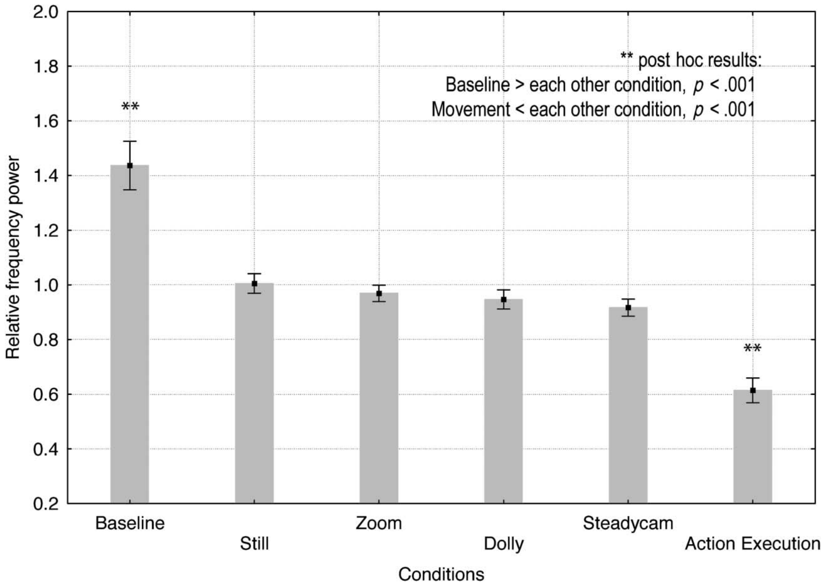


Figure 4. Central alpha frequency power during baseline, action observation, and action execution.
Condition effect: F(5, 80) = 17.882, p<0.001. Figure taken from Heimann et al., 2014.

Post-hoc comparisons showed that frequency power for baseline was significantly higher than for all other conditions and that frequency power for action execution was significantly lower than for all other conditions (for all these comparisons p<0.001). Differences between different observation-conditions were not significant (for all p>0.3).

To control for similar effects in occipital regions we repeated the analysis just described in occipital electrodes. The 2x6 ANOVA with 2 factors of Hemisphere and 6 factors of Condition (baseline, still, zoom, dolly, steadicam and action execution) in the occipital region showed a main effect of Condition (F(5,80)=21.08, p<0.001). Post-hoc comparisons showed that frequency power for baseline was significantly higher than for all other conditions (p<0.001), while there was no significant difference among observation conditions and action execution condition. These results support the notion that alpha ERD in occipital regions is discriminable from alpha ERD in central regions.

The ANOVA also showed a significant effect of Hemisphere x Condition (F(5,80)=3.3, p<0.01). Post-hoc comparisons showed that this difference was due to differences of the baseline values only, with the right hemisphere having a higher baseline value than the left hemisphere (p<0.01). We therefore computed the lateralization index (Baseline right – Baseline left)/ (Mean of Baseline right and left) for all participants in central and occipital electrodes for this condition and applied a paired samples t-test to check for differences between the two regions. Descriptive analysis showed that the mean lateralization index of baseline values in central electrodes was 0.082, while the mean lateralization index of Baseline values in occipital electrodes was 0.096 (indicating a slightly higher lateralization to the right hemisphere in Occipitals). However, the paired sampled t-test did not show a significant difference between the two regions (t(16)=-0.240921, p > 0.8). This means that our data cannot support a hemispheric difference between the two regions of interest in the occipital cortex regarding the mean power frequency over the whole time of video observation.

In order to generally assess *beta* ERD in central sensory-motor areas (as part of the mu-rhythm), we compared the frequency power of three different beta-ranges (low, middle and high beta) extracted from wavelet transformation during baseline with frequency power during observation conditions and during action execution (see **Figure 5**, next page).

Descriptive analysis for the low beta range (14-20 Hz, see **Figure 5a**, next page) showed that compared to baseline ERD was present in all four observation conditions as well as during action execution, with a maximum in the steadicam condition. A 2x6 ANOVA (Hemisphere x Condition) accordingly showed a significant main effect for Condition. Post-hoc comparisons revealed that the significant differences occurred between baseline and all other conditions ($p<0.001$), as well as between steadicam and action execution ($p<0.01$).

Descriptive analysis for the middle beta range (18-24 Hz, see **Figure 5b**, next page) as well as for the high beta frequency range (24-30 Hz, **see Figure 5c**, next page) showed no ERD for observation or action execution and the 2x6 ANOVAs (Hemisphere x Condition) showed no significant main effect for Condition or Hemisphere x Condition interaction. The results are contrasted to each other by the following figure (**Figure 5**).

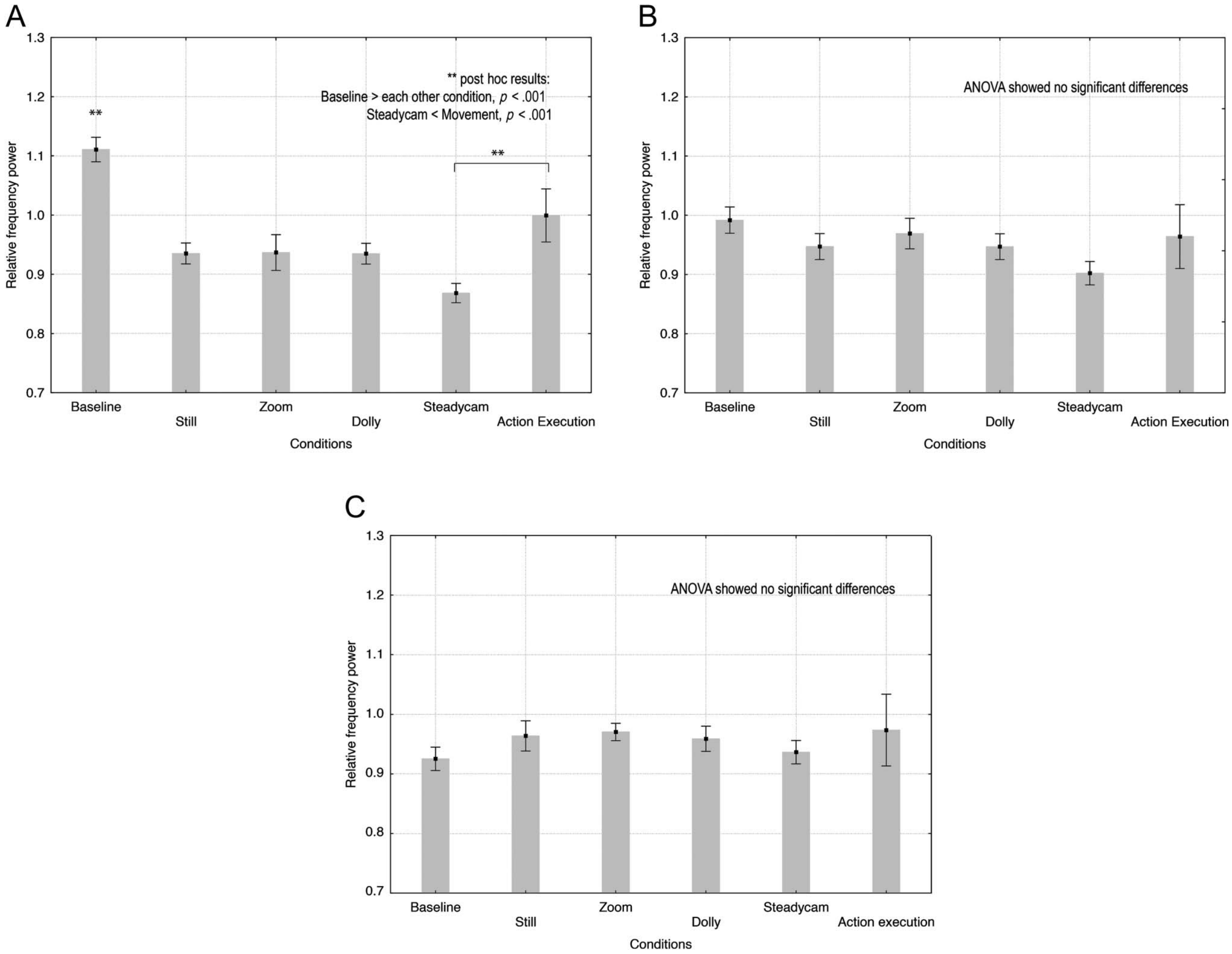


Figure 5. (A) Central lower beta frequency power during baseline, action observation and action execution. Condition effect: F(5, 80) = 7.0747,p <0.001. (B) Central middle beta frequency power during baseline, action observation, and action execution. Condition effect: F(5, 80) =.93231, p = .46468. (C) Central high beta frequencypower during baseline, action observation, and action execution. Condition effect: F(5, 80) = 0.453, p = .809618. Figure taken from Heimann et al., 2014.

Due to these results, we decided to further investigate only the two ranges that gave significant results: the selected alpha frequency range, as well as the low beta frequency range.

In order to assess the precise time-course of ERD/ERS in sensory-motor areas during the four different observation conditions in the frequency ranges of interest, we compared the log values of the frequency power of the chosen frequency ranges extracted from wavelet transformation during the different observation conditions divided by baseline. For each frequency range, a 2x4x4 ANOVA design was created with 2 levels of Hemisphere (left versus right), 4 levels of Condition (still, zoom, dolly, steadicam) and 4 levels of Time (4 time windows representing early ERD, beginning of ERS, rebound (over-synchronization) and return-to-baseline stages). For a better illustration of the time course, in the graphs we show the 20 epochs separately, marking the selected windows on the bottom.

Alpha range

The results of the 2x4x4 ANOVA (Hemisphere x Condition x Timewindow) for the selected alpha frequency ranges revealed a significant main effect of Time ($F(3,48)=20.1$ $p<0.001$) as well as significant interactions of Condition x Time ($F(9,144)=1.97$, $p<0.05$). Descriptive analysis regarding the main effect of Time showed a typical ERD-ERS-rebound pattern (see also Avanzini et al., 2012). Descriptive and Post-hoc analysis regarding the interaction of Condition x Time showed the following characteristics (see **Figure 6**, next page). 1) In the ERD phase (first second of video observation) in the selected alpha range, no significant differences among conditions were observed. 2) In the following time window of further ERD/beginning ERS (second and third second of video observation) conditions differentiated. Descriptive analysis showed that ERD was strongest for the steadicam, followed by dolly, then zoom, then still condition. Post-hoc comparisons showed that differences were significant between still and steadicam condition ($p<0.001$), zoom and steadicam condition ($p<0.05$), and between still and dolly condition ($p<0.05$). 3) In the rebound window (first second after stimulus offset) order of power values changed with descriptive analysis showing that zoom had the highest power, followed by dolly, then still, then steadicam. Post-hoc comparisons showed that differences were significant between zoom and still ($p<0.01$) and zoom and steadicam ($p<0.05$). 4) In the returned-to-baseline window (second second after stimulus offset) no significant differences between conditions were measured.

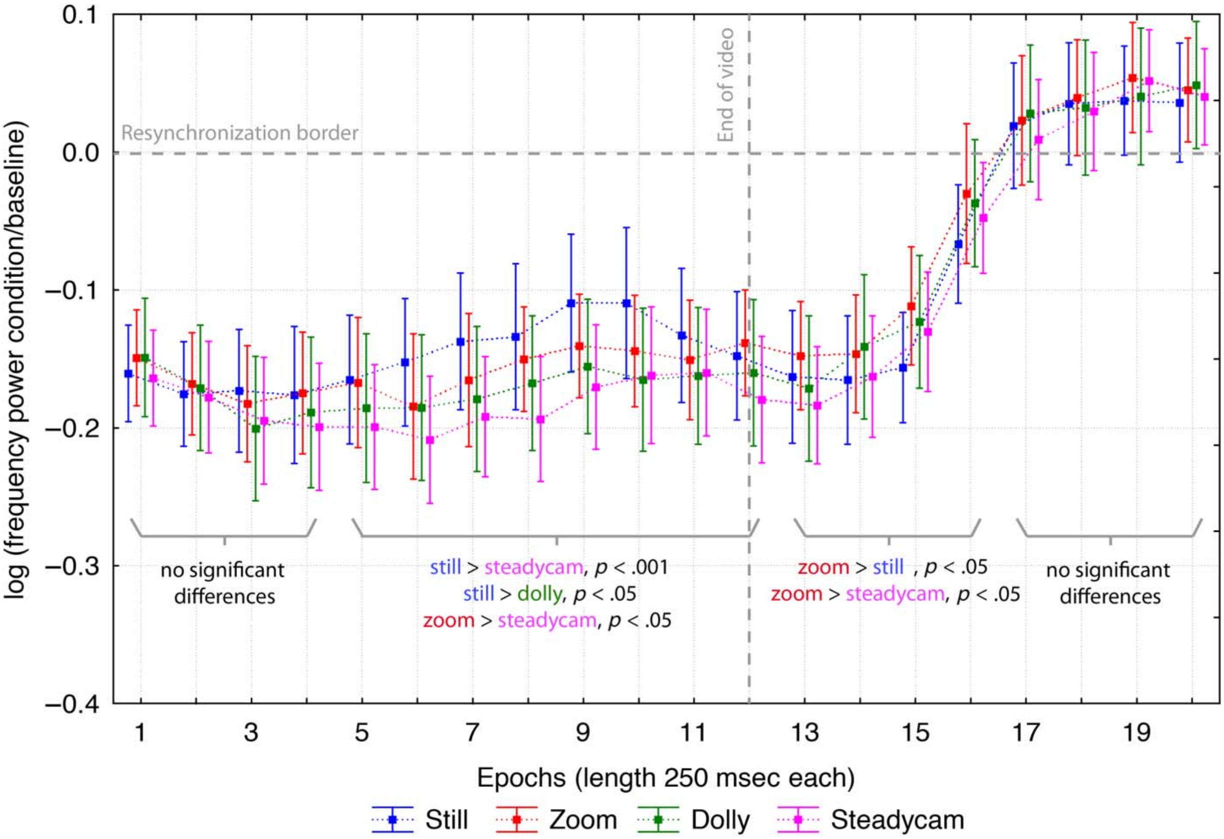


Figure 6. Central alpha frequency power over 20 epochs. Statistics over four time windows as indicated. Condition x Time interaction, **F**(9, 144) = 1.97, p < .05. Figure taken from Heimann et al., 2014.

To control for effects in occipital regions we repeated the analysis done in central electrodes for occipital electrodes. The 2x4x4 ANOVA (2 levels of Hemisphere (left vs right), 4 levels of Condition (still, zoom, dolly, steadicam) and 4 levels of Time (see above) showed a main effect of Hemisphere (F(1,16)=4.74, p<0.05), a main effect of Time (F(3,48)=66.17, p <0.001) and a significant Hemisphere x Time interaction (F(3,48)=4.98, p <0.01). No effect was found regarding Condition x Time (p>0.4). Descriptive analysis regarding the effect of hemisphere showed that, as already reported for the 2x6 ANOVA, ERD was stronger in the right hemisphere. Descriptive analysis regarding the effect of Time showed an ERD pattern as expected (due to the presentation of a visual stimulus). Descriptive analysis regarding the effect of Hemisphere x Time showed that ERD was stronger in the right hemisphere for the first 3 time windows. Post-hoc comparisons showed that all of these differences were significant (for window 1 and 2 p <0.001, for window 3 p<0.05).

Due to these results, we repeated the comparison of the lateralization index for occipital and central electrodes. For this, for each participant and each region, we took the mean value of alpha frequency power over the four conditions of the right hemisphere in each time window, subtracted the mean value of alpha frequency power over the four conditions of the left hemisphere and divided the result by the mean value of these two numbers. Then, with the results of this calculation, we conducted a repeated measure 2x4 ANOVA with 2 factors of Region and 4 factors of Time.

Still, as results did not show any significant Condition effect or Condition x Time interaction they support the notion that alpha ERD in occipital cortices is discriminable from alpha ERD in motor areas, showing only effects likely due to visual attention without differing results for the conditions of interest.

Low Beta range (14-20 Hz)

The results of the 2x4x4 ANOVA for the low beta range (14-20 Hz) revealed a significant main effect of Time (F(3,48)=41.03, $p<0.001$) as well as a significant interaction Condition*Time (F(9,144)=2.02, $p<0.05$). Descriptive analysis regarding the main effect of Time showed a typical ERD-ERS-rebound pattern (see also Avanzini et al., 2012).

Descriptive and Post-hoc analysis regarding the interaction of Condition x Time showed the following characteristics (see **Figure 7**, next page). 1) In the beta range, already in the first ERD phase (first 1000 ms of stimulus presentation) conditions differed. Descriptive analysis showed that early ERD was strongest for steadicam, followed by dolly, then zoom, then still condition. Post-hoc comparisons showed that differences were significant between still and steadicam ($p<0.05$), zoom and steadicam ($p<0.05$) and dolly and steadicam ($p<0.05$). 2) Also in the following time window of further ERD/early ERS conditions differentiated. Descriptive analysis showed that ERD was still strongest for the steadicam, followed by dolly, then zoom, then still condition. Post-hoc comparisons showed that differences were significant between still and steadicam condition ($p<0.001$), still and dolly condition ($p<0.05$), zoom and steadicam condition ($p<0.05$), as well as dolly and steadicam condition ($p<0.01$). 3) In the rebound window (first second after stimulus offset) the order changed. Descriptive analysis showed that zoom had the highest power, followed by dolly then still, then steadicam. Post-hoc comparisons showed that differences were significant between zoom and steadicam ($p<0.05$) and dolly and steadicam ($p<0.05$) condition. 4) In the returned-to-baseline window no significant differences among conditions were measured.

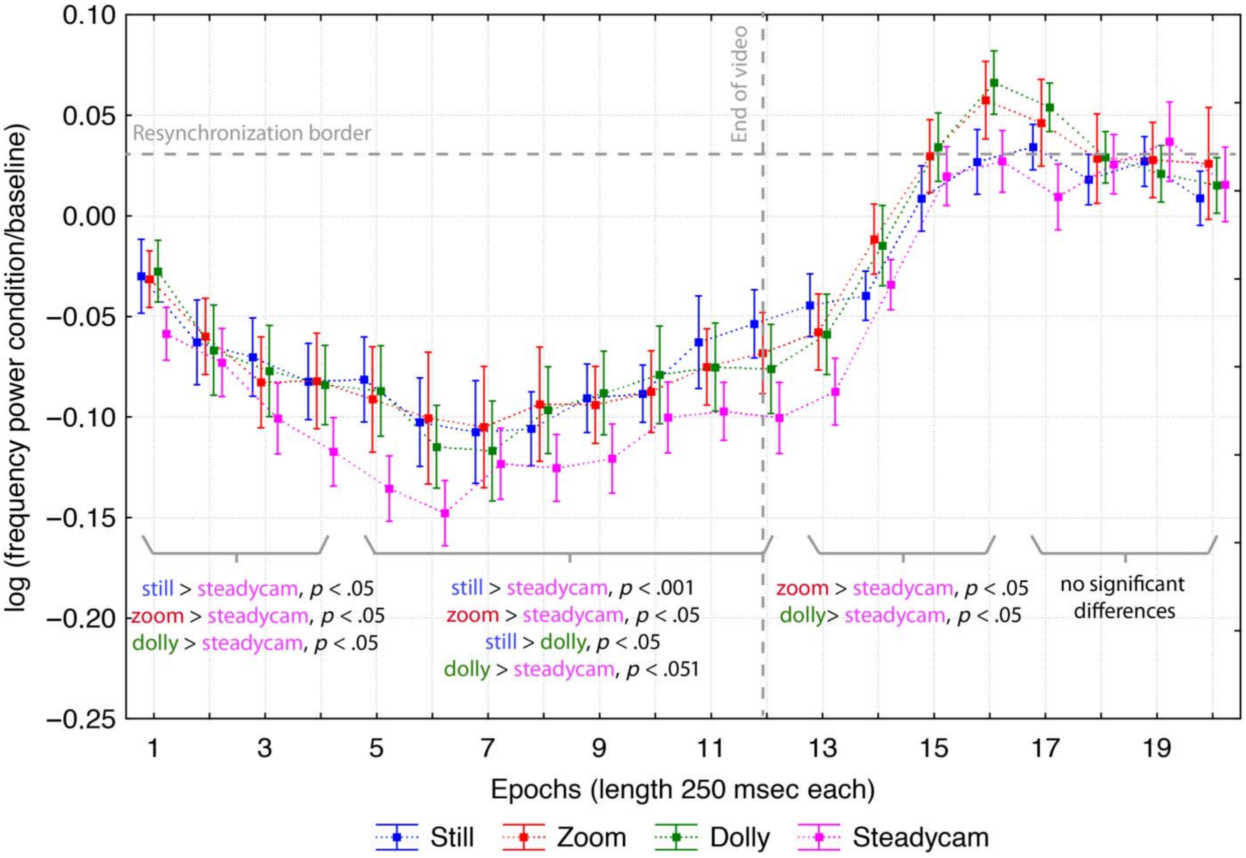


Figure 7. Central low beta frequency power over 20 epochs. Statistics over four time windows as indicated. Condition * Time interaction, F(9, 144) = 2.02, p < .05. Figure taken from Heimann et al., 2014.

III.1.4.2 Results – Rating task

Results of the rating task for question one (see **Figure 8A**, page 61) showed that participants felt more involved in the scene when the camera was approaching the agent in comparison to when the still camera was used. A One-way ANOVA with the single factor of Condition (still, zoom, dolly, steadicam) showed a significant main effect (F(3,48)=13.54, p<0.001). Post-hoc comparisons showed significant differences between still and all other conditions (p<0.001).

Similarly, results of the rating task for question two (see **Figure 8B**, page 61) showed that participants felt more like the actor, that is, like being in the position of the actor in the scene when the camera was approaching the agent in comparison to the still camera. A One-Way ANOVA with the single factor of Condition (still, zoom, dolly, steadicam) showed a significant main effect (F(3,48)=11.29, p <0.001). Post-hoc comparisons showed significant differences between still and all other conditions (p<0.001).

Results of the rating task for question three (see **Figure 8C**, next page) showed that participants felt the zoom as less effective in making them feel like they themselves were approaching the scene.

A One-way ANOVA with the single factor of Condition (zoom, dolly, steadicam) showed a significant main effect (F(2,32)=6.77, p<0.01). Post-hoc comparisons showed significant differences between zoom and all other conditions (p<0.01).

Results of the rating task for question four showed that participants did not feel any difference among conditions in terms of the way they felt at ease while watching the video clips. A One-way ANOVA with the single factor of Condition (still, zoom, dolly, steadicam) showed no significant main effect (F(3,48)=0.59, p>0.6).

Results of the rating task for question five (see **Figure 8D**, next page) showed that participants found the camera movement more realistic when the steadicam was used. A One-way ANOVA with the single factor of Condition (zoom, dolly, steadicam) showed a significant main effect (F(2,32)=6.91, p<0.01). Post-hoc comparisons showed significant differences between steadicam and zoom (p<0.01) and between steadicam and dolly (p<0.05).

The results for question six (see Figure 8E) showed that participants found the camera movement more resembling a person's movement approaching the scene when the steadicam was used**.** A One-way ANOVA with the single factor of Condition (zoom, dolly, steadicam) showed a significant main effect (F(2,32)=16.14, p<0.001). Post-hoc comparisons showed significant differences between zoom and steadicam as well as between dolly and steadicam (p<0.001).

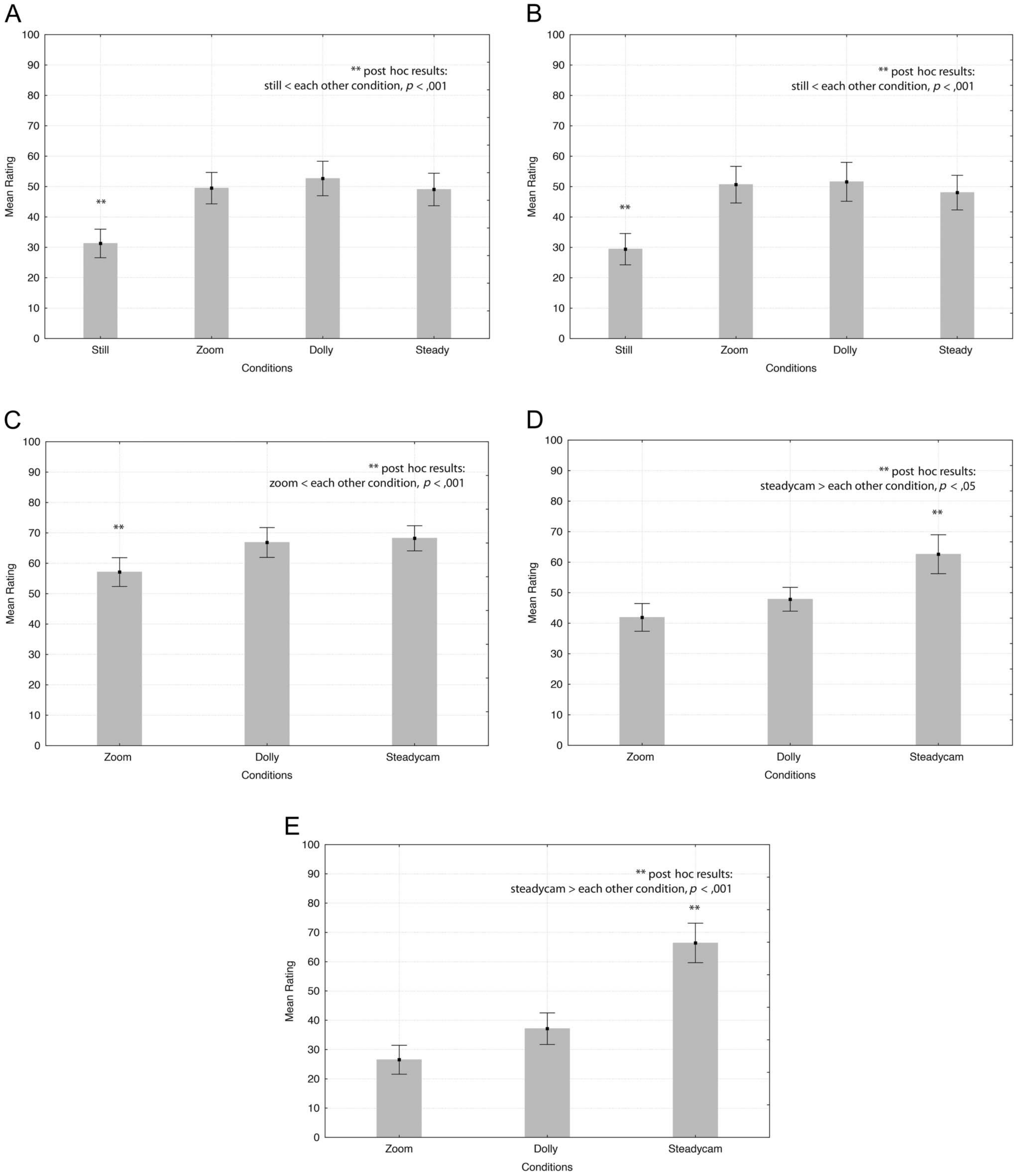


Figure 8.Results of the rating task including several questions assessing spectator's subjective experiences with the movie clips. (A) How much did you feel involved in the scene? F(3, 48) = 13.544, p <0.001. (B) How much did you feel like the actor? F(3, 48) = 11.291, p <0.001. (C) How much did you feel as if you yourself would approach the scene? F(2, 32) = 6.7766, p <0.01. (D) How realistic did you find the camera movement? F(2, 32) = 6.9131, p<0.01. (E) How much did you feel the camera movement resembled a person's movement when approaching the scene? F(2, 32) = 16.138, p <0.001. Figure taken from Heimann et al., 2014.

III.1.5 Summary and discussion of results

Previous studies showed that during the execution and the observation of goal directed actions rolandic mu-rhythm shows ERD in both of its supposed components: central alpha frequency range and central lower beta frequency range (Derambure et. al., 1993; Leocani et al., 1997; Pfurtscheller & Lopez Da Silva, 1999; Pfurtscheller & Aranibar, 1979; Pfurtscheller & Berghold, 1989; Stancak & Pfurtscheller, 1996; Toro et al., 1994). In the present study we specifically investigated if this typical ERD/ERS pattern is modulated by the observation of video clips produced via four different uses of the camera. All videos showed an actor grasping an object from a table in front of him. In 25% of the video clips, the camera recording the scene stayed fixed on a tripod 260 cm away from the object; in another 25% of the video clips a zoom was applied slowly magnifying the scene until the size of the object correspondded to the size it would have if the camera would have been placed in a distance of 80 cm from it; in the third 25% of the video clips during recording the camera was moved towards the scene by means of an automatic transport on dolly tracks (preprogrammed, steady pace) stopping when the camera reached a distance of 80 cm from the object; in the last 25% the camera was carried towards the scene until reaching the same distance of 80 cm from the object by means of a steadicam construction used by a professional camera man. This experimental setup was based on the before portrayed hypothesis, that camera movements play a fundamental role in the spectators' immersion in movies (see Chapter II.3). Precisely, descriptions of such camera movements underline that especially camera movements able to closely simulate vision in real life can activate the spectator's whole body, by this enabling an evaluation of the moving images in terms of the vision of an intended consciousness exploring world. The statements of several scholars regarding the role of embodied cognition and especially action-perception links in art perception (see for example Gallese & Freedberg, 2007; Fingerhut, 2012; Gallese & Guerra, 2014) have furthermore suggested that this process could be crucially supported by the mirror mechanism, proposed to play a major role not only in action understanding (see Gallese & Sinigaglia 2010, 2011) or action preparation in real life (for a critiques of the simulation theory of Gallese & Sinigaglia see Gallagher, 2001, 2007) but also in the immediate feeling of presence of things and beings seen, if depicted or real (see Fingerhut, 2014). This led to the hypothesis that moving images closer resembling the vision we entertain in real life would be correlated with a stronger ERD of the mu-rhythm, a commonly used marker of the mirror mechanism in humans.

Analysis of the present EEG data showed the following:

1) As shown in Figures 4 and 5, execution and observation of goal-directed hand movements produced significant ERD in both hemispheres for selected central alpha frequency ranges (8-14 Hz) as well as central lower beta frequency ranges (14-20 Hz) when comparing the mean value of frequency power of the whole time of action observation. No significant ERD for middle and high beta ranges (24-30 Hz) was detected. Furthermore, central alpha ERD was strongest during the participants' own hand action execution, while for lower beta action execution did not evoke stronger ERD than during action observation. In contrast, the steadicam condition evoked significantly stronger lower beta ERD than action execution.

These findings in general corroborate previous research describing central alpha and lower beta ERD during the execution as well as the observation of goal-directed hand actions.

2) Remarkably, as shown in Figures 6 and 7, in both frequency ranges that had been described to show ERD during execution and observation of goal-directed actions, two common characteristics of the EEG time-course were observed. First, during the time of video observation (time windows 1 and 2), descriptive analysis showed the strongest ERD for the steadicam, followed by the dolly, then the zoom, then the still condition. Statistically, in the alpha frequency range, significant differences were found between still versus steadicam as well as between zoom versus steadicam, and still versus dolly from 1000 ms after stimulus onset until stimulus offset. In lower beta frequency range, significant differences to steadicam included still versus steadicam, zoom versus steadicam as well as dolly versus steadicam. These differences reached significance from the very beginning of stimulus presentation. Furthermore, in the second time window, also in the beta range the difference between still and dolly became significant.

Second, during the rebound phase (third time window), in both ranges descriptive analysis showed the steadicam condition having the lowest power, but now first followed by still, then dolly, and lastly zoom condition. In the alpha range significant differences occurred between zoom and still and between zoom and steadicam, while in the lower beta frequency range differences between zoom steadicam and dolly and steadicam were significant.

Taken together, these results indicate that reducing the distance between observer and observed agent, realized by moving the camera towards the scene, evokes stronger ERD of the mu-rhythm during the observation of goal-directed hand actions. This difference was most pronounced when the camera movement was realized by using the steadicam (significantly different from the zoom and the still in the second and third window in alpha and from all other conditions in the first three windows in beta, except for the difference from the still in time window three. Results of occipital control recordings did not support the interpretation that this difference might be due to increased overall attention evoked by the observation of these specific filmed actions.

Results of our Rating task, as shown in Figure 8, furthermore showed that participants clearly rated those movies in which the camera approached the scene as more involving than those filmed by a still camera. Furthermore, they perceived the movements of the steadicam as being the most natural and most resembling the movements of an approaching observer, thus eliciting the feeling that the observer him/herself would walk towards the scene. However, it was not always possible to distinguish between steadicam and dolly cam regarding the last question.

These results indicate that approaching the scene by means of a camera enhances the observers' mirror mechanism when watching the video clip (showing a goal related action). However, this effect only appears if the perceptual experience induced by the different video clips and the visual experience we normally employ while moving ourselves actually resemble each other. Such similarity seems to depend on the filming technique and appears to be strongest when video clips are filmed with the steadicam.

Furthermore it seems as if the observation of images produced by other camera movements and uses, such as the dolly cam and the zoom, correlate with a faster rebound of the mu-rhythm. Indeed, in this phase (third time window) the differences between still and steadicam are no longer significant in either band range, while there are significant differences between still versus zoom as well as between steadicam versus zoom and steadicam versus dolly. Interestingly, Koelewijn et al. (2008) reported that beta oscillations were modulated in a similar way by the correct/incorrect nature of the observed action. In their study, they recorded EEG during an execution/observation task. In the execution task participants were asked to execute button responses according to instructional cues. In the observation task they saw other persons performing the same task, giving correct or incorrect responses. Results showed that beta oscillations during action observation were more strongly modulated if the action observed was "incorrect" according to the given cue. This stronger modulation was especially visible in the rebound phase, with beta showing a significantly stronger rebound when the answer was incorrect. We suggest that the results of our study might be interpreted as showing a similar effect - though not so much based on a training performed directly beforehand, but rather on participants' previous perceptual experience of the real world. That is we propose to consider the zoom and dolly conditions as presenting the observer with an "incorrect" representation in so far as the movement of the camera does not resemble the movement of an actual person. Hence, the still and steadicam conditions would correspond to natural, thus "correct", visual experiences.

In conclusion, we propose the existence of a familiarity effect regarding visual traces of camera movements in filmed stimuli. This effect can modulate the mirror mechanism activation during observation of goal-related hand actions. That is, among videos dynamically reducing the distance between the observer and the observed agent, only videos simulating the "natural" vision of a human observer approaching an agent can elicit a significantly stronger ERD in comparison to videos showing the same scene from a fixed distance. Furthermore, the artificiality of other ways of simulating the dynamic distance reduction (such as zoom or dolly) might be reflected in differences in the time course of the rebound phase. This shows that the time-course of mu-rhythm ERD/ERS/rebound is modulated by the resemblance between the effect of camera movements and ordinary human vision. Familiarity with the visual experience provided by the video predicts mu ERD/ERS/rebound time-course.

Nevertheless, several questions remain open.

1) The results indeed cannot sufficiently support the hypothesis that the stronger activation of the mirror mechanism supports the spectator's involvement in the fictional world. Though the rating task showed that video-clips produced by the steadicam were subjectively judged by participants as more giving the feeling of being in place of a person approaching the scene (that is quasi embodying the camera), involvement ratings in general did not differ among zoom, dolly and steadicam condition. Admittedly, subjective involvement in a scene is hard to assess. In fact, a conscious rating of of the involving nature of videos seen dozens of times in the last hour might not be the optimal method to do so. Clearly, the results of the rating task might be influenced by these circumstances and future research should try to find better ways to measure this factor.

2) Moreover, participants' persistent judgment of feeling more in the place of a person approaching the scene when the rated video was produced by means of a steadicam, raises another question, formulated in the following.

As already reported in Chapter II.2 of this thesis, motor cortex activations attributed to the mirror mechanism have not only been found during the observation of actions, but also during the observation of traces of actions such as writing, brush painting etc. (see Sbriscia-Fioretti et al., (2013); Heimann et al. (2013); Umiltà et al. (2012); Longcamp et al. (2006, 2008)).

This proposes two possible causes for the stronger ERD we found for video-clips produced by a steadicam. Firstly it might be, that the usual activation of the motor cortex during action observation resembling the activation of the motorcortex during our own grasping movement of the right hand is simply enhanced by a more "realistic" representation, that is a representation that by including more depth cues etc. closer resembles the vision we are used to from real life. On the other hand we might also hypothesize that the camera movement itself, as a trace of a walking movement, could elicit an additional motor cortex activation not similar to motor cortex activations during our own grasping actions with the right hand, but rather resembling such activations during actions involving muscles involved in walking. Indeed, this is reminiscent of descriptions from film theory talking about an "embodiment" of the camera in the specific sense of the camera taking over the place of an actual person in the fictional world, "through whose eyes" the scene is seen and lived (see for example Nielson, p. 242 and following). Future studies need to clarify this issue by investigating if camera movements produced by a steadicam filming an empty room (without performing actor) can equally elicit motor cortex activations of the reported nature. It might also be possible to further distinguish the activations found by individually separating the intrinsic mu-rhythms for foot and hand areas, then comparing the results of respective analysis (see Pfurtscheller, Neuper, Andrew, & Edlinger, 1997). Importantly, if such studies would show the predicted activations this would mean the discovery of a third kind of cases involving the mirror mechanism. Not only the perception of an action performed by another agent, or the perception of a trace of an action, in the past performed by an agent, but even the perception of perceptual movements of another agent could then be declared to involve the activation of the motor cortex. I will come back to these issues in Chapter IV of this thesis, when combining the results of both presented studies.

III.2 *A Cut in the Mirror?!* – A high density EEG study investigating the neural correlates of different montage techniques in film

This chapter and the study presented in it focuses on the second type of movement defining motion pictures, according to Vivian Sobchack: "the movement between projected images called editing" (Sobchack, 1982). Specifically, our experiment investigated spectators' brain activity during the observation of short movie clips produced via the editing of two shots showing a continuous scene but being filmed from different angles. Half of these clips were edited according to a set of rules commonly referred to as continuity editing, while the edits in the other half violated this system by not complying to the 180° rule, one of the most important guidelines of continuity editing.

Again, before reporting about experimental setup and results, the following section will clarify the precise idea behind the investigation by framing how editing can be and has been considered as a fundamental means of the creation of a fictional world via activating and extending the spectator's body.

III.2.1 Introduction: editing, continuity editing and its grounding in embodied perception

As Bordwell & Thompson argue in their introduction to film art, film editing is the "coordination of one shot with the next" (Bordwell & Thompson, 2004). The term "shot" refers to a single, continuous piece of recording produced with a moving-picture camera. An ordinary Hollywood film typically is made of around one thousand shots, action movies might even contain more than two thousand shots. Taking the average length of a Hollywood movie to be approximately ninety minutes, this makes an average shot length of 2.7-5.4 seconds (Bordwell & Thompson, 2004). According to Hochberg & Brooks, here lies one of the features most obviously distinguishing film from experiences in the real. At the same time it is also here that we find one of the most crucial capacities of film: to be able to represent scenes and events in a piecemeal fashion, that is juxtaposing scenes not spatially or temporally connected thus acting highly associative while at the same time avoiding any redundant sequences (Hochberg & Brooks, 1978). As it will be further illustrated, the success of these representations is not taken to be granted of course but rather highly depending on skill and experience in handling film's raw material.

Usually, it is the director guided by his narrative intentions who decides at what point in time a running shot has to be “cut” to best fulfill its purpose and which other shot should be edited to make the story proceed. In former times this indeed entailed cutting the celluloid film between two frames and glueing the end of the current shot to the beginning of the next selected shot. But even when using modern digital techniques, the result of the editing with respect to the flow of the movie is still the same: the old shot suddenly ends and the new one begins. It seems a curious fact, that this “sensorially brutal act of replacing the entire visual scene”, as Tim Smith in his elaborate thesis about continuity editing puts it (Smith, 2005, p.22), does not create conscious experiences significantly imbuing the spectator’s experience of the movie as a whole. Indeed, empirical research has shown that viewers in fact find it hard to recall specific details of the editing after a movie presentation. Instead, they apparently recall the film’s events as one continuous sequence (Messaris, 1994).

It has been repeatedly suggested, that this “invisibility of cuts” (see Reisz & Millar, 1953) must be attributed to the prominent use of continuity editing. It is commonly agreed upon that continuity editing, synonymous with the Hollywood Style of filmmaking emerging during the earliest days of cinema and conventionalized in the grand years of film between 1930 and 1940, is the most dominant style of editing today (Smith, 2005, p.22). The basic principle of this kind of editing lies in the compliance to a set of rules, or, as Smith pronounces, “abstract heuristics” or “rules of thumb”, describing how a scene should be staged, filmed and edited so that the viewer can comprehend the event flow with minimum effort” (Smith, 2005, p.22). Regarding montage, this “minimal effort” has been declared to correspond with the avoidance of any major discontinuities regarding space, time and causation between the two edited shots (thus the name continuity editing, see also Magliano & Zacks, 2011). The further explanations will illustrate this.

The continuity style of filmmaking is based on a system of guidelines regarding camera placement and editing commonly referred to as the 180° System. According to the 180° System, the initial shot of a scene (establishing shot) crucially orientates the spectator by drawing an imaginary line, called the “axis-of-action”, which divides the action space in two halves: the one the camera is located in (as being placed within a circle orthogonally focusing on the (180°) division line, with the action taking place at the centre) and the one being on the other side of that line (see Figure 9, next page).

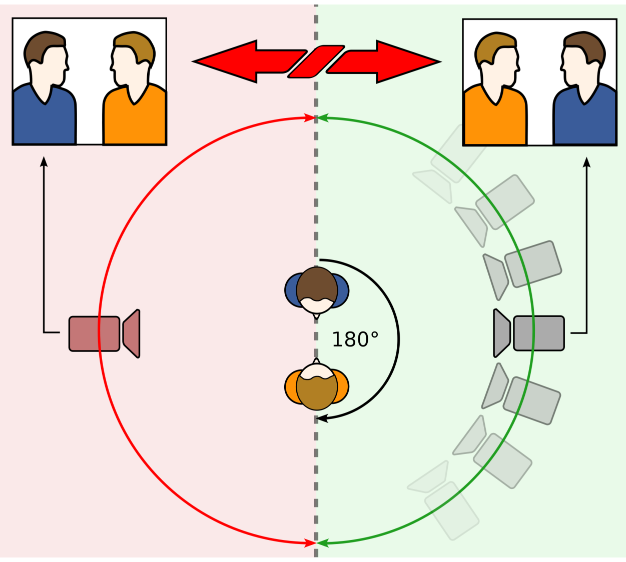


Figure 9: An Illustration of camera displacements between shots complying to or violating the 180° rule. Taken from Wikipedia: 180-degree rule. (2014, December 31). In *Wikipedia, The Free Encyclopedia*. Retrieved 15:40, January 15, 2015 from http://en.wikipedia.org/w/index.php?title=180-degree_rule&oldid=640401700.

As a first rule deriving from this "staging" (commonly referred to as the 180° rule), the position of the camera can be varied between two shots only as long as this line is not crossed. Interesting is the explanation given for this rule: Despite the actual displacement, so the statement goes, the experience of this kind of "continuity edits" does not cause a disturbing violation, precisely, as it stays "close to what we know from real world perception" (Magliano & Zacks, 2011, p.2). Indeed, bodily movements in combination with blinks and saccades causing short interruptions of the processing of currently given visual information, can lead to experiences possibly comparable with what happens to us during a cut (see also Murch, 2001). A breaking of the 180° rule, on the other hand – in its result called "reverse shot" – supposedly crucially violates our perceptual habits acquired in a spatially, temporally and causally stable world. This is due to that fact that, if a spectator would like to change his perspective on a scene for 180° in real life, he would have to get on the other side of the setup. Such a bodily displacement is possible of course, but in the real world depending on bodily movements (such as walking or running) that need time. Furthermore, on the way to the new standpoint, the spectator would most probably not close his eyes, so that the bodily movement would be reflected in continuous perception. Even if taking into account short interruptions elicited by blinks or fast eye-movement, it is very unlikely, that our perception in real life provides us with an experience even close to a reverse shot in film. To be precise, such edits (that is reverse shots) violate spatial continuity as, without a trace of bodily displacement of the camera, suddenly left becomes right and vice versa. They also violate action continuity (related to causation) as due to the spatial violation described before the motion flow across shots is disturbed and we cannot be sure if the shot before, with regard to its content, links directly to the shot after. Lastly, they possibly also violate temporal continuity as the spatial jump needed to explain the new perspective in our bodily habits, possibly links to a bodily movement needing time. It thus represents a jump in time. While continuity edits might not even be noticed, so the theory of continuity editing states, these kinds of cuts are not easily overseen, unless embedded in other editing techniques (such as cross-cutting etc., see for example Bordwell, 1985). Indeed, it has been demonstrated by a number of studies that the different editing techniques are correlated to different experiences and cognitive behaviors in a number of different tasks. Already in 1990, it was shown that participants that are asked to press a button as soon as they see a cut respond significantly faster to discontinuity edits than continuity edits (D'Ydewalle & Vanderbeeken, 1990; Schröder, 1990). In 2009, Zacks et al. even reported, that continuity edits, much more then discontinuity edits, were sometimes missed by spectators performing the same task. Both of these studies supported the hypothesis that continuity rules produce "invisible cuts". Furthermore however, it was also shown that 200-400 ms after discontinuity cuts, participants showed significantly more saccadic eye-movements, possibly indicating the need to extract visual information (d'Ydewalle, Desmet, & Van Rensbergen, 1998; Hochberg & Brooks, 1978; May, Dean, & Barnard, 2003). This period was shown to be preceded by a period of about 150ms in which no new visual information is processed to the level of awareness, possibly indicating cognitive overload (Geiger & Reeves, 1993). As related to this, because of indicating an attentional drain caused by discontinuity edits, has been interpreted the finding that participants that needed to perform a secondary task whilst watching a movie, showed prolonged response times to the secondary task after discontinuity edits (Geiger & Reeves, 1993; Lang, Geiger, Struckwerda, & Sumner, 1993). Also, participants in a memory task about watched movies showed better and faster recognition memory for information originally presented after a discontinuity edit (Lang, 1991; Frith & Robson, 1975). However, interestingly, recalls of entire film sequences have shown to suffer rather than profit from frequent discontinuity edits (Frith & Robson, 1975; Kraft, 1987).

Taken together these findings clearly indicate that continuity edits and discontinuity edits are not only experienced differently, but they correlate with different further processing with crucial impact on most important cognitive functions, such as attention and memory. Regarding the effect of montage on film experience, the reported results furthermore suggest that discontinuity edits might have the power to interrupt the filmic illusion by explicitly marking a violation of visual habits as they are known from real world perception. At the same time it seems possible that continuity edits can counterbalance this effect, while creating an experience similar enough to vision in normal life to activate our body in a way also overcomes certain violations. As Smith puts it:

“Film can be viewed as a deviant form of visual experience. Its level of deviancy is controlled by continuity editing” (Smith, 2005, p.18)

III.2.2. Experimental idea and hypothesis

It is the suggestion of this thesis, that the idea that movies can involve us they way they do by using images close to our experience to activate our normal skills of embodied perception, can be tested by neuroscientific experiments deriving their three hypotheses from the framework outlined in Chapter II of this thesis (see especially Chapter II.4). These hypotheses also relate to the application of montages creating experiences more or less closely resembling perception in real life. Specifically, we would like to suggest, that, firstly, comparing the onset of visual stimuli per se with continuity edits and discontinuity edits (violating the 180° rule), early brain activity signifying stimulus detection should significantly differ between the first stimulus type and the latter two. While the onset of a visual stimulus should be precisely correlated with the detection of a clearly and unambiguously new stimulus, cuts should rather be registered as violations of the action continuously perceived across the cut. Furthermore, secondly, there should be differences detectable regarding the further processing between continuity edits (fulfilling basic habits of visual perception in real life) and discontinuity edits (violating these habits), reflecting their difference in resembling real world perception as well as regarding spectators’ experiences. Lastly, especially montages creating moving images supposedly closely resembling vision in real life, can be predicted to stronger activate action perception links (in real life proposed to serve action understanding, action preparation and the feeling of presence of the surrounding world).

These hypotheses were approached via two different analyses of EEG data and one behavioral task comparing three different stimuli involving a visual change: the plain onset of a new visual stimulus, a continuity edit and a discontinuity edit.

Firstly, we performed an Event Related Potential (ERP) analysis looking for components in the event related time-course of the brain activity, in their modulation possibly reflecting differences between conditions in early and later stimulus processing phases. ERP analysis is a well established method in cognitive research, which led to the development of a catalogue of temporally and spatially described components reliably associated with specific cognitive functions. Continuity editing has been specifically hypothesised to depend on the substantiation of spectators' habitual expectation regarding spatial, temporal and causal constancy of the world or the perceived action (see Chapter III.2.1). In fact, a lot of work in EEG research has looked at violations of expectations, especially distinguishing semantic and syntactic violations in language (sentence) and music perception (see Hahne and Friederici 1999; Friederici, Pfeifer, & Hahne, 1993; Friederici 2002, Grodzinsky and Santi, 2008; Steinhauer & Drury, 2012; Koelsch, 2011). Most recently, the findings of this field have been picked up and further explored by research looking at action perception. Most interestingly, Maffongelli, Bartoli, Sammler, Koelsch, […], & D'Ausilio (submitted) found similar components indicating semantic and syntactic violations also for the visual perception of action sequences (including either unexpected content or a scrambling of temporal order). Precisely, Maffongelli et al. (submitted) found that semantic violations in action observation (regarding sequences of static images of an ongoing action including an unexpected element, such as brushing your hair with a toothbrush) elicited a frontally distributed EEG negative deflection around 400 ms after stimulus-onset. On the other hand, structural violations in action observation (sequences of static images of an ongoing action scrambled in their logical order) elicited first an early left anterior deflection, indicating the detection of the structural violation, followed by a late left anterior positivity, reflecting further cognitive processing. Due to the specific edits looked at in our experiment, not changing the content of the action seen but rather the perspective on it, we expected to find ERP components specifically marking a cut in general (that is, independently from the angle of camera displacement in simple comparison to a plain visual onset) as a structural violation in the above declared sense. Furthermore, we supposed that the two different cut conditions would differ among each other, if not already in the first component of the complex described by Maffongelli et al., then at least in the second component, indicating further cognitive processing. More specifically, we expected to find differences illuminating spectators' different experiences of and responses to continuity edits and non-continuity edits as described in Chapter III.2.1. While continuity edits tend to be rather overseen, non-continuity edits are easily remarked and seem to restrict our attention to other cognitive tasks simultaneously performed.

Second, we performed a time-frequency analysis focusing on the mu-rhythm in central (motor) areas (with its alpha and beta components) as a common marker for the mirror mechanism in humans (see Chapter III.1.2.1). Reflections on film experience based on 4EA approaches to cognition, as outlined in Chapter II.3, gave rise to the hypothesis that the spectator's strong involvement in the fictional environment created by film might be due to the activations of visual habits and experiences from the real world elicited by moving images, and especially by those closely resembling what is known to us from perception in real life. It was furthermore proposed that this process might crucially be based on action perception links that previous research suggested may activate the body, even if something is not (yet) in our reach or if it is completely unavailable to be bodily explored. A mechanism that has been repeatedly claimed to enable such an action-perception or emotion-perception link is the mirror mechanism as introduced in Chapter II.2 of this thesis. In the study "Moving mirrors – A high density EEG study exploring the effect of different camera movements on motor cortex activation during action observation" (see Heimann et al. (2014) as well as Chapter III.1 of this thesis) we found that indeed, the mirror mechanism was more strongly activated when spectators watch a video clip produced by use of a steadicam, the result of which best resembles vision during the spectators' own movement. We hypothesized that we would see a similar effect also in this study. To be precise, we expected a stronger and possibly shorter lasting ERD of the mu-rhythm during action observation after continuity edits, which in film theory are described as substantiating habits from real world perception, than after discontinuity edits, described to crucially violate our usual habits of perceiving derived from real world interactions.

The following paragraphs will describe the details of setup, and the analysis results of the respective experiment.

III.2.3 Participants, materials and methods

The participants in the experiment consisted of 20 healthy volunteers (10 male, 10 female, mean age 24.8 (SD 2,31), all right handed as assessed by (an Italian adaption of) the Edinburgh Handedness Inventory (Oldfield 1971)). Participants were recruited by public announcement and were paid 25 euros as reimbursement. Before the experiment, they received written and oral experimental instructions. After the experiment each participant was debriefed. Written informed consent was obtained from all participants before entering the study. The study was approved by the local Ethical Committee.

Stimuli consisted of video clips of 5 seconds length each. Each video was a montage of two single clips. In the first clip of 2 seconds length, two agents (one male, one female) stand in front of a table, at the center of which a small object (either a salt shaker or an espresso cup) is placed. The actors first look at each other before one of them (either the man or the woman) directs the gaze down to the object, the other following this gaze with her/his own one. In the second clip of 3 seconds length, the actor leading the gaze in clip one then grasps the object, picks it up and places it right in front of the other actor. The variable specifically manipulated for the experiment is the angle with which the camera is displaced from the first to the second shot edited together. In 50 % of the clips this angle is 90°, in the other 50 % this angle is 180° (in both cases the jump is once realized clockwise, and once counterclockwise). This made up for a total of 8 clips (always showing the same scene, but using two different objects (a cup and a salt shaker), two displacement angles (90° and 180°), and two directions of displacement: once counterclockwise, once clockwise). Video clips were recorded and edited in a professional film studio, enabling us to film the same scene under highly controlled conditions. **Figure 10** below shows 3 still frames demonstrating the difference in the camera displacement.

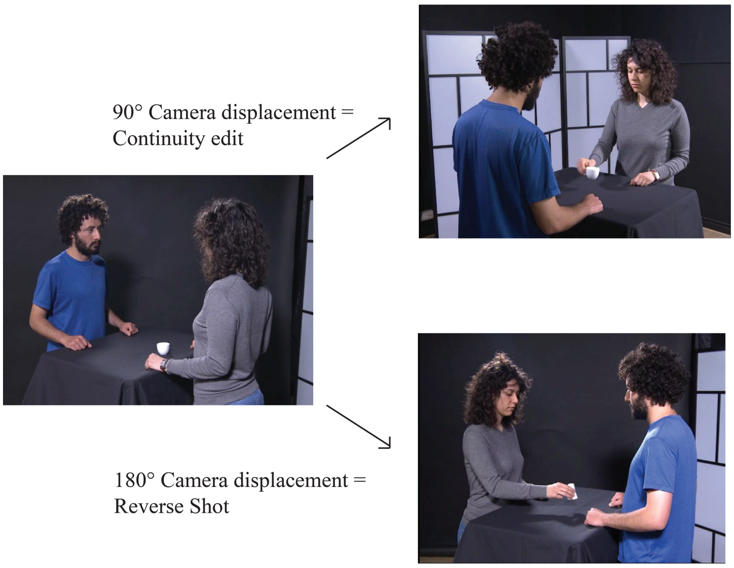


Figure 10: Illustrating the differences between applied edits, once involving a 90° camera displacement, once a 180° camera displacement across the edit.

The experiment consisted of two different parts comprising: 1) a 60 minutes EEG recording session (including breaks); 2) a 10 minutes rating task. With preparation and debriefing the whole procedure took about 1.5 hours per participant.

The EEG was recorded during three blocks of about 15 minutes length each. After each block participants were given the possibility of a short break. Two blocks consisted of 53, and one of 54 trials, leading to a total of 80 trials per condition (continuity edit: 90° displacement / reverse shot: 180° displacement). Participants were seated in an isolated EEG lab in front of a computer screen placed on a table at a distance of 50 cm. Participants were then instructed about the experimental procedure: Each trial began with a fixation cross (of random duration between 500-1000 ms), followed by one of the video stimuli (of 5 sec. length each (see before), presented in random order). In 50% of the trials, after stimulus presentation, first a light green screen (2 seconds), then a grey screen (intertribal interval ITI, 3 seconds) were displayed (guaranteeing the return of brain activity to baseline). Participants were asked to blink only during the grey screen period, to minimize eye movement artifacts in the periods of interest. In the other 50% of the trials, after stimulus presentation and before the grey screen, participants were presented a question on screen referring to either the object or to one of the actors in the video displayed before (for example: "Was the object that was handed over an espresso cup?" or "Was it the male actor who handed over the object?"). The answer had to be given by clicking the mouse on the left button with the right index finger or the right button with the right medium finger. This allowed participants to answer with "Yes" or "No" As we did not compare reaction times answer-button matches were not balanced among participants, but left button always indicated a "Yes", right button a "No". The mouse was positioned on the table in front of participants at a distance of 15 cm from their right hand. During video observation, participants were asked to keep the right hand at a fixed position indicated by a physical marker. When the question slide appeared, participants were asked to move their hand from start position to the mouse and give their answer, and then move back the hand to start position. If participants gave a wrong answer or did not answer within 5 seconds, they were told the trial was incorrect or the answer was given too slowly, and the trial was repeated. This action execution condition served both as control for attention and to record participants' ERD during action execution. See **Figure 11**.

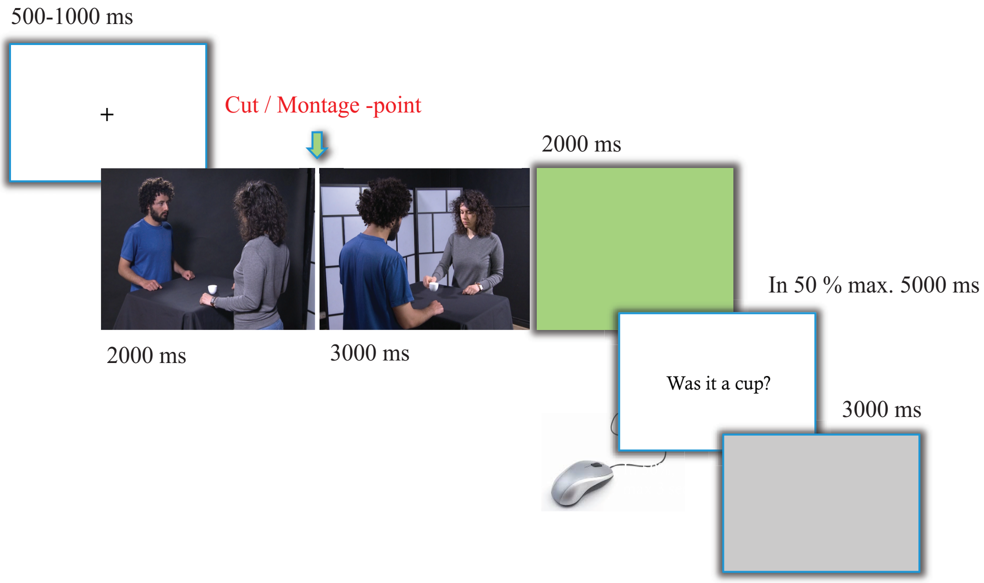


Figure 11: Experimental paradigm during the EEG recording.

The second part of the experiment consisted of a rating task. Participants, still sitting in front of the screen as during the EEG recording session, were again shown the video clips and for each of the clips they were asked five different questions (in 5 separate blocks (one for each question) always conducted in the same order). The questions were:

1) How much did you feel involved in the scene?
2) How much did you feel like the actor (the one handing over the object)?
3) How easy to watch did you find this video?
4) How natural did you find the representation of the scene in this video?
5) Which of the two videos was longer (regarding its duration)?

The first two questions were designed to investigate participants' potential feeling of involvement with the observed scene in terms of empathy with the actor. The third question was designed to explore how much at ease participants were with the different ways of editing. The fourth question was designed to measure participants' estimation of the ecological plausibility of the different types of editing used. The last question was designed to investigate if the different montage techniques have an influence of participants' time duration estimation.

EEG data was acquired by a 128-channel Sensor Net (Electrical Geodesic, Eugene, USA) and recorded within the standard EGI package Net Station 4.3.1. EEG was sampled at 500 Hz, and band-pass filtered online at 0.1-100 Hz. The electrodes impedance was kept less than 50 kΩ (checked after each block). The raw EEG data was recorded with the vertex (Cz) as the online reference and re-referenced off-line to the common average (Muthukumaraswamy, Johnson, & McNair, 2004). Stimuli were presented with E-Prime 2.0 and at the beginning of each trial, and all event markers were sent to Net Station. Participants' motion was monitored by the experimenter and video-recorded for off-line analysis; if participants moved during the observation or rest conditions the trial was excluded from further analysis. All further processing was done using the Matlab Toolbox FieldTrip (Oostenveld et al., 2011). Statistical analysis, if not further specified, were done using the STATISTICA software of StatSoft.

For ERP analysis, data were filtered off-line with an additional low passfilter of 45 Hz (resulting in 0.1-45 Hz data) and segmented into specific time epochs for baseline and montage conditions, each of one second length. For the montage conditions we selected the first second after the editing point in the videos – in 50% representing a montage complying to, and in 50% violating the 180° rule (further refered to as continuity edit (referring to 90° camera displacement) and reverse shot (referring to 180° camera displacement)). As baseline we took the first second after video onset. We decided for this baseline as it allowed us to distinguish ERPs caused by the onset of a new visual stimulus in general, from those specifically caused by the camera displacement across the edit. That means, that baseline and montage conditions were similar as both involve a visual change (in case of the baseline from black screen to the scene, in case of the montage conditions, that is the edit, from one shot to the other). The difference though is that the two shots following each other in our within-scene-montages bear a contextual relationship, as they show the same ongoing scene, only filmed from different angles.

Artifacts for all conditions were removed through visually inspected Independent Component Analysis as implemented in FieldTrip, considering temporal, topographic and spectral distribution of the component. Subsequent additional visual inspection of each segment led to the exclusion of all trials with still remaining artifacts. A minimum number of 50 trials for each condition was kept (fulfilled by all but two participants, who were consequently excluded from further analysis, resulting in 18 data sets finally analyzed).

ERPs were computed by averaging over trials and participants using time-locked-analysis over the segmented epochs as implemented in FieldTrip. As a first comparison of baseline and montage conditions in FieldTrip (cluster based permutation test using dependent samples t-test statistics, Monte Carlo method, based on 500 randomizations) did show significant differences for almost all electrodes, thus not allowing a sensible cluster identification, Regions of Interest (ROIs) were selected on the basis of the findings of prior studies investigating ERPs due to content/structure violations. Previous studies looking at sentences (Kutas and Hillyard, 1980; Kutas and Federmeier, 2000; Lau et al., 2008), musical phrases (Steinbeis and Koelsch, 2008; Koelsch, 2011) and goal-directed actions (Maffongelli et al., submitted) dissociated ERPs indicating content violations from ERPs indicating structure violations. While content violations were found to correlate with a negative component around 400 ms after stimulus onset (N400), structure violations were reported to elicit an earlier and specifically anterior negativity (ELAN). Furthermore, when task relevant, these early anterior left negativities were reported to be followed by a late positivity, in the action domain, being equally located in left anterior regions (see Maffongelli et al., submitted).

In these studies, usually four ROIs were defined: anterior left, anterior right, posterior left and posterior right. We adopted these ROIs and additionally defined two further central ROIs to allow a comparison with results of the ERD analysis when considering differences between the two Montage conditions. The resulting six ROIs are indicated in **Figure 12**-

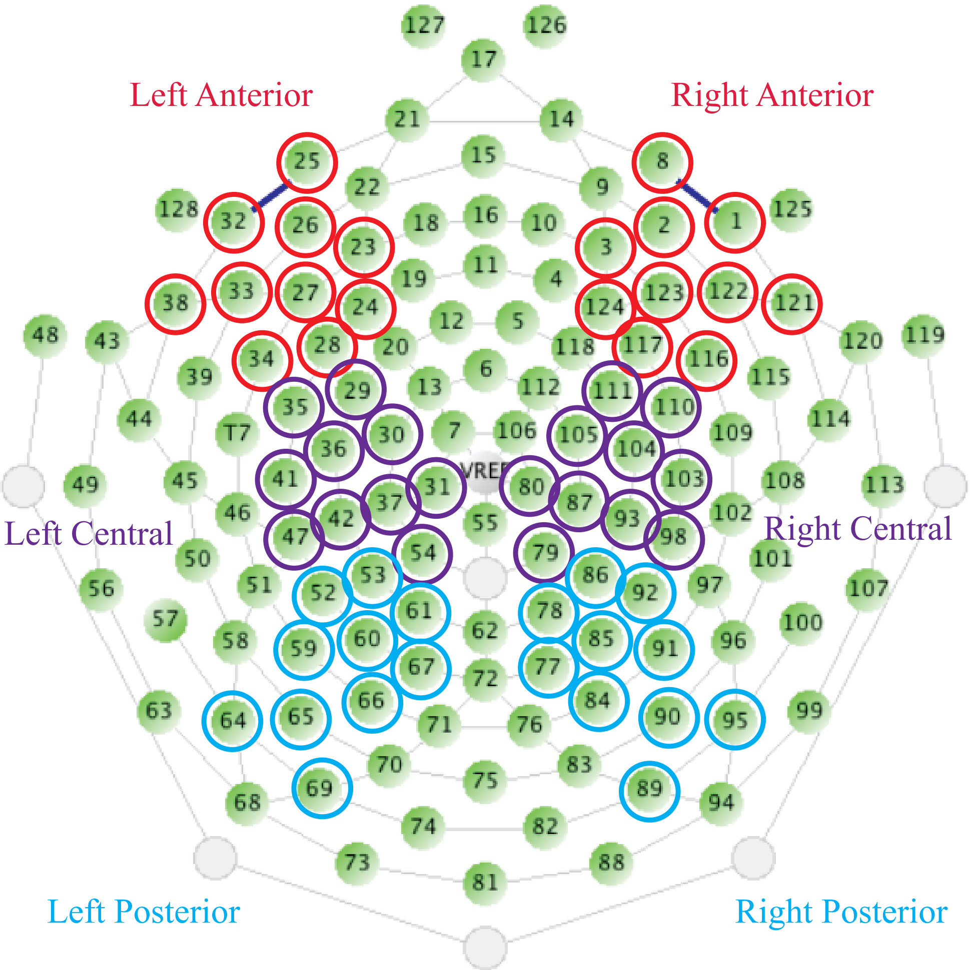


Figure 12: Selection of electrodes for the six ROI chosen.

Time-windows of interest were detected based on where (on scalp surface) ERPs reached their maximum values. (See also **Figure 13**, next page)

Time-window 1 is 140-190 ms after stimulus onset – centered around an early negativity (N1) in anterior and central regions and an early positivity (P1) in posterior regions.

Time-window 2 is 180-220 ms after stimulus onset - centered around an early positivity in anterior and central regions (P2) and an early negativity in posterior regions (N2)

Time-window 3 is 250-380 ms after stimulus onset – centered around a strong negativity in anterior and central regions (N3) and a strong positivity in posterior regions (P3)

Time-window 4 is 400-650 ms after stimulus onset – centered around a late positivity (P4-6) in all regions.

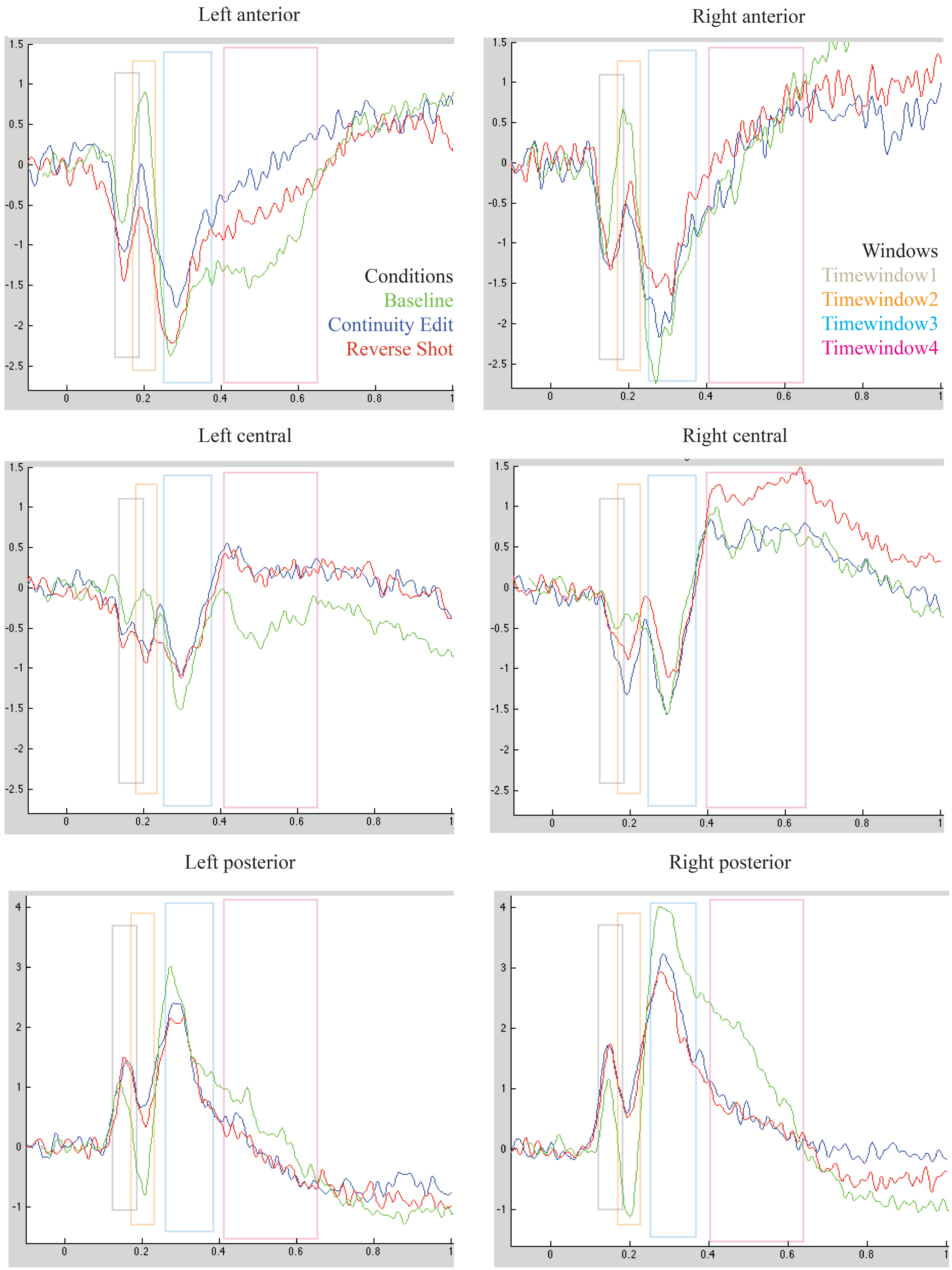


Figure 13: ERP course over the six selected ROI. Frames mark the 4 time-windows chosen due to maximum values of observable components and previous literature.

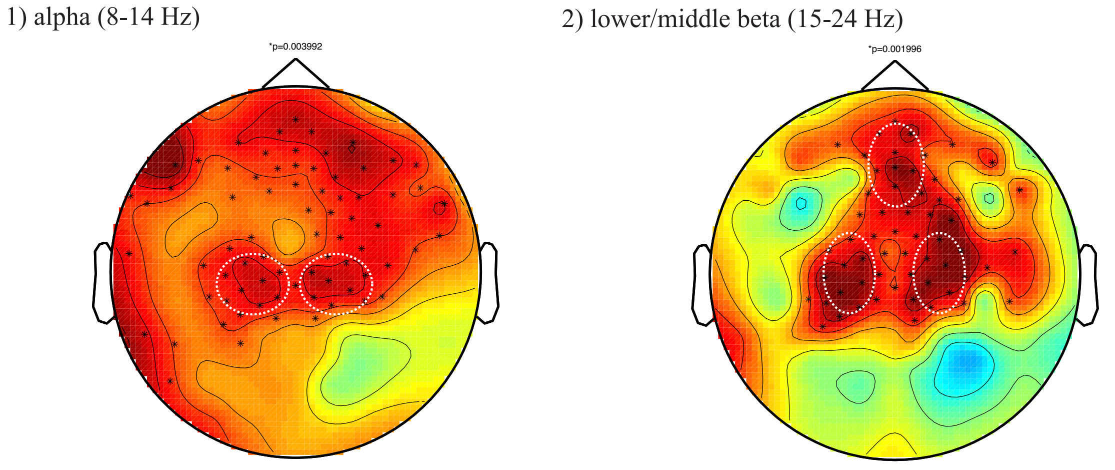
Data was extracted averaging potential over defined ROIs and time windows. Outliers were replaced with the corresponding average values of all participants (the criterium being ± 2.5 standard deviations from mean, resulting in the replacement of less than 1 % of all data). A final Kolmogorow-Smirnov test showed no significant results for all data, thus justifying the use of parametric statistics.

Regarding statistical analysis, in a first analysis, conducted to assess differences caused by a complete change of visual input in comparison to within-scene edits, we compared average potentials of baseline with average potential after within-scene-cuts (without separating continuity edit and reverse shot conditions). For each window a 2x2x3 ANOVA (2 Conditions x 2 Hemispheres x 3 ROIs) was conducted.

In a second step. we then compared continuity edit and reverse shot conditions using the same ROIs, time windows and statistical methods.

For all ANOVAs, Duncan post-hoc tests were used to further explore significant factors and interactions. Reported results in these tests were automatically adjusted for multiple comparisons by the software used.

For ERD analysis data were filtered offline with a bandpassfilter of 1-30 Hz and were segmented into specific time epochs. From observation trials (videos edited according to or violating the 180° rule), the three seconds after the editing point of the video clips plus the two seconds of green screen were analyzed. As baseline (time without any ERD expectation), 1000 ms of grey screen ending 250 ms before the start of the new trial (appearance of the fixation cross) were selected from the observation trials. From action execution trials, segments of 1000 ms were selected, starting 500 ms before the motor response (button press) and ending 500 ms after it. Artifacts were removed through visually inspected Independent Component Analysis as implemented in FieldTrip, considering temporal, topographic and spectral distribution of the component. Subsequent additional visual inspection of each segment led to the exclusion of all trials with still remaining artifacts. A minimum number of 50 trials for each condition were kept (fulfilled by all but three participants, who were consequently excluded from further analysis, resulting in 17 datasets finally analysed). Frequency bands of interest were the different components of the rolandic mu-rhythm, consisting of an alpha and a beta band (see for example Avanzini et al., 2012). The time–frequency analysis was performed for each participant on 1.5 second long segments for all conditions (baseline, continuity edit (90°), reverse shot (180°), action execution), using Hanning tapers in 1 Hz intervals with a sliding time window of 0.5 seconds, in the frequency range from 7 to 30 Hz. Frequency-power coefficients were calculated by taking the average across trials for each of the 128 channels. Electrode-clusters of interest were chosen for each of the two frequency bands of interest by means of a first exploratory statistical analysis comparing baseline and action execution condition in a cluster-based permutation test as implemented in FieldTrip (using dependent samples t-test statistics, Monte Carlo method, based on 500 randomizations). On the basis of the results showing the electrodes with significant differences (see topoplots below), for alpha as well as for beta bands two symmetrical central clusters (one in the left, the other in the right hemisphere) were chosen. As previous research described the source of the beta component of the mu-rhythm sometimes lying more frontal (see also Stancak, A., & Pfurtscheller, G., 1996), for beta bands we additionally selected one frontal cluster that also showed a strong difference between conditions. See **Figure 14** below.



Plotted results of cluster based permutation analysis comparing baseline and action execution condition (average of 100 ms before and after button press). Chosen electrode clusters for further analysis marked by white dashed line.

left central cluster: E 30, 31, 36, 37, 41, 42, 53, 54
right central cluster: E 79, 80, 86, 87, 93, 103, 104, 105

left central cluster: E 30, 36, 37, 41, 42, 47, 52, 53
right central cluster: E 87, 93, 98, 103, 104, 105, 110, 111
frontal cluster: E 4, 10, 11, 14, 15, 16, 17, 19, 21,

Figure 14: Electrode cluster chosen for the two band ranges of interest on the base of dependent samples t-test statistics using Monte Carlo Method as implemented in FieldTrip package.

Using these clusters, specific alpha- and beta-frequency bands were selected for each participant following the procedure described in previous studies (Oberman et al., 2007a,b; Babiloni, Del Percio, Rossini, Marzano, [...], & Eusebi, 2009). The individual peak (F) of attenuated frequency was determined by calculating the ratio between the frequency power in action execution trials and during baseline in the following sub-frequency bands: 8–9, 9–10, 10–11, 11–12, 12–13, 13–14 Hz for alpha, 15-16, 16-17, 17-18, 18-19, 19-20, 20-21, 21-22, 22-23, 23-24 for beta. Each value was then transformed into a log-ratio and the frequency that corresponded to the log-ratio with the most negative value was taken as F. A 3 Hz range frequency band was chosen for each participant (F - 1; F + 1) and frequency type (alpha and beta). For the following statistical analyses, the frequency power in this 3 Hz range was extracted in all conditions (number of participants selected per range: *Alpha*: 7-9 Hz: 5; 8-10 Hz: 4; 9-11 Hz: 1; 10-12 Hz: 3; 11-13 Hz: 2; 12-14 Hz: 2; *Beta*: 16-18 Hz: 1; 17-19 Hz: 4; 18-20 Hz: 7; 20-22 Hz: 4; 21-23Hz: 1). ROIs and specific frequency bands of each participant were then used to extract the final data from all four conditions and chosen time windows (see below). Before statistical analysis all data was corrected for normality by using a log transformation resulting in non-significant Kolmogorov-Smirnov tests for all conditions, justifying the use of parametric statistics.

For statistical analysis, in order to assess central alpha ERD in sensory-motor areas during different observation and action execution trials, we compared the extracted frequency power for the different conditions using several ANOVAs.

1) In order to generally assess central alpha and beta ERD in sensory-motor areas, for every band range considered (alpha, beta) we compared the frequency power extracted from baseline (average of 1000 ms ending 250 ms before new fixation cross) with its value during observation conditions (average of 1000 ms: selected was the second second after the cut (4th second of video observation) in continuity edit and reverse shot condition, in which full ERD can be expected) and action execution condition (average of 1000 ms: 500 ms before and 500 ms after button press). Specifically we used a repeated measures 4x2 ANOVA with four levels of Condition (baseline, two observation conditions – continuity edit and reverse shot – and action execution condition) and two levels of Hemisphere (left and right). For the beta bands an additional ANOVA for the frontal cluster was performed (comparing the four conditions only)
2) To assess differences between the two observation conditions (continuity edit and reverse shot) another set of ANOVAs was performed, also taking into account possible differences in the time-course of the activation. As it was shown that also the ERS following ERD of the rolandic mu-rhythm shows a specific pattern that can be modulated by contextual conditions (Muthukumaraswamy et al., 2004; Pfurtscheller et al., 2005, Heimann et al., 2013, Heimann et al., 2014), we separately extracted the data

of the 5 seconds after the cut (including 2 seconds of green screen after video offset), resulting in two 2x2x5 (2 Conditions x 2 Hemispheres x 5 Time Windows) ANOVAs (one for central alpha, one for central beta) with two levels of Condition (continuity edit and reverse shot), two levels of Hemisphere (left and right) and five levels of Time (each of the five seconds following the stimulus onset) as well as one 2x5 (2 Conditions x 5 Time Windows) ANOVA for the frontal beta cluster.

3) Since the central alpha frequency band (8–14 Hz) overlaps with the posterior alpha band, it is possible that recordings in central areas might be affected by this posterior activity. In order to check for such an influence, for the alpha-range selected for each participant we extracted the respective power values from 8 electrodes per hemisphere in occipital areas (electrodes 65, 66, 68, 69, 70, 71 73, 74 in left occipital lobe & electrodes 76, 82, 83, 84, 88, 89, 90, 94 in right occipital lobe) using the same frequency bands as previously described. We then repeated the described ANOVAs described also for these data.

For all ANOVAs, Duncan post-hoc tests were used to further explore significant factors and interactions. Reported results are automatically adjusted for multiple comparisons. Error bars in all the graphs represent standard errors.

Results of the rating task partly showed a violation of normality not correctable through transformation. In consequence non-parametric statistics (Wilcoxon Signed Ranks) were used for statistical analysis, resulting in 5 related sample test (one for each of the questions).

All Means and Standard Errors are noted in the table at the end of the results section.

III.2.4 Results

III.2.4.1 Results EEG: ERP

Comparison Baseline – Montage

We first compared baseline and montage conditions (continuity edit and reverse shot conditions taken together) in the four selected time windows.

For the first timewindow (140-190 ms) a 2x2x3 ANOVA (2 Conditions x 2 Hemispheres x 3 ROIs) showed significant interactions of Condition x Hemisphere ($F(1,17)=5.47$, $p<0.05$) and Conditions x ROI ($F(2,34)=4.41$, $p<0.05$) Post-hoc tests regarding the Condition x Hemisphere interaction and the Condition x ROI interaction showed that the effects were driven by left hemispheric ($p<0.05$) and anterior regions ($p<0.01$). This shows a significant difference between baseline and montage conditions pronounced in anterior and left hemispheric regions, because in the grand average ERP a negative component (N1) was found, which was more strongly pronounced for montage conditions. See potential course in anterior and left regions Figure 13, page 79, and for a representative example **Figure 15**.

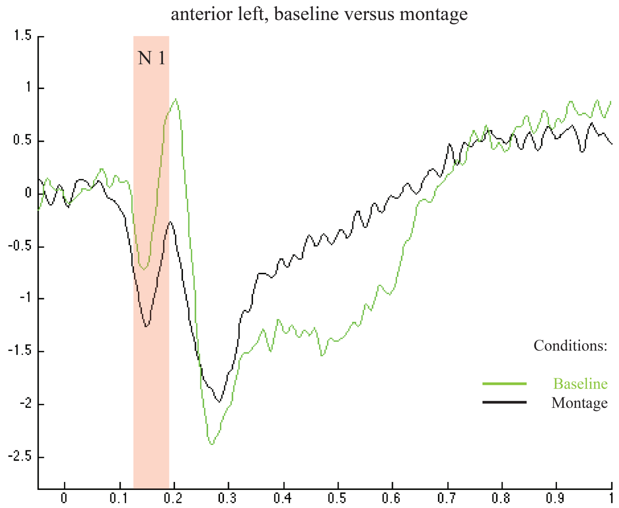


Figure 15: ERP time-course in left anterior ROI. Marked in red the N1 component pronounced for montage conditions in anterior and left hemispheric regions. Condition x Hemisphere interaction: $F(1,17)=5.47$, $p<0.05$ and Conditions x ROI interaction: $F(2,34)=4.41$, $p<0.05$. In anterior regions the significance level is $p<0.01$; in left hemispheric regions $p<0.05$.

For the second time window (180-220 ms), a 2x2x3 ANOVA (2 Conditions x 2 Hemispheres x 3 ROIs) showed significant main effects of Condition ($F(1,17)=10.47$, $p<0.01$), ROI ($F2,34)=13.02$, $p<0.001$) and a significant interactions for Condition x ROI ($F(2,34)=9.72$, $p<0.001$).

Post-hoc tests showed that, regarding the Condition effect, average potential was higher for baseline (p<0.01). Regarding the effect for ROI, we found that average potential in anterior regions and central regions was significantly smaller than in posterior regions (p<0.001). Finally, Post-hoc tests regarding the Condition x ROI interaction showed that in anterior regions, average potential of baseline was significantly higher than that of montage conditions (p<0.01), while in posterior regions this relation was reversed (p<0.01). This shows: a) a significant difference between baseline and montage conditions in anterior regions, which in the Grand Average ERP shows up as an early positivity (in the following referred to as P2), more strongly pronounced for baseline; b) a significant difference between baseline and montage conditions in posterior regions, which in the Grand Average ERP shows up as an early negativity (in the following referred to as N2) that is also more strongly pronounced for baseline. See **Figure 16** below.

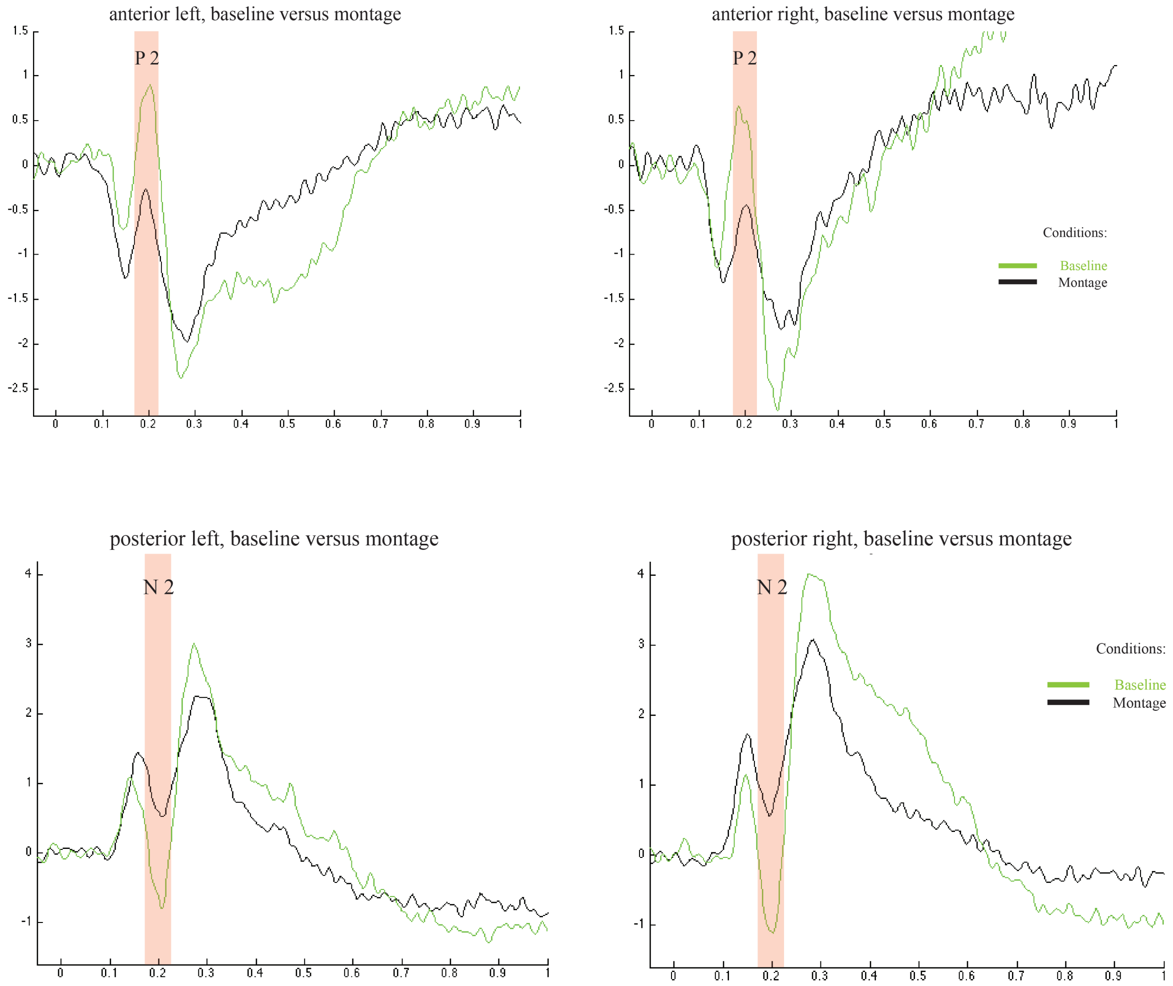


Figure 16: ERP time-course in anterior and posterior ROIs. Marked in red the P2 and N2 components that are pronounced for baseline. Condition x ROI interaction: F(2,34)=9.72, p<0.001; differences between conditions in anterior and posterior regions p<0.01.

For the third time window (250-380 ms), a 2x2x3 ANOVA (2 Conditions x 2 Hemispheres x 3 ROIs) showed significant main effects for ROI ($F2, 34)=45.09$, $p<0.001$) only.
Post-hoc tests showed that that average potential in anterior and in central regions was lower than in posterior regions (all $p<0.001$). No difference between conditions was found.

For the fourth time window (400-650 ms), a 2x2x3 ANOVA (2 Conditions x 2 Hemispheres x 3 ROIs) showed significant main effects for Hemisphere ($F(1,17)=23.5$, $p<0.001$), ROI ($F2,34)=12.29$, $p<0.001$) and significant interactions for Condition x Hemisphere ($F(1,17)=4.57$, $p<0.05$) and Condition x ROI ($F(2,34)=4.19$, $p<0.05$).
Post-hoc tests showed that, regarding the Hemisphere effect, average potential was lower in the left hemisphere ($p<0.001$). Regarding the effect for ROIs we found that average potential in anterior regions was lower than in central and posterior regions and average potential in central regions was lower than in posterior regions (for the difference between anterior and posterior regions: $p<0.001$; between anterior and central regions: $p<0.01$; between central and posterior regions: $p<0.05$). Post-hoc tests on the Condition x Hemisphere interaction only supported the findings of hemispheric differences in both conditions (for baseline $p<0.001$; for montage conditions, $p<0.05$). Post-hoc tests on the Condition x ROI interaction showed that in posterior regions, average potential in baseline was higher than average potential after montage ($p<0.05$). This shows a significant difference between baseline and montage conditions in posterior regions, showing a late positivity (in the following referred to as P4-6) more strongly pronounced for baseline. See **Figure 17** below.

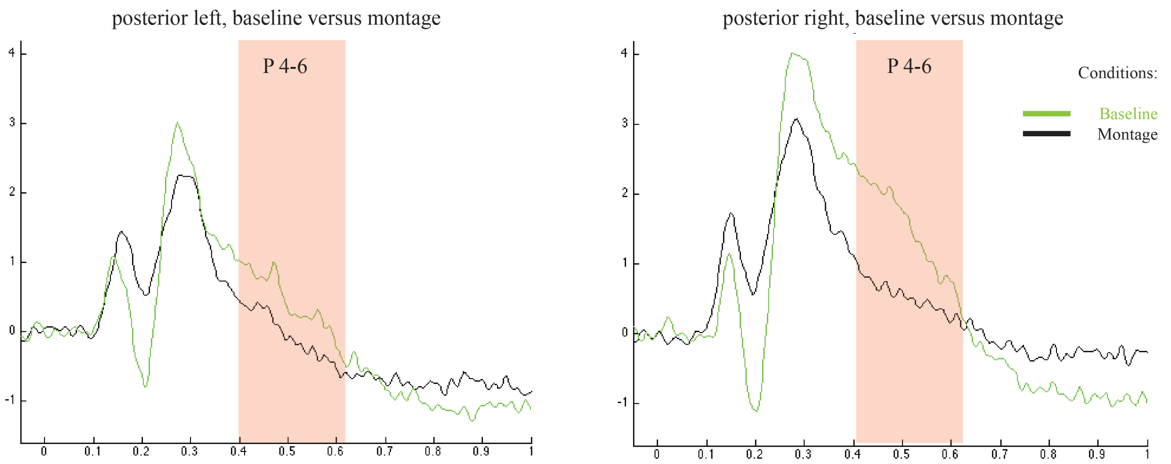


Figure 17: ERP time-course in posterior ROIs. Marked in red the P4-6 component pronounced for baseline. Condition x ROI interaction: $F(2,34)=4.19$, $p<0.05$; differences in posterior regions $p<0.05$.

Comparison Continuity Editing – Reverse shot

The analyses described above were repeated in a second step, now comparing ERPs after montage that involved either continuity edit or reverse shot.
For the first time window (140-190 ms), a 2x2x3 ANOVA (2 Conditions x 2 Hemispheres x 3 ROIs) showed a significant main effect only for ROI (F(2,34)=3.64, $p<0.05$). Post-hoc tests showed that average potentials in anterior regions were significantly lower than in posterior regions ($p<0.05$). No effect or interaction involving Condition was found.

For the second time window (180-220 ms), equally, a 2x2x3 ANOVA (2 Conditions x 2 Hemispheres x 3 ROIs) showed a significant main effect only for ROI (F(2,34)=36.04, $p<0.001$. Post-hoc tests showed that average potentials in anterior and central regions were significantly lower than in posterior regions (all $p<0.001$). No effect or interaction involving Condition was found.

For the third time window (250-380 ms) a 2x2x3 ANOVA (2 Conditions x 2 Hemispheres x 3 ROIs) showed significant main effects for ROI (F(2,34)=54,79, $p<0.001$) and significant interactions for Hemisphere x ROI (F(2,34)=5.12, $p<0.05$). Post-hoc tests showed that average potential in anterior and central regions was lower than in posterior regions (all $p<0.001$). Furthermore, for posterior regions, average potential was lower in the left hemisphere ($p<0.001$). No effect or interaction involving Condition was found.

For the fourth time window (400-650 ms), a 2x2x3 ANOVA (2 Conditions x 2 Hemispheres x 3 ROIs) showed significant main effects for Hemisphere (F(1,17)=6.25, $p <0.05$), ROI (F(2,34)=10.12, $p<0.001$), and a significant interaction of Condition x Hemisphere x ROI (F(2,34)=3.91, $p<0.05$). Post-hoc tests showed that average potential in left hemisphere was significantly lower than in right hemisphere ($p<0.05$). Furthermore, average potential in anterior regions was lower than in central and posterior regions ($p<0.001$). Most importantly for this comparison, continuity edit condition significantly differed from reverse shot condition in left anterior ($p<0.01$) as well as right central ($p<0.05$) ROIs. This supports: a) a significant difference between the two conditions in a late left anterior positivity (P4-6 ant left), which is more pronounced for the continuity edit condition; b) a late right central positivity (P4-6 cen right), which is more pronounced for the reverse shot condition. See **Figure 18**, next page.

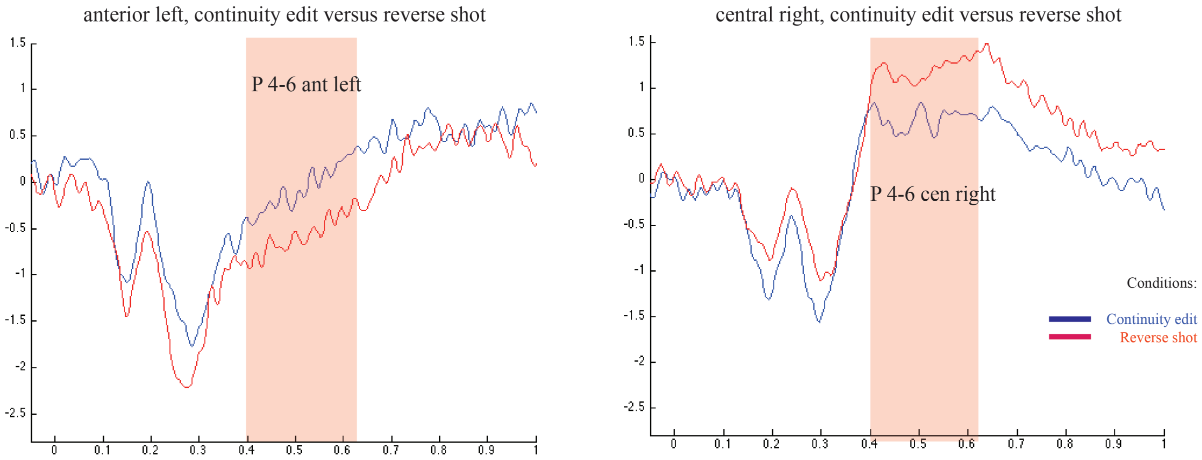


Figure 18: ERP time-course in anterior left and central right ROIs. Marked in red the P4-6, which in anterior left region is more pronounced for continuity edit condition, and in central right region is more pronounced for reverse shot condition. Condition x Hemisphere x ROI interaction: F(2,34)=3.91, p<0.05. Differences between conditions in anterior left region p<0.01, in central right region p<0.05.

III.2.4.2 Results EEG: ERD

Analysis ERD in observation and action execution condition

In order to generally assess mu-rhythm ERD in sensory-motor areas, we compared the frequency power for selected alpha and beta frequency ranges during baseline, with observation conditions (continuity edit and reverse shot condition) and the action execution condition.

Descriptives for alpha power in central regions showed that, compared to baseline, ERD was present in both observation conditions as well as during action execution (button press), with a maximum in the latter condition. A 4x2 ANOVA (Condition x Hemisphere) showed a significant main effect of Condition (F(3,48)=20.67, p<0.001). Post-hoc comparisons showed that frequency power for baseline was significantly higher than for all other conditions (p<0.001). Furthermore, frequency power for both observation conditions was significantly higher than for action execution (for continuity edit versus action execution p<0.01, for reverse shot versus action execution p<0.05). See **Figure 19**, next page.

Descriptives for beta power in central and frontal areas showed that compared to baseline, ERD was present in both observation conditions as well as during action execution. A 4x3 ANOVA (Condition x Region) accordingly showed a significant main effect for Condition (F(3,48)=25.63, p<0.001), for Region (F(2,32)=16.66, p <0.001) and a significant Interaction Condition x Region (F(6,96)=3.6, p<0.01) Post-hoc comparisons revealed that the significant differences for conditions again ocurred between baseline and all other conditions (all p<0.001), for Regions between the two central clusters and the frontal cluster (all p<0.001). Post-hocs for the Condition x Region Interaction showed that for left and right central electrodes as well as for frontal electrodes baseline beta power was significantly higher than the respective values for the observation conditions and for action execution (p<0.001). For right central electrodes we also found a significant difference between beta power in observation conditions and action execution (p<0.01, for central left electrodes this difference is not significant (p=0.55)). See **Figure 19** below.

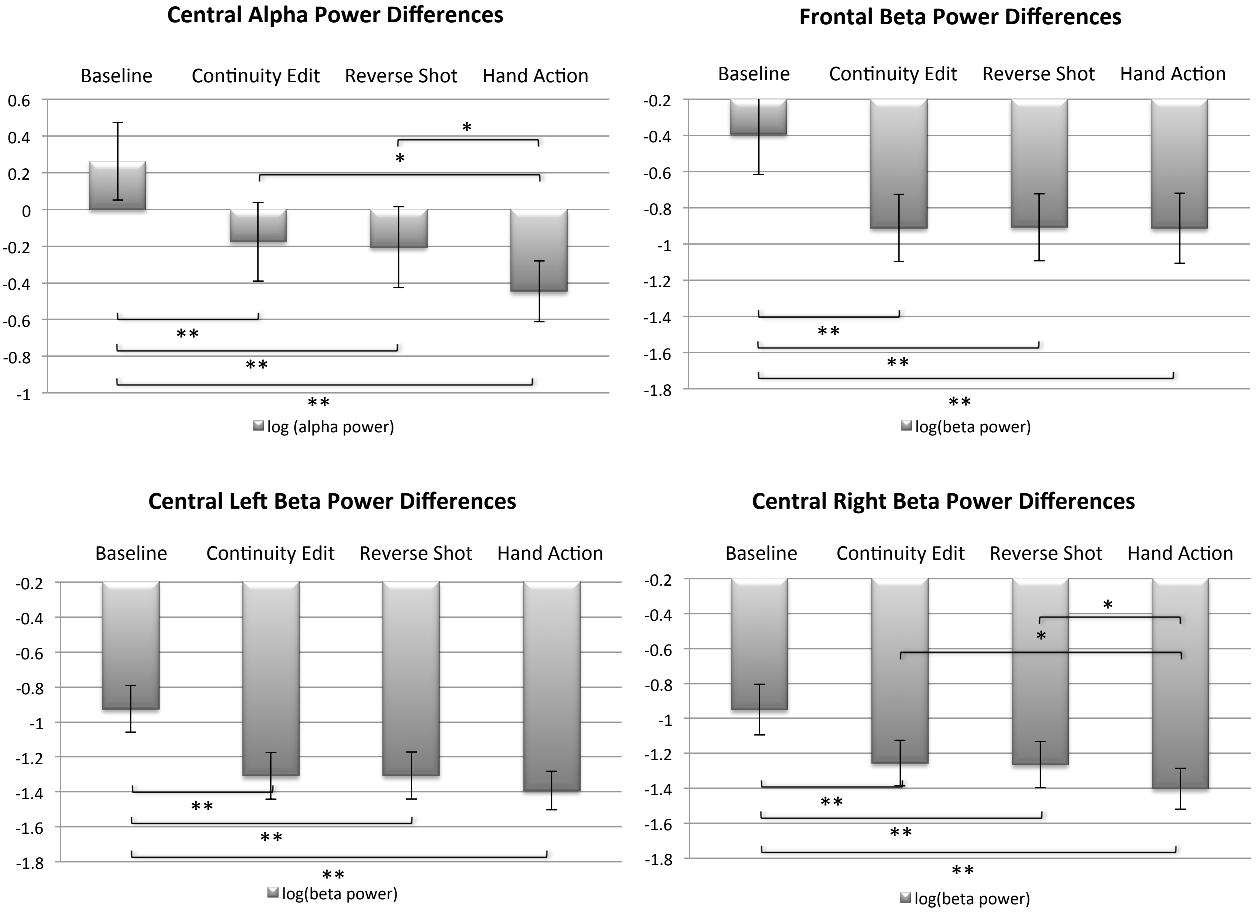


Figure 19: Results of first analysis assessing ERD of the three suggested components of the mu-rhythm during movie watching (action observation after edit) and during the participants' own hand action. Condition effect for alpha range: F(3,48)=20.67, p<0.001, for beta ranges: F(3,48)=25.63, p<0.001. Results show typical desynchronization (indicated by the significant difference with baseline) for action observation and action execution for all band ranges and regions (p<0.001), with significant differences between observation conditions and action execution for central alpha and central right beta power (p<0.01/p<0.05). Furthermore, beta bands in central regions show stronger desynchronization than in frontal regions: F(2,32)=16.66, p <0.001.

Due to the differences in the strength of desynchronization between frontal and central clusters, reflected in the effect for Region, the following analyses assessing ERD time-course were done separately for frontal and for central electrodes (with two levels of Hemisphere).

ERD-time-course comparison of continuity edit and reverse shot conditions

In order to assess the precise time-course of ERD/ERS in sensory-motor areas during the two different observation conditions we compared the log values of the frequency power of the chosen frequency ranges during the different observation conditions divided by baseline.

For central alpha power values the 2x2x5 ANOVA (2 levels of Condition (continuity edit and reverse shot montage), two levels of Hemisphere (left versus right), and 5 levels of Time (5 seconds from cut on)) showed a significant main effect of Time (F(4,64)=26.10, p<0.001) as well as a significant interaction Condition x Hemisphere (F(1,16)=6.22, p<0.05).

Post-hoc comparisons showed that for the time-course, significant differences occurred between all time windows during video observation (desynchronization phase) and all time windows after video observation (resynchronization phase) (p<0.01). Results regarding time-course differences are illustrated together with beta band results in **Figure 21**, see next page.

Post-hocs further investigating the significant Condition x Hemisphere Interaction showed that the significant differences occurred between left and right hemisphere for continuity edit conditions (with the left hemisphere being stronger desynchronized than the right, p<0.001) and between continuity edit and reverse shot conditions in the right hemisphere (with continuity edits being less desynchronized than reverse shots, p <0.01). See **Figure 20** below.

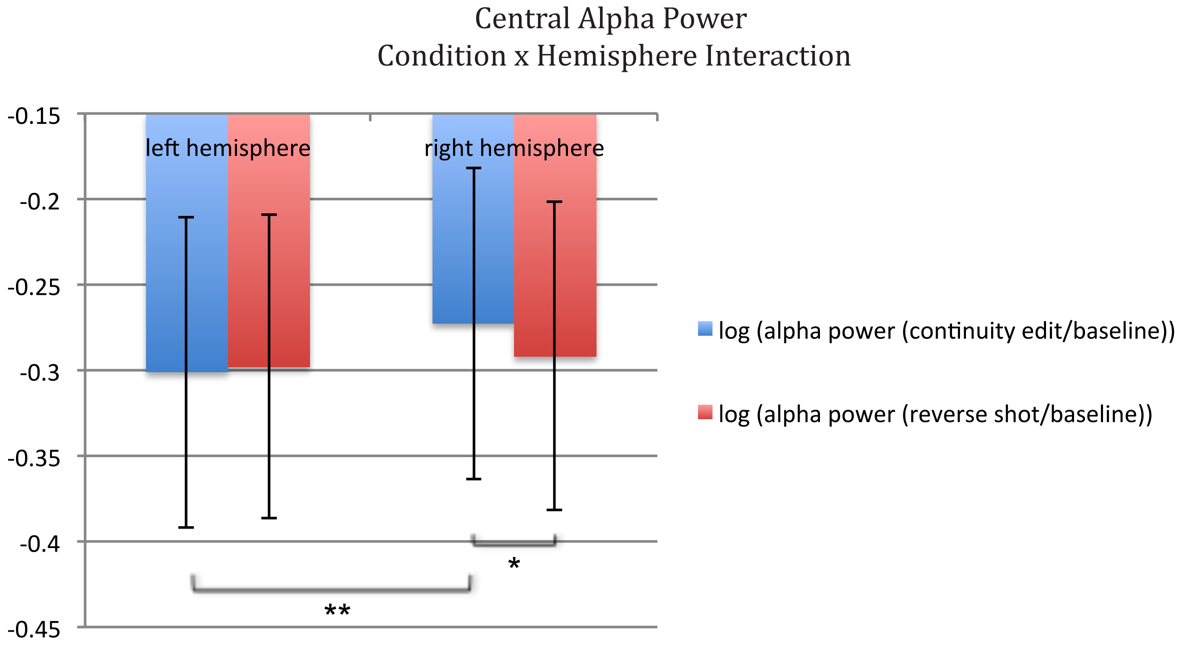


Figure 20: Central alpha power results regarding Condition x Hemisphere interaction: F(1,16)=6.22, p<0.05. Indicated are significant differences between alpha power for continuity edit condition in left and continuity edit condition in right hemisphere p<0.001, as well as between continuity edit condition and reverse shot condition in right hemisphere p<0.01.

For central beta power values the 2x2x5 ANOVA (2 levels of Condition (continuity edit and reverse shot condition), 2 levels of Hemisphere (left versus right), and 5 levels of Time (5 seconds from cut on)) showed a significant main effect of Time only (F(4,64)=23.53, p<0.001). Post-hocs showed that for the time-course the significant differences occurred between all time windows during video observation (desynchronization phase) and all time windows after video observation (resynchronization phase) (p<0.001). See **Figure 21** below.

For frontal beta power values, the 2x5 ANOVA (2 levels of Condition (continuity edit and reverse shot condition) and 5 levels of Time (5 seconds from cut on)) showed a significant main effect of Time only (F(4,64)=17.36, p<0.001)

Post-hocs showed that for the time course the significant differences occurred between all time windows during video observation (desynchronization phase) and all time windows after video observation (resynchronization phase) (all p<0.001). See **Figure 21** below.

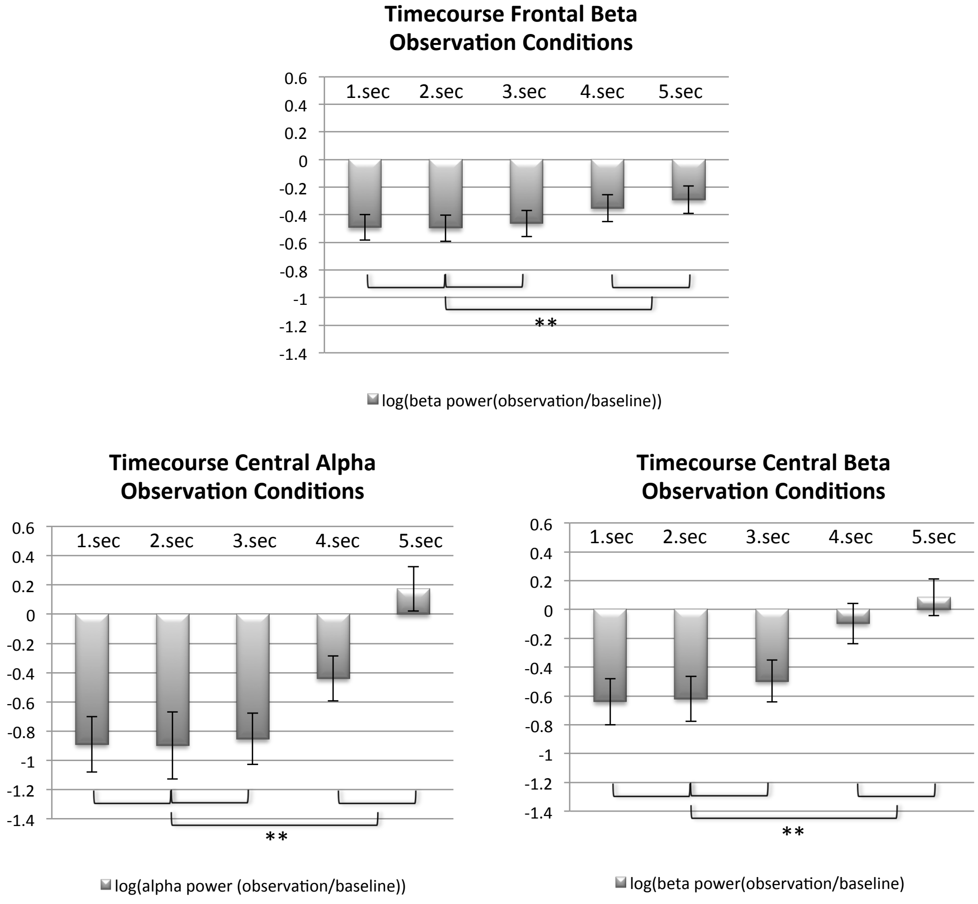


Figure 21: Time-course differences in central alpha, central beta and frontal beta power over 5 seconds after an edit (3 seconds of movie showing goal related action, 2 seconds resynchronization phase). Time effect in central alpha: F(4,64)=26.10, p<0.001, central beta: F(4,64)=23.53, p<0.001, frontal beta: F(4,64)=17.36, p<0.001. Differences between desynchronization and resynchronization periods were significant in all bands and regions (p<0.01 for central alpha, p<0.001 for central and frontal beta).

Control in occipital regions

To control for similar effects in occipital regions we repeated the analysis just described in 1) and 2) in occipital electrodes. The 4x2 ANOVA with 4 factors of Condition (baseline, continuity edit, reverse shot and action execution) and 2 factors of Hemisphere in the occipital region showed a main effect of Condition only (F(3,48)=21.44, p<0.001).

Post-hoc comparisons showed that frequency power for baseline was significantly higher than for all other conditions (p<0.001). There was no difference between frequency power for observation conditions and action execution. See **Figure 22** below.

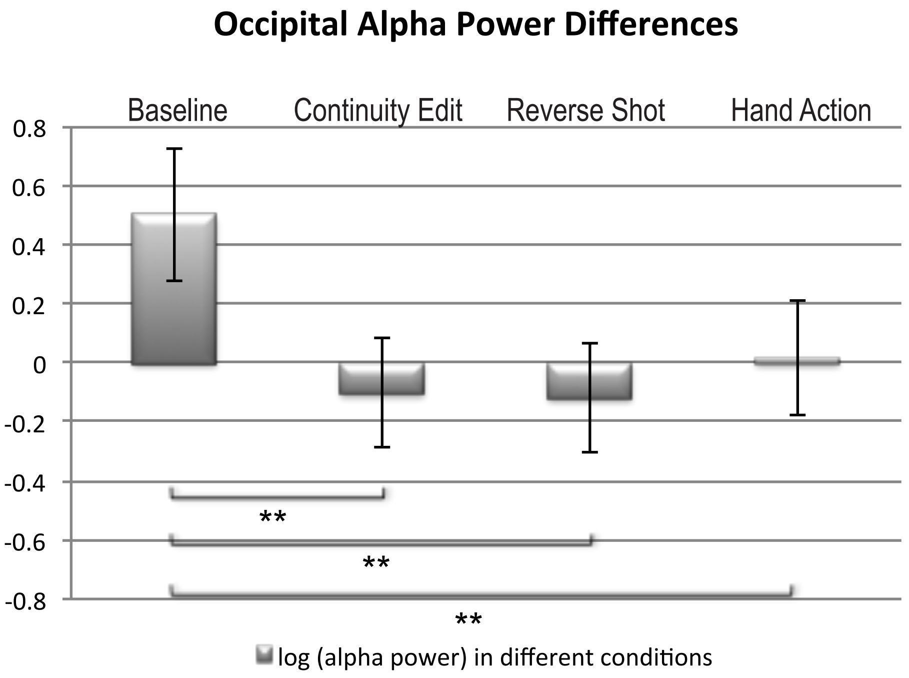


Figure 22: Results of a control analysis assessing occipital alpha power ERD. Condition effect: F(3,48)=21.44, p<0.001. Differences between baseline and all other conditions were significant with p<0.001.

The 2x2x5 ANOVA (2 levels of Condition (continuity edit and reverse shot), 2 levels of Hemisphere (left versus right) and 5 levels of Time (5 seconds from cut on)), investigating differences between the two observation conditions, showed a significant main effect of Time only (F(4,64)=23.93, p<0.001).

Post-hocs showed that for the time course the significant differences occurred between all time windows during video observation (desynchronization phase) and all time windows after video observation (resynchronization phase) (all p<0.05). See **Figure 23**, next page.

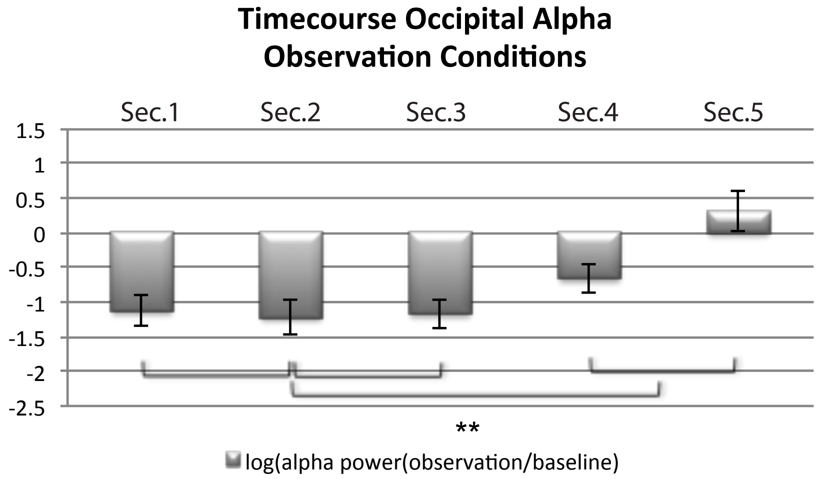


Figure 23: Results of a control analysis assessing time-course of occipital alpha power ERD. Condition effect: F(4,64)=23.93, p<0.001. Significant differences occur between desynchronization and resynchronization periods: p<0.05.

III.2.4.3 Results Rating Task

Only the results of question 4 of the rating task revealed significant differences between observation conditions. For the question "How natural did you find the representation of the scene in this video?" participants rated videos with a continuity edit as more natural in their representation of the scene than videos including a reverse shot (N=17, T= 25, Z= 2,22, p<0.05). See **Figure 24** below.

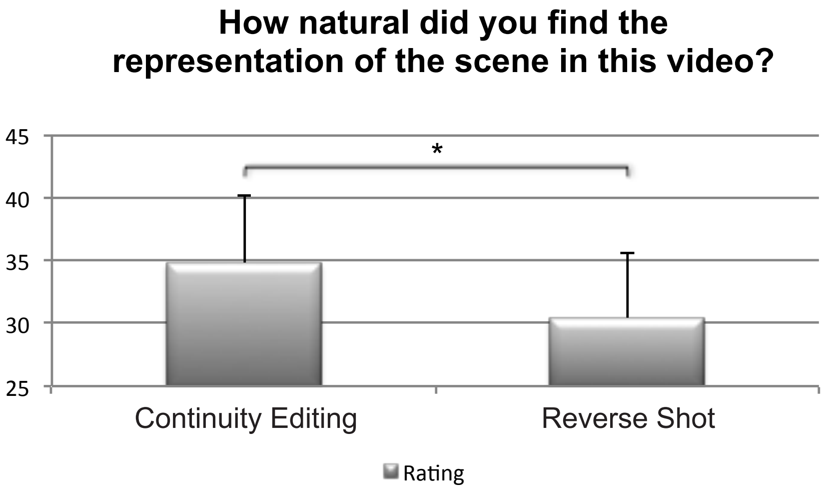


Figure 24: Results of the rating task, question: "How natural did you find the representation of the scene in this video?", showing that movie clips containing a continuity edit were judged as more naturally representing the scene than movie clips containing a reverse shot edit. Results of Wilcoxon Signed Ranks: N=17, T= 25, Z= 2,22, p<0.05.

The following tablet provides means and standarderrors for all significant results reported.

MEANS AND STANDART ERRORS RESULTS

ERP RESULTS	**Mean**	**STE**
TIME WINDOW 1 (140-190ms)		
Comparison Baseline-Montage Conditions		
Significant potential differences for Condition x Hemisphere Interaction		
Left hemisphere, Baseline	0.037	0.047
Left hemisphere, Montage Conditions	-0.164	0.067
Significant potential differences for Condition x ROI Interaction		
Anterior Regions, Baseline	-0.006	0.081
Anterior Regions, Montage Conditions	-0.212	0.092
Comparison Continuity Edit – Reverse shot Conditions		
Significant potential differenced for ROI effect		
Anterior Regions	-0.227	0.106
Posterior Regions	0.174	0.076
TIME WINDOW 2 (180-220 ms)		
Comparison Baseline-Montage Conditions		
Significant potential differences for Condition effect		
Baseline	-0.006	0.042
Montage Conditions	-0.123	0.042
Significant potential differences for ROI effect		
Anterior Regions	-0.666	0.241
Central Regions	-0.504	0.126
Posterior Regions	0.977	0.238
Significant potential differences for Condition x ROI Interaction		
Anerior Regions, Baseline	-0.245	0.296
Anterior Regions, Montage Conditions	-1.087	0.241
Posterior Regions, Baseline	0.61	0.325
Posterior Regions, Montage Conditions	1.34	0.198
Comparison Continuity Edit-Reverse Shot Conditions		
Significant potential differences for ROI effect		
Anterior Regions	-1.086	0.209
Central Regions	-0.625	0.12
Posterior Regions	-1.344	0.199
TIME WINDOW 3 (250-380ms)		
Comparison Baseline-Montage Conditions		
Significant potential differences for ROI effect		
Anterior Regions	-1.205	0.202
Central Regions	-0.72	0.132
Posterior Regions	1.675	0.227
Comparison Continuity Edit-Reverse Shot Conditions		
Significant potential differences for ROI effect		
Anterior regions	-1.324	0.216
Central regions	-0.784	0.142
Posterior regions	1.719	0.178

Significant potention diffrences for Hemisphere x ROI Interaction		
Posterior left regions	1.451	0.159
Posterior right regions	1.985	0.277
TIME WINDOW 4 (400-650ms)		
Comparison Baseline-Montage Conditions		
Significant potential differences for Hemisphere effect		
Left hemisphere	-0.212	0.092
Right hemisphere	0.54	0.086
Significant potential differences for ROI effect		
Anterior regions	-0.678	0.165
Central regions	0.233	0.191
Posterior regions	0.938	0.221
Significant potential differences for Condition x Hemisphere Interaction		
Left hemisphere, Baseline	-0.343	0.132
Left hemisphere, Montage Conditions	-0.081	0.113
Right hemisphere, Baseline	0.714	0.14
Right hemisphere, Montage Conditions	0.367	0.077
Significant potential differences for Condition x ROI Interaction		
Posterior Regions, Baseline	1.299	0.319
Posterior Regions, Montage Conditions	0.578	0.184
Comparison Continuity Edit – Reverse shot Conditions		
Significant potential differences for Hemisphere effect		
Left hemisphere	-0.097	0.12
Right hemisphere	0.408	0.099
Significant potential differences for ROI effect		
Anterior regions	-0.427	0.139
Central regions	0.554	0.169
Posterior regions	0.38	0.188
Significant potential differences for Condition x Hemisphere x ROI Interaction		
Anterior left, Continuity Edit	-0.571	0.171
Anterior lef, Reverse Shot	-1.104	0.291
Central right, Continuity Edit	0.514	0.252
Central rigtht, Reverse Shot	0.892	0.25
ERD RESULTS	**Mean**	**STE**
First analysis central alpha		
Significant lot(power) differences for Condition effect		
Baseline	0.262	0.21
Continuity Edit	-0.176	0.214
Reverse Shot	-0.206	0.22
Action Execution	-0.446	0.17
First analysis central and frontal beta		
Significant lot(power) differences for Condition effect		
Baseline	-0.776	0.124
Continuity Edit	-1.159	0.109
Reverse Shot	-1.16	0.114
Action Execution	-1.236	0.095

Significant log(power) differences for Region effect		
Central left	-1.248	0.123
Central right	-1.218	0.127
Frontal	-0.782	0.1
Significant log(power) differences for Condition x Region interaction		
Baseline left central	-0.983	0.142
Continuity Edit left central	-1.309	0.132
Reverse Shot left central	-1.307	0.134
Action Execution left central	-1.392	0.11
Baseline right central	-0.95	0.144
Continuity Edit right central	-1.256	0.131
Reverse Shot right central	-1.264	0.133
Action Execution right central	-1.392	0.11
Baseline frontal	-0.394	0.14
Continuity Edit frontal	-0.912	0.098
Reverse Shot left frontal	-0.908	0.105
Action Execution left frontal	-0.913	0.094
First analysis occipital alpha		
Significant log(power) differences for Condition effect		
Baseline	0.501	0.224
Continuity Edit	-0.104	0.186
Reverse Shot	-0.118	0.185
Action Execution	0.017	0.194
Time course analysis, comp. Continuity edit and reverse shot – central alpha		
Significant log(power) differences for Time effect		
1st second	-0.445	0.09
2nd second	-0.449	0.11
3rd second	-0.427	0.08
4th second	-0.22	0.07
5th second	-0.086	0.07
Significant lot(power) differences for Condition x Hemisphere interaction		
Continuity Edit, left hemisphere	-0.3	0.082
Continuity Edit, right hemisphere	-0.273	0.079
Reverse Shot, left hemisphere	-0.298	0.08
Reverse Shot, right hemisphere	-0.291	0.078
Time course analysis, comp. Continuity edit and reverse shot – central beta		
Significant log(power) differences for Time effect		
1st second	-0.323	0.074
2nd second	-0.317	0.075
3rd second	-0.251	0.064
4th second	-0.053	0.065
5th second	-0.042	0.058
Time course analysis, comp. Continuity edit and reverse shot – frontal beta		
Significant log(power) differences for Time effect		
1st second	-0.502	0.093
2nd second	-0.516	0.096
3rd second	-0.47	0.097
4th second	-0.362	0.099
5th second	-0.301	0.102

Time course analysis, comp. Continuity edit und reverse shot – occipital alpha		
Significant lot(power) differences for Time effect		
1st second	-0.565	0.11
2nd second	-0.612	0.12
3rd second	-0.586	0.102
4th second	-0.333	0.102
5th second	-0.152	0.146
Rating Task Results		
Question „How natural did you find the representation oft he scene?“		
Continuity Edit	34.912	5.313
Reverse Shot	30.257	5.221

III.2.5 Summary and Discussion of Results

ERP results

For the first time window (140-190 ms) we found significantly lower potentials for montage conditions than for baseline in anterior and left hemispheric regions. Together with the descriptive information of the Grand Average ERP this shows an amplitude increase of an early anterior negativity (N1) for montage conditions, especially pronounced in the left hemisphere. Early anterior negativities have previously been described as appearing in response to structural irregularities in language as well as in music and action observation. In the language domain they appear to be correlated with the initial syntactic processing of local dependencies. They often show a left hemispheric bias and, if task relevant, were described to be followed by a late posterior positivity (Hahne and Friederici 1999; Friederici et al., 1993; Friederici 2002, Grodzinsky and Santi, 2008; Steinhauer & Drury, 2012). In studies in the domain of music in contrast, a right hemispheric bias for the early negativity has been observed, while the following late positivity was again localized in posterior regions (Koelsch, 2011). Most importantly, in recent studies investigating content and structure violations in action observation, it was reported that structure violations correlate with a pronounced early negativity and a late positivity both localized in anterior left regions (Maffongelli et al. (submitted)).

Our results of the first time window thus suggest that both the visual onset of a video as well as the visually different stimulus after a cut elicit an ERP associated with the detection of a syntactic violation. The significant difference between baseline and montage condition, however, indicates that the violation caused by the within-scene cuts is perceived as being much stronger. Interestingly, no significant differences between the two Montage conditions (continuity edit versus reverse shot) were found in this time window, indicating that the quantitatively/qualitatively different visual interruption caused by these two conditions (continuity edit versus reverse shot) is not traced by this component.

For the second time window (180-220 ms) we found significantly higher potential for baseline than for montage conditions in anterior regions, as well as significantly lower potential for baseline than for montage conditions in posterior regions.

Together with the descriptive information of the Grand Average ERP this shows an amplitude increase of an early anterior positivity (P2) as well as early posterior negativity (N2) for baseline. Previous research has correlated raising amplitude sizes of P2 with decreasing stimulus ambiguity. Specifically, Kommeier et al. (2012) reported that when participants watch ambiguous stimuli (as the Necker lattice) as well as slightly disambiguated stimuli of the same kind, an early centrofrontal positivity around 200 ms after stimulus onset and a late centroparietal positivity around 400 ms after stimulus onset can be observed, with the amplitudes of both increasing with decreasing ambiguity. The change in vision produced by a stimulus onset is a clear event of rather non-ambiguous nature. The change caused by a cut involving the montage of two shots filmed from a different camera position, in contrast, causes what can be described as an ambiguous event. Precisely, the same scene is still observed though the camera has been displaced, that is, the visual continuity is broken. It is possible that these kinds of stimuli elicit a response precisely signaling an ambiguity as some visual features are still kept, while others have changed.

Interpretations of the differences found in posterior regions are reminiscent of a study by Proverbio & Riva (2009). In this study, which explores ERP responses to static visual representations (photos) of actions, the authors found a pronounced posterior negativity, around 250 ms after stimulus onset, elicited by all stimuli representing comprehensible in comparison to incomprehensible actions. The differences in our data might thus stem from the fact that montages in general cause a slight disturbance in action comprehension, traced by the modulation of the reported component.

Again, interestingly, also in the second time window, no difference between continuity edit and reverse shot condition was found, indicating that within early processing there are no differences between the two conditions.

For the third time window (250-380 ms) no differences were found either between baseline and montage conditions or between continuity edit and reverse shot montage conditions. Nevertheless, the Grand Average ERP for this time window showed a negativity in anterior regions and a positivity in posterior regions. Fronto-central negativities around 300 ms after stimulus onset have been reported to be associated with the detection of faces (Barrett, Rugg, & Perrett, 1988; Debruille, Pineda, & Renault, 1996; Debruille, Brodeur, & Porras 2012). Faces are indeed present in all of our stimuli conditions. Positivities (in temporal/parietal regions) around 300 ms after stimulus onset have been associated with working memory activities, indicating a perceived change of the environment (Donchin & Coles, 1988; Patel & Azzam, 2005; Polich, 2007). The fact that we did not find significant differences between conditions here is likely to indicate that face detection and working memory access do not differ among our conditions.

For the fourth time window (400-650 ms) we found a significantly higher potential for baseline than for montage conditions in posterior regions. Together with the descriptive information of the Grand Average ERP this shows an amplitude increase of a late posterior positivity (P4-6) for baseline. As reported above, late posterior positivities have been described in language research to correlate with structure violations. However, it was also reported that strong content violations do correlate with a pronounced posterior P4-6 (Meerendonk et al., 2010). It can therefore be hypothesized that the strong visual change from black screen to video onset elicits the registration of a semantic rather than syntactic violation leading to a contextual update indicated by a pronounced posterior P4-6. Another possibility might be that the observed P4-6 is the one reported to appear after non-ambiguous stimuli (see results for the second time window), though in the studies of Kommeier et al. (2012) this component was found in central-parietal regions rather than in posterior regions.

Comparing continuity edit and reverse shot conditions within this time window we found two significant differences: We measured significantly higher potential (P4-6) for continuity edit condition than for reverse shot condition in anterior left regions. As reported above, recent studies exploring content and structure violation in action observation observed a late P4-6 in left anterior regions following syntactic violations (indicated by early left anterior negativities), associated with postperceptual processes possibly serving an adjustment to the detected violation, most likely helping to overcome the change without reaching visual awareness. Indeed, in our data both montage conditions show an early anterior negativity, though this component is not different in amplitude when comparing continuity editing and reverse shot. Interestingly then however, the later P4-6 component indicating updating/ reanalysis is stronger for the continuity editing condition, possibly indicating a suppression of the adjusting response for reverse shots. In contrast to this, in central right regions we found a significantly higher potential (P4-6) for the reverse shot condition when compared to the continuity editing condition. Interestingly, previous ERP studies investigating the neural correlates of change blindness found a pronounced late positivity (300-700 ms) over central parietal lobes only when a change was detected (Niedeggen, Wichmann, & Stoerig, 2001; Koivisto & Revonsuo, 2003). Koivisto & Revonsuo suggested that these changes are likely not to be associated to phenomenal visual awareness, but to postperceptual processes such as conscious evaluation of change and decision making. With reference to Block (2001) they write “this later positivity may be related to access consciousness or reflexive consciousness (a special kind of access). It does not correlate to the subjective experience of seeing but to perceiver's other beliefs about the experience with the seen event." (Koivisto & Revunsuo, 2003, p.428)

Interestingly, taken together these findings can further illuminate, and on the other hand be informed by a prior fMRI study on the topic, done by Magliano and Zacks in 2011. Using film sequences taken out of a real movie Magliano and Zacks distinguished BOLD responses regarding three different types of stimuli: a) Film sequences including montages that, according to the descriptions of the Magliano and Zacks, supposedly do not violate time and space or action continuity as the edited shots show the same ongoing scene filmed from different camera positions but not violating the 180° rule (continuity edits); b) Film sequences including montages that supposedly violate time and space continuity as the edited shots show the same scene but the difference in position of the filming camera violates the 180° rule (reverse shots); c) Film sequences including montages that clearly violate space, time and action continuity by editing shots showing completely different scenes. BOLD responses registered during the watching of continuity edits revealed a transient increase in networks associated with early visual processing (inferotemporal cortex, posterior superior temporal sulcus and precentral sulcus) as well as in the precuneus right after the cut, which the authors interpreted as stimulus-driven processing serving the remapping of visual features across the editing boundaries. On the other hand, for montages violating spatial and/or temporal discontinuities by crossing the 180° axis, fMRI data revealed a contemporary transient bilateral increase of activation in the parahippocampal cortex, interpreted by the authors as a stimulus-driven response to the reverse shot. Moreover, for the same condition, they also observed transient decreases of activation in early visual cortices that they suggested reflects attention-driven down regulation of activity in regions involved in the adjustive processing of the changed dimensions, thus being responsible for the perceptual salience of the cut. Lastly, for clear scene breaks they found a reduction of activation in the lateral parietal cortex, a region formerly associated with the representation of action goals, possibly serving a complete contextual update due to the salient difference of the new stimulus (Maglioni & Zacks, 2011).[42]

---

[42] Notably, Magliano and Zack's interpretations are driven by Event Segmentation Theory (EST). EST models how the brain manages to dissect the continuous flow of perceptual input into hierarchically organized parts and sub-parts. Important in this context are prior findings demonstrating that in general the brain's responses to discontinuities depends on the degree of spectators' attention towards to the violations they represent (see Grill-Spector & Malach, 2001, for stimulus driven discontinuities and Maunsell & Treue, 2006 for attention driven discontinuities). While changes of a stimulus in general have been shown to increase activity in regions associated with the processing of the changed dimensions, hence likely helping the

fMRI as a technique allows excellent spatial resolution but cannot provide precise temporal information, as is accessible via EEG recordings. Furthermore, previous EEG research has accumulated substantial knowledge associating certain components with certain cognitive functions, that might be of crucial help in further investigating the issues at hand.

To be precise, our findings could not only support but also refine Maglioni & Zacks hypothesis that moving images are processed by the brain using networks for action observation processing, crucially depending on real world experience. Indeed, our findings indicated that these networks seem to be sensitive to violations of constancy expectations, represented in the ERP by components previously associated with syntactic violations during the perception of action sequences. Furthermore, regarding the temporal dimension of the processes involved, our study revealed that earliest stimulus detection processes, at least as far as reflected by the modulation of the ERP components observed, do not differ significantly between continuity edits and reverse shots. In contrast to that, the modulation of later components, as observed in our study, does show differences between montage conditions, thus supporting Magliano & Zacks hypothesis that late postperceptual processing makes up for the different experiences connected with continuity edits and reverse shots. To be precise, our findings indicate, just as suggested by Magliano and Zacks, that continuity edits elicit postperceptual processes probably helping to adjust to the new visual inputs, preventing a distortion of the visual flow. Futhermore, our findings additionally showed that reverse shots seem to activate access to reflexive consciousness, possibly directing awareness from the content of the image to its representation, thus the conscious perception of the cut. We will come back to these findings in the integrative discussion of ERP and ERD results below.

ERD results

ERD results in general show the expected desynchronization of the mu-rhythm during action observation and execution in both alpha and beta bands. More specifically, when analysing band power of both bands in selected electrodes during the second second of action observation as well as in the second surrounding participants' button press, a significant desynchronization during action observation as well as, more pronounced, during action

---

perceiver to bridge the discontinuity, stronger stimuli changes across cuts have been reported to produce attention driven decrease of activation of the same regions, hence possibly not leading to an assimilation of the change but instead to a further enhanced salience of stimulus differences and, consequently, to the perception of an event break. Magliano and Zacks (2011) hypothesize that precisely these principles underlie the 180° rule, with continuity edits depending on spatial remapping processes, which for reverse shots as well as for real scene breaks are suppressed.

execution was found. In contrast to this, in occipital alpha no difference between observation and execution conditions was found, supporting the notion that alpha ERD in occipital cortices is discriminable from alpha ERD in motor areas, with alpha ERD in occipital cortices showing only effects likely due to attention.

Further investigation of the whole time course of mu-rhythm desynchronization and resynchronization for observation conditions (that is, three seconds of movie observation following the point of montage and two seconds of the subsequent ITI), in beta bands did not show any further difference between conditions (continuity edit versus reverse shot montage). In contrast, in alpha bands a significant Condition*Hemisphere interaction was detected. To be precise, post-hoc tests revealed that for continuity edits the significant difference was between hemispheres, with the left hemisphere showing stronger ERD than the right hemisphere. In contrast, for reverse shots, ERD was equal in both hemispheres, leading to a significant difference between continuity edit condition and reverse shot condition in the right hemisphere.

Previous studies demonstrated that the mu-rhythm shows ERD during voluntary action execution as well as action observation (Derambure et. al., 1993; Stancak and Pfurtscheller, 1996; Toro et al., 1994; Pfurtscheller et al., 1999). Moreover, it has been reported that in alpha bands, mu-rhythm suppression during the observation of hand actions is larger in the hemisphere contralateral to the moving hand (Perry & Bentin, 2009; for respective findings in fMRI see also Shmuelof & Zohary, 2005). When interpreting our findings we would like to reiterate that a camera displacement of 180° (reverse shot) causes a mirroring of the original spatial relations of the viewer to the scene: what has been presented as left before, becomes right after the cut. We suggest that this might cause a short orientation-deficit, disturbing the clear identification of the observed hand actually executing the action. In consequence, for reverse shots no clear contralateral lateralization of the ERD might occur, but rather both hemispheres should show equally strong ERD – as indeed was found in the present experiment. Again, analyses of alpha power in occipital regions do not reveal comparable results, supporting the notion that alpha ERD in occipital cortices is discriminable from alpha ERD in motor areas.

We would furthermore like to point out that it is possible that these diverging activations have an influence on further processing of the stimulus as indicated by our ERP results and previous studies of Magliano & Zacks (2011). Indeed, the location of the late P4-6 in the reverse shot condition, lying in central right electrodes, is reminiscent of the differences found in our analysis of the ERD of the central mu-rhythm, precisely happening in central right regions. Future research needs to adress the possibility that the different processing of stimuli might be due to deviating processes in motor rather than visual regions.

Also, it has to be clearly stated at this point that these results do not match our original hypothesis predicting a stronger ERD for stimuli that better resemble vision as known from real world experience (that is, continuity edits). This might be due to several reasons. Firstly, it could be argued that discontinuity edits, when compared to continuity edits, elicit an increase of attention in general, influencing also the action perception links investigated via analysis of the mu-rhythm. Such attentional drain could make up for the differences expected, that is enhancing the activation of the motor cortex during action observation, originally assumed to be attenuated due to the artificiality of the representation. Indeed, previous studies reported that the mirror mechanism can be enhanced by attentional processes (see Bach, Peatfield & Tipper, 2007 for behavioral results and Chong, Williams, Cunnington, & Mattingley, (2008) for an fMRI study). Furthermore Reeves, Thorson, Rothschild, McDonald, Hirsch, & Goldstein, (1985) by comparing potential of occipital alpha as registered during EEG recording, reported that attention was directed toward the source of a perceived discontinuities. These findings can also be supported by studies showing enhanced accuracy and faster response times in recognition memory tasks for information originally presented after a discontinuity (Carroll & Bever, 1976). Indeed, considering this alternative, the heightened activation of the motor cortex during the watching of videos produced by the steadicam could also be attributed to this raise of attention due to the additional movement in the scene. However, such an attentional drain should be traceable by occipital alpha showing the same differences as central alpha. Remarkably, this control analysis did not show such a result in either of the two studies conducted.

In consequence I would like to propose another explanation for the surprising lack of additional activation of the motor cortex after continuity edits. In comparison to camera movements enabling moving images to closely resemble vision in the real world, continuity edits might only serve to produce moving images not crucially violating habits from real world perception. Indeed, one may doubt that the explanation that our own bodily movements in the world could lead to visual impressions that closely resemble cuts in the movies, even when talking about continuity edits. Of course blinks and saccades might elicit interruptions of vision reminiscent to cuts. However, it feels as if this insight takes much more time and reflexion on when and how we move, blink and saccade and how this could lead to similar experiences as a cut, than considering about how a steadicam produces a movement close to our own. Indeed, wouldn't be the equivalent of a steadicam for montage be the use of a steadily moving camera whose lens sometimes gets shut by a lid, just as we sometimes blink? Most likely however, the effect of such an artificial blink, does not correspond to most edits, not even those complying to the 180° rule!

Rather than actually "simulating" vision in the real world, we might therefore consider continuity edits as being merely "less violating" it. This would be also supported by the fact that the first component (N1), indicating the detection of structural violation is indeed modulated in the same way by both edits AND followed for continuity edits by a component suggested to serve overcoming of this violation. In consequence, only camera movements would have the possibility to help creating moving images closely resembling vision in the real. This might also mean, that only camera movements would be able to crucially enhance the mirror mechanism, possibly fundamentally underlying the spectator's immersion in the movie. Clearly, future research has to find ways to further illuminate this issue.[43]

I will come back to this thought in the last chapter of this thesis.

Integrative discussion of ERD, ERP and behavioural data

With our study we aimed to illuminate how different kind of editing techniques affect spectators' film experience. Reflections on film influenced by 4EA approaches to cognition led to the hypothesis that film's obvious power to completely involve the spectator in a fictional world might be due to an activation of all habits and experiences of an embodied visual activity (see also Chapter II.3 and II.4 of this thesis). Furthermore, such activation might depend on moving images closely resembling visual perception in real life. A prominent line in cognitive film theory (see Chapter II.3 of this theses) has repeatedly described continuity editing as being a system creating such kind of smooth images. Also based on the findings reported in our study about camera movements (see Chapter III.1), we hypothesized that 1) edits in general should be detected, in contrast to the plain onsets of new visual stimuli, as violations of continuously perceived action sequences. Precisely, for the special case investigated, that comprised edits produced via filming the same scene from different angles, we predicted the detection of structural violations regarding the actions perceived; 2) continuity edits and discontinuity edits should differ at least in later stages of (post-perceptual) processing with the precise differences reflecting the degree of deviation from real world perception as well as spectators' different awareness of the edits; and finally 3)

---

[43] One idea to further test the immersive effect of camera movements might be to use a "priming" of the spectator: If camera movements can really help spectator's immersion into the movie, an establishing shot produced via the use of the steadicam might be able to attenuate the differences between continuity edits and reverse shot as observed in our study.

continuity edits should enhance motor cortex activations during action observation, as such action perception links were suggested to play a major role in feeling the presence of and reacting to a world on distance.

Indeed, our findings firstly showed that in contrast to plain onsets of a new visual stimulus, within-scene cuts correlated with modulations of early components previously associated with structural violation, ambiguous character and hard comprehensibility. Remarkably, no difference between continuity edits and reverse shots was found for these components. These results in fact support the notion that within-scene cuts in general are detected in their violation of certain expectations regarding spatial/temporal/causal constancy, which are derived from real life. More concretely, they suggest that the visual changes evoked by a cut are categorized as structural/syntactic violations of the action perceived.

Next, our results showed a late anterior left positivity more pronounced for continuity edits than for reverse shots. In previous research such a component has been found to occur after an early left anterior negativity, which was associated with the detection of a structural violation in visually presented action sequences if this violation was task relevant. The late anterior left positivity was therefore interpreted as possibly representing updating processes that serve the reanalysis of the scene (see Maffongelli et al., submitted). The absence of this component for reverse shots in our study supports Zacks and Magliano's hypothesis of the importance of such later remapping processes for spectators' experiences of the differente edits and their interpretation, which is that such processes might be suppressed for montages violating continuity editing rules. Furthermore, we found a late central right positivity (P4-6) more pronounced for reverse shot montages, possibly indicating the access of reflexive consciousness, leading to the awareness of the perceptual process rather than its content – thus a rupture of spectators' immersion in the movie.

Regarding ERD data, for continuity edits we found a stronger ERD of the mu-rhythm in the left hemisphere, likely to reflect a selective response to the (right) hand moving in the observed video. On the contrary, for reverse shot condition, left and right hemispheres showed equal ERD. We suggested that this might be due to a short orientation-deficit following the strong spatial violation caused by the reverse shot, that render left to right and right to left. In consequence, for reverse shots no clear contralateral lateralization of the ERD might occur, but rather both hemispheres should show equally strong ERD – as indeed it was found in the present experiment. We also suggested that there might be a link between this lack of lateralization and the different postperceptual processing of reverse shots described before.

Remarkably, contrary to our initial hypotheses, ERD responses did not differ in their general strength between the two montage conditions. We proposed that this possibly indicates that continuity edits play less of an activating and more of a sustaining role with regard to action perception links possibly involved in spectators' film experiences. This can be corroborated by reflections about how similar edited moving images can in fact get to human vision. Indeed, it is relatively intuitive to assume that a camera carried by a human (thus a steadicam), records images closely resembling such a humans vision. However, though we might try to think about cuts as blinks or suppression of vision during saccades, these physiological phenomena are not enough to explain the jump in time and space elicited by edits – even when these are produced with compliance to the rules of continuity editing.

Indeed, the consequent interpretation that continuity edits should be considered as "less violating" rather than "simulating" vision in the real world can also be supported by the results of our rating task which indicated that participants experience continuity edits as producing a more natural representation of the observed scene than that achieved by reverse shots, but that did not support any differences between the different representations regarding a feeling of involvement in the scene etc. Nevertheless, the critique against using a rating task of the kind performed in our experiment as a method to access the involvement in the scene, as uttered in Chapter III.1.3.4, still holds, and asks for future research finding a better approach to this question.

Taken together these findings indicate that indeed film seems to be perceived in activation of habits and experiences of an embodied vision, developed in interaction with the real world. This is supported by our data, which shows that visual ruptures such as those caused by edits are not perceived as plain onsets of a new visual stimulus, but rather as structural violations of a continuously perceived action. Moreover, images closely matching the circumstances of vision in the real world, including edits according to the rules of continuity editing, seem to be processed via the activation of adjustment processes in later visual processing. Such processes possibly allow an undisturbed involvement in the movie, that is, the experience of the fictional world just as an active embodied activity exploring it. On the other hand, images vehemently violating the circumstances of vision as known from the real world, as for example reverse shots not complying to the 180° rule of continuity editing, seem to elicit brain responses associated with the access of reflexive consciousness. In consequence spectators might be aware of the violation, and with that of the movie as a representation with its own narrative means. Such processes simultaneously might cause a rupture of spectators' movie experience, that is, the undisturbed experience of the content and plot, rather than the experience of the film as a medium. Exactly this experience is what is reported by cognitive film theory as the typical effect of reverse shots, if not embedded in circumstances helping to "hide" the cut[44]. Further research should explore if this difference in processing might be linked to the lack of hemispheric lateralization of the mu-rhythm ERD observed in our study.

Not supported was our hypothesis that continuity edits, though judged by participating spectators, as producing images more natural in character than reverse shot, do enhance motor cortex activations during action observation, as was observed for camera movements producing experiences close to human vision in real life (see Chapter III.1 of this thesis). As explicated before, we do not take this as evidence against our hypothesis in general. Rather we suggest that this might indicate that continuity edits play less of an activating than sustaining role regarding action perception links that are possibly involved in enabling spectators' bodily involvement in the movie. We consider this a very interesting finding, characterizing the different narrative devices of film that we looked at so far.

I will come back to these interpretations in the following summary and conclusive discussion of the contents of this thesis.

---

[44] Another rule of continuity editing for example proposes to always let reverse shots, if needed, coincide with moments in a shot that include a lot of fast and large movements possibly eliciting saccades or even blinks in the perceiver, which are able to hide the cut. See for example Smith (2005) or Murch (2001).

IV. Summary and Conclusion

This thesis is an illustration of how 4EA approaches to cognition can be used by empirical cognitive neuroscience to develop new experimental hypotheses and setups delivering new insights regarding key questions about the human mind and in particular about human film perception. The main objective of the research presented here was to illuminate the cognitive processes underlying film perception, that is, the special relation between edited live action moving images and the spectator. As underlined in Chapter I, films and movies are ever more present in today's life. Indeed, most recent surveys reported that for about 1/5 of our lifetime we encounter the world exclusively through movies. Nevertheless, our knowledge about the cognitive processes supporting this encounter is still rather limited. This is a curious fact, given the almost bipolar nature of film. On the one hand, objectively seen, the world as represented in film, due to the use of camera and montage, looks very different from the way we perceive it in real life interactions. On the other hand, film is also widely known for its strong illusive power, sometimes even described as eliciting a bodily surrender, drawing the spectator in the fictional world without any trace of constant perceptual violations. This makes film perception one of the mysteries of human cognition, and suggests that filmmakers could be conceived as implicit specialists of perception, thus making their products – movies – worth being studied also by Cognitive Neuroscience.

Remarkably, hypotheses and designs guiding empirical investigations crucially depend on the theoretical framework from which they originated. Traditional cognitivist views on human cognition as outlined in Chapter II.1 conceptualize cognition as relying on symbolic transformations designed to combine and fill in the holes of passively registered data from the environment, to create detailed inner representations of the world and of our possibilities of acting within it. Though so far this framework has led to enormous progress in research exploring the human mind and brain, it often neglects certain aspects – that is the human body and its interactions with the world and other animate beings – that should be considered as equally important for cognition as the architecture of the human brain. Following this insight, embedded, embodied, enactive, extended and affective approaches to the mind describe cognition as a dynamical interaction between an animate body and an enlived world, relying on experience and skillful behavior rather than on innate and learned knowledge describable in terms of rules and conventions. In the last decades such approaches have been considered to be of use also when trying to understand human perception of arts and media. Chapter II.2 outlined this development with a special focus on new hypotheses generated within theories on the perception of visual images. Interestingly, it was a neuroscientific finding that recently boosted positions emphasizing the role of the body in art perception: the discovery of mirror neurons. Precisely, Freedberg and Gallese (2007) proposed that at a basic level beholders experience their relationship with visual art as being bodily affected both by the represented content and by its representation. They proposed that certain aesthetic experiences at a basic level rely on inner embodied simulation of what the image represents or of the hand gestures that created it (as, for example, the trace of the artist's creative act captured in expressive brushstrokes). They furthermore suggested that such resonance-phenomena might crucially rely on direct action-perception links, like those provided by the mirror mechanism. These links can be automatically activated not only during real world observation but also during the perception of images portraying the world. Precisely, such activation consists of the articulated coming into play of motor or sensory regions of the brain, despite the lack of actual movements or sensations – an event possibly serving direct understanding of the represented via the support of simulation processes. Some scholars have also argued that such activations might be responsible for our feeling of the presence of things (even in depictions) while at the same time correlating with bodily engagements (explorative movements) not only exploring the depicted, but also the depiction itself. This, in turn, might lead to the development of new perceptual skills via the interaction with artworks (see for example Fingerhut, 2014). Remarkably, research carried out during the last years has indeed started to empirically support Gallese & Freedberg's claims by showing that also static images and traces of actions have the capacity to activate and distinctly modulate the cortical motor system of observers' brains.

Film experiences, as most of us would admit, even more than other visual media experiences, seem to involve strong bodily affection. This might be the reason why cognitive film science of the last decades, more willingly than other subjects within arts and media science, have quickly picked up on the idea that film perception might crucially rely on embodied mechanisms of human cognition in the real world. Remarkably the discussion within this field did not focus only on the role of action perception links. Rather, ecological film theory has argued that film derives its illusionary power from the complete recruitment of embodied perception, characterized by crucial experiences with perceptual stimuli that, judged from an objective stance, are anything but of a continuous and transparent nature. In line with 4EA approaches as portrayed in Chapter II.1, scholars supporting this theory assume that these fundamental experiences derived from a continuous bodily exploration of the world can be used not to substitute information lacking in the movies or to transform "wrong" representation of film into something known, but rather to equally enter in interaction with a fictional world. Such an interaction, so it is further suggested, is substantiated by moving images closely simulating vision as known from us from real world perception. On the base of this grounding then, even violations of visual habits, such as weird perspectives or spatial and temporal jumps, do not stop the interactive relationship completely, but are only registered in their deviating nature. Interestingly, traditional narrative film devices such as camera movement and montage, according to the descriptions found (see Chapter II.3, III.1.1 and III.2.1), can play a fundamental role in these processes. Via the creation of moving images that crucially substantiate experiences from real world perception by closely resembling vision known from this normal interaction with the world, some camera movements and montages are suggested to lay the very groundings of spectators' immersive experience. On the other hand, the same narrative devices can also be used to create experiences different from what is known to us in real life, thus generating surprise, shock or estrangement, while at the same time maybe step by step extending our perceptual habits and skills.

On the basis of these collected assumptions, several hypotheses guided the first neuroscientific experiments investigating the neural underpinnings of film perception, focusing on the effect of camera movement and montage on brain activity.

These hypotheses comprise (see Chapter II.4):

1) Detecting Violations of Visual Habits: New Stimulus or Violation of Action

   Moving images, on the basis of the recruitment of embodied vision as a perceptual activity exploring a constant and continuous world known from experience, are evaluated as the experiences of an animate being exploring its enlived environment rather than a row of snapshots of different quality in need of interpretation. Violations of visual habits caused by the special narrative devices of film, such as camera movements and montage, should thus be detected as violations of the observed actions rather than as plain visual ruptures or changes.

2) Processing Violations of Visual Habits: Repair or Reflexion

   Nevertheless, the degree of similarity of the moving images with visual impressions as known to us from our normal interaction with the real world should significantly modulate the activations found. Specifically, different montage techniques, described to more or less violate constancy expectations derived from real world interactions, should be associable to differences in late stages of cognitive processing, possibly illuminating the diverging descriptions of these kind of montages in film theory and the spectators' experience.

3) Familiar Vision and Action Perception Links

   Lastly, on the basis of the suggestions on a possible role of action-perception links such as the mirror mechanism in visual art and media perception, it can be hypothesized that moving images more strongly resembling real world visual experiences elicit a stronger activation of the mirror mechanism, possibly underlying spectators' immersion in the movie by initiating bodily simulations or explorations with a (fictional) world – a world that exactly on the basis of these activations is felt as being in our grasp.

Indeed, the experiments designed on the basis of these hypotheses provided crucial new information about neural activities during movie watching that were interpreted within the given framework.

Precisely, the first study, focusing on hypothesis 3 by investigating the effect of different kinds of camera movements and uses on the activation of the motor cortex, revealed that the time-course of mu-rhythm ERD/ERS/rebound during the observation of actions represented in short movie clips correlated with the camera movements employed to create the movie. More precisely, the modulation observed seem to be dependent on the resemblance between the effect of camera movements and ordinary human vision. Visual experience during the movie judged as being closest to visual experience in real life, elicited the strongest ERD, indicating a higher activation of the mirror mechanism. Furthermore, the more artificial ways of moving or using the camera, such as dolly cam and zoom, correlated with an earlier rebound of the mu-rhythm, indicating an earlier shut down of the mirror mechanism. We therefore proposed that the familiarity effect of the visual traces of camera movements in filmed stimuli modulates the activation of the mirror mechanism, particularly during observation of goal-related hand actions.

Unfortunately experimental setups and results did only partly supported the hypothesis that the activation of the mirror mechanism has a significant effect on spectator's involvement with the scene, and therefore his immersion in the fictional world. Furthermore, findings raised the crucial question if the additional activations found are actually due to an enhanced activation of hand areas in the motor cortex due to the familiar vision, or to an additional activation of areas concerning muscles needed for walking, such as the feet areas. This latter possibility would indicate that the mirror mechanism can not only be activated by the perception of goal related actions of other agents, or the traces of those actions performed in the past, but also by the perception of the perceptual movements of another agent, by this crucially marking perception as a goal related activity. Clearly, future research needs to be employed to further illuminate these issues, particularly relevant to the understanding of why we experience film the intense way we do.

Despite these open questions, our study gave first support for Hypothesis 3 on our list, predicting action perception links to have a fundamental role in film perception by supporting spectators' immersion in a fictional world, by simulating visual conditions known from real life. Specifically, we found that the mirror mechanism is modulated by the spectator's familiarity with the vision simulated by the effect of camera movements and uses.

Next, the second study investigating the effect of different kind of montages on brain activity (ERP and ERD analysis, including tests of Hypotheses 1, 2 and 3) showed that, in contrast to plain onsets of a new visual stimulus, within-scene cuts correlated with modulations of early components of brain electrical activity, which previously have been associated with structural violation, ambiguous character and hard comprehensibility. Remarkably, no difference between continuity edits and reverse shots was found for these components. These results in fact support the notion that within-scene cuts in general are detected in their violation of certain expectations regarding spatial, temporal and causal constancy derived from real life (Hypothesis 3). More concretely, the results suggest that the visual changes evoked by a cut are first of all categorized as structural/syntactic violations of the action perceived, independently of the degree of violation.

Next, our results showed a late anterior left positivity that was more pronounced for continuity edits than for reverse shots. In previous research, such a component has been found to occur after an early left anterior negativity, which was associated with the detection of a structural violation in visually presented action sequences, if the violation is task-relevant. The late anterior left positivity was therefore interpreted as possibly representing updating processes that serve the reanalysis of the scene (see Maffongelli et al., submitted). The absence of this component for reverse shots in our study supports a hypothesis first proposed by Zacks and Magliano who suggested that a suppression of postperceptual processing in visual regions might cause the perceptual salience of non-continuity edits in regarding spectators' experiences. Furthermore, we found a late central right positivity (P4-6), which was more pronounced for reverse shot montages, possibly indicating an enhancement of this salience-effect due to the access of reflexive consciousness leading to the awareness of the perceptual process rather than of its content (all Hypothesis 2).

Regarding ERD data, for continuity edits we found stronger ERD of the mu-rhythm in the left hemisphere, likely due to a selective response to the (right) hand moving in the observed video. On the contrary, in reverse shot condition left and right hemispheres showed equal ERD. We suggested that this might be due to a short orientation-deficit, following the strong spatial violation caused by the reverse shot, who, regarding the spectator's experience abruptly turns left to right and vice versa. In consequence, for reverse shots no clear contra-lateral lateralization of the ERD might occur, but rather both hemispheres should show equally strong ERD – as indeed was found in the present experiment. Future research has to qualify if this activation contributes to the differences between continuity edits and reverse shots in later stages of cognitive processing as reported above.

Taken together this study provided strong evidence in support of Hypotheses 1 and 2 of our list, predicting that violations of visual habits would be cognitively detected as violations of actions, and that later postperceptual processing underlies spectators' subjective experiences that differentiate between moving images similar to vision from the real world and moving images clearly deviating from what is known by bodily experience.

As maybe one of our most interesting findings in this study we furthermore discovered that, contrary to our initial hypotheses (Hypothesis 3), ERD responses did not differ in their general strength between the two montage conditions. This might indicate that camera movements and montage might have different influences on cognition and, most importantly, on film perception. While ecological camera movements, such as those elicited by use of the steadicam producing moving images very close to perception in the real world, might have the power to actually enhance or even initiate spectators' immersion in the movie, continuity edits, as the declared equivalent of an ecological edit, might have only a sustaining role in this process. Again, future research has to further explore this hypothesis.

Lastly, it should be stressed that the findings of our studies are obviously relevant also outside the specific field of film theory.

One of the fields that could profit most from our findings might indeed be empirical research about cognition itself. As a matter of fact, most experiments exploring cognition and perception crucially rely on films and movies, which are assumed to represent a way to present to participants the real world under highly controlled conditions. Our results indicate that the representations themselves might have a crucial effect on the results as differences with visual perception in the real world, elicited by the representation chosen, proved able to not only elicit quantitative differences in the measurements (such as enhanced or attenuated activations), but even elicit different postperceptual processing of the stimuli. This argues for caution in interpretating results possibly affected by such confounding factors, as well as for the application of more ecological studies in cognitive neuroscience.

Furthermore, given the fact that our findings essentially support an ecological view on film perception (by supporting Hypotheses 1, 2 and 3), they should stimulate more research following this line. This includes:

a) Other Violations and Familiarities

Needed are studies extending the investigations presented here to other uses of film's special narrative devices. Obvious examples of such investigations lie in further explorations regarding the rules given by continuity editing. Such comparisons include continuity edits versus jump cuts, match action edits versus non match action edits etc. It should also be illuminating to investigate the influence of measures that continuity editing describes as helping to overcome certain violations, such as strong movement in the scene hiding cuts via eliciting saccades or special effects that attenuate the abruptness of a cut such as fade ins or fade outs. It could be equally interesting to investigate the possibilities of digital film such as slow and fast motion, matting, virtual backgrounds and so forth, all representing experiences not known to us from real life.

b) Moving Image and Sound

A very interesting topic consists in the interaction of moving image and sound. Questions to be asked here for example comprise the possibility of a smoothing effect of continuous sounds overlaying discontinuity edits. It is thinkable that such combinations have a crucial effect on brain activity turning discontinuity into perceived continuity also regarding brain activity.

c) New Perceptual Habits

One should also seriously consider the claim, repeated throughout the whole thesis, that film perception, thought of an interaction between a spectator and a fictional world, should have the power to evoke the development of new perceptual habits. As also Sobchack stresses, film finally should be conceptualized as part of our extended mind incorporated into one's being. Just as a far-sighted person is not able to read a book without the help of glasses, films enable us a vision more and more deviating from what we are capable of doing as a result of our learning in real life. It seems highly interesting to try to trace the transient changes in brain responses, that should be involved in these extensions, possibly rendering effects initially perceived as strong violations into slight deviations and lastly even into known perceptual circumstances. Something like this might be realizable by comparing the film perception of experts (that is filmmakers, cameramen or editors) with that of amateurs. It might also be worth to try simulating the development of expertise by repeatedly presenting to amateurs new kind of moving images, comparing neural activations at the beginning and at the end of this kind of training, while looking for the enduing changes. Lastly it should be highly interesting to trace perceptual habits most likely developed in the interaction with movies in their influence on life off the screen, such as in imagination, memory, dreams or even visual behavior of images and the world in general.

In a nutshell, there is a lot more to do on our way to understanding what it means to watch a movie.

This thesis suggested and elaborated a new theoretical background for human film perception on the basis of which first empirical research was designed and conducted. Results of the reported experiments delivered strong evidence for the basic hypotheses made. This showed that the embedding of neuroscientific research in the framework of 4EA approaches to cognition has the potential to establish an important and new framework for future interdisciplinary research on our relation to one of the most important media of today: Film.

IV. References

Altschuler, E.L., Vankov, A., Hubbard, E.M., Roberts, E., Ramachandran, V.S., & Pineda, J.A. (2000). Mu wave blocking by observation of movement and its possible use as a tool to study theory of other minds. Poster session presented at the 30th Annual Meeting of the Society for Neuroscience. New Orleans, LA: Society for Neuroscience Conference Abstracts.

Altschuler, E.L., Vankov, A., Wang, V., Ramachandran, V.S., & Pineda, J.A. (1997). Person see, person do: human cortical electrophysiological correlates of monkey see monkey do cell. Poster session presented at the 27th Annual Meeting of the Society for Neuroscience. New Orleans, LA: Society for Neuroscience Conference Abstracts.

Anderson, J.D. (1998). The Reality of Illusion: An Ecological Approach to Cognitive Film Theory. Carbondale: Southern Illinois University Press.

Avanzini, P., Fabbri-Destro, M., Dalla Volta, R., Daprati, E., Rizzolatti, G. & Cantalupo, G. (2012). The Dynamics of Sensorimotor Cortical Oscillations during the Observation of Hand Movements: An EEG Study. *PLoS ONE 7(5)*: e37534. doi:10.1371/journal.pone.0037534.

Babiloni, C., Del Percio, C., Rossini, P.M., Marzano, N., Iacoboni, M., Infarinato, F., Lizio, R., Piazza, M., Pirritano, M., Berlutti, G., Cibelli, G., & Eusebi, F. (2009). Judgment of actions in experts: a high-resolution EEG study in elite athletes. *Neuroimage*, *45*, 512–521.

Bach, P., Peatfiel, N.A., & Tipper, S.P. (2007). Focusing on body sites. The role of spatial attention in action perception. *Experimental Brain Research*, *178*, 509–517.

Baron-Cohen, S. (2001). Theory of Mind in Normal Development and Autism. *Prisme*, *34*, 174–183.

Baron-Cohen, S., Leslie, A.M., & Frith, U. (1985). Does the Autistic Child Have a 'Theory of Mind'? *Cognition*, *21*, 37–46.

Barrett, S.E., Rugg, M.D., & Perrett, D.I. (1988). Event related potentials and the semantic matching of familiar and unfamiliar faces. *Neuropsychologia*, *26*, 105–117.

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral Brain Sciences*, *22*, 577–660.

Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645.

Barthes, R. [1980] (1981).Camera Lucida. New York: Hill and Wang.

Benjamin, W. [1935]. Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit. In: Tiede-mann, R., & Schweppenhäuser, H. (1980). Walter Benjamin: Gesammelte Schriften Band I, Werk-ausgabe Band 2. Frankfurt am Main: Suhrkamp Verlag. pp. 431–469.

Berenson, B. (1896). The Florentine Painters of the Renaissance, with an Index to their Works. London.

Bernier, R., Dawson, G., Webb, S., & Murias, M. (2007). EEG mu rhythm and imitation impairments in individuals with autism spectrum disorder. *Brain & Cognition*, *64*, 228–237.

Bértolo, H. (2005). Visual imagery without visual perception? *Psicológica*, *26*, 173–188.

Bettetini, G. (1973). The Language and Technique of the Film. The Hague: Mouton.

Bettman, G. (2013). Directing the Camera. How Professional Directors Use a Moving Camera to Energize Their Films. Studio City (CA): Michael Wiese Productions.

Bhalla, M., Proffitt, D.R. (1999). Visual-motor recalibration in geographical slant perception. *Journal of Experimental Psychology. Human Perception and Performance, 25*, 1076–1096.

Blackmore, S.J., Brelstaff, G., Nelson, K., & Troscianko, T. (1995). Is the richness of our visual world an illusion? Transsaccadic memory for complex scenes. *Perception*, *24*, 1075–1081.

Block, N. (2001). Paradox and cross purposes in recent work on consciousness. *Cognition*, *79*, 197–219.

Bordwell, D., & Thompson, K. (2004). Film Art: An Introduction (7th ed.). N.Y.: McGraw-Hill.

Boronat, C.B., Buxbaum, L.J., Coslett, H.B., Tang, K., Saffran, E.M., Kimberg, D.Y., & Detre, J.A. (2005). Distinctions between manipulation and function knowledge of objects: evidence from functional magnetic resonance imaging. *Cognitive Brain Research*, *23*, 361–373.

Bridgeman, B., van der Heijden, A.H.C., & Velichkovsky, B.M. (1994). A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences*, *17*, 247–292.

Brown, G. (2003). The Moving Camera Part I. *Zerb*, *58*. Also available on Garrett Brown's website: www.garrettcam.com

Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G. & Freund, H.-J. (2001). Action Observation Activates Premotor and Parietal Areas in a Somatotopic Manner: a fMRI Study. *European Journal of Neuroscience*, *13*, 400–404.

Carr, L., Iacoboni, M., Dubeau, M.C., Mazziotta, J.C., & Lenzi, G.L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Science USA*, *100*, 5497–5502.

Carroll, J. M., & Bever, T. G. (1976). Segmentation in Cinema Perception. *Science*, *191*, 1053–1055.

Carson, G. (dir.) (2002). Vilmos Zsigmond Flashes 'The Long Goodbye.' MGM Home Entertainment.

Casetti, F. (2015). The Lumière Galaxy. Seven Key Words for the Cinema to Come. New York: Columbia University Press.

Chao, L.L., & Martin, A. (2000). Representation of manipulable manmade objects in the dorsal stream. *Neuroimage*, *12*, 478–484.

Chong, T.-J., Williams, M.A., Cunnington, R., & Mattingley, J. B. (2008). Selective attention modulates inferior frontal gyrus activity during action observation. *NeuroImage*, *40*, 298–307.

Clark, A. (1997). Being There. Putting Brain, Body, and World Together Again. Cambridge/London: MIT Press.

Clark, A., & Chalmers, D. (1998). The Extended Mind. *Analysis*, *58*, 10–23.

Cochin, S., Barthelemy, C., Lejeune, B., Roux, S., & Martineau, J. (1998). Perception of motion and qEEG activity in human adults. *Electroencephalography in Clinical Neurophysiology*, *107*, 287–295.

Cochin, S., Barthelemy, C., Roux, S., & Martineau, J. (1999). Short Communication – Observation and execution of movement: similarities demonstrated by quantified electroencephalography. *European Journal of Neuroscience*, *11*, 1839–1842.

Collingwood, R. G. (1938). The Principles of Art. Oxford: Clarendon Press.

Colombetti, G. (2014). The Feeling Body: Affective Science Meets the Enactive Mind. Cambridge, MA: MIT Press.

Comuntzis, G.M. (1987). Children's Comprehension of Changing Viewpoints in Visual Presentations. Visual Communication Conference.

Comuntzis-Page, G.M. (1991). Perspective-taking theory: shifting views from Sesame Street. Visual Communication Conference.

Curtis, R., & Koch, G. (eds.) (2009). Einfühlung. Zur Geschichte und Gegenwart eines ästhetischen Konzepts. München: Fink Verlag.

Cutting, J.E. (2004). Perceiving Scene in Film and in the World. In: Anderson, J.D., & Anderson B. F. (eds.) Moving image theory: ecological considerations. Carbondale: Southern Illinois University Press.

Cutting, J.E., DeLong, J.E., & Nothelfer, C.E. (2010). Attention and the Evolution of Hollywood Film. *Psychological Science*, *21*, 440–447.

Damasio, A.R. (1994). Descartes' Error: Emotion, Reason and the Human Brain. New York: Penguin Group.

Damasio, A.R. (2008). Cinéma, esprit et émotion: la perspective du cerveau. *Trafic*, *67*, 94–101.

Debruille, J.B., Brodeur, M.B., & Franco Porras, C. (2012). N300 and Social Affordances: A Study with a Real Person and a Dummy as Stimuli. *PLoS ONE, 7(10)*: e47922. doi: 0.1371/journal.pone.0047922

Debruille, J.B., Pineda, J., & Renault, B. (1996). N400-like potentials elicited by faces and knowledge inhibition. Brain Research. *Cognitive Brain Research, 4,* 133–144.

Decety, J., Grèzes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., & Fazio, F. (1997). Brain activity during observation of actions. Influence of action content and subject's strategy. *Brain*, *120*, 1763–1777.

De Jaegher, H., & Di Paolo, E. (2007). Participatory SenseMaking: An Enactive Approach to Social Cognition. *Phenomenology and the Cognitive Sciences*, *6*, 485–507.

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Does social interaction constitute social cognition? *Trends in Cognitive Sciences*, *14*, 441–447.

Derambure, P., Defebvre, L., Dujardin, K., Bourriez, J.L., Jacquesson, J.M., Destee, A., & Guieu, J.D. (1993). Abnormal cortical activation during planning of voluntary movement in patient with epilepsy with focal motor seizures: Event-related desynchronization study of electroencephalographic mu rhythm. *Epilepsia*, *38*, 655–622.

Di Dio C., & Gallese V. (2009). Neuroesthetics: a review. *Current Opinion in Neurobiology*, *19*, 682–687.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research*, *91*, 176–180.

Donald, M. (1991). Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition. Cambridge, MA: Harvard University Press.

Donchin, E., & Coles, M.G.H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 357–374.

Dudai, Y. (2012). The cinema-cognition dialogue: a match made in brain. *Frontiers in Human Neuroscience, 6*, art. 248.

D'Ydewalle, G., Desmet, G., & Van Rensbergen, J. (1998). Film Perception: The processing of film cuts. In: G. Underwood (ed.). Eye guidance in reading and scene perception. Oxford: Elsevier. pp. 357–367.

D'Ydewalle, G., & Vanderbeeken, M. (1990). Perceptual and cognitive processing of editing rules in film. In: Groner, R., D'Ydewalle, G., & Parnham, R. (eds.), From eye to mind: Information acquisition in perception, search, and reading. Amsterdam: Elsevier (North Holland). pp. 129–139.

Eco, U. (1987). Semiotik – Eine Einführung. München: Wilhelm Fink Verlag.

Eco, U. [1995] (2000) Kant and the Platybus: Essays on Language and Cognition. New York: Mariner Books.

Eisenstein, S. (2006). Jenseits der Einstellung. Schriften zur Filmtheorie. Frankfurt am Main: Suhrkamp Verlag.

Fadiga L., Fogassi L., Pavesi G., & Rizzolatti G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, *73*, 2608–2611.

Fingerhut, J. (2012). Das Bild, dein Freund. Der fühlende und der sehende Körper in der enaktiven Bildwahrnehmung. In: Feist, U., & Rath, M. (eds.) (2012). Et in imagine ego. Facetten von Bildakt und Verkörperung. Berlin: Akademie Verlag. pp. 177–198.

Fingerhut, J. (2014). Extended Imagery, Extended Access, or Something Else? Pictures and the Extended Mind Hypothesis. In: Marienberg, S., & Trabant, J. (eds.) Bildakt at the Warburg Institute. Berlin/New York: de Gruyter.

Forgas, J.P. (2000): Handbook of Affect and Social Cognition. Mahwah: Taylor & Francis.

Freedberg, D., & Gallese, V. (2007). Motion, Emotion and Empathy in Aesthetic Experience. *Trends in Cognitive Science*, *11*, 197–203.

Friederici, A.D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*, 78–84.

Friederici, A.D., Pfeifer, E., & Hahne A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*, 183–192.

Frith, U., & Robson, J. E. (1975). Perceiving the language of films. *Perception*, *4*, 97–103.

Gallagher, S. (2001). The Practice of Mind. Theory, Simulation or Primary Interaction? *Journal of Consciousness Studies*, *8*, 83–108.

Gallagher, S. (2005). How the body shapes the mind. Oxford: Oxford University Press.

Gallagher, S. (2007). Simulation Trouble, *Social Neuroscience*, *2/3*, 353–365.

Gallagher, S. (2008a). Direct perception in the intersubjective context. *Consciousness and Cognition*, *17*, 535–543.

Gallagher, S. (2011a). Strong interaction and self-agency. *Humana Mente Journal of Philosophical Studies*, *15*, 55–76.

Gallagher, S. (2011b). Empathy, simulation, and narrative. *Science in Context, 25*, 355–381.

Gallagher, S. (2011c). Aesthetics and Kinaesthetics. In: Bredekamp, H., & Krois, J. M. (eds.) (2011), Sehen und Handeln. Berlin: Akademie Verlag GmbH, John Wiley and Sons, Inc. pp. 99–113.

Gallagher, S. (2012). In defense of phenomenological approaches to social cognition: interacting with the critics. *Review Philosophical Psychology*, *3*, 187–212.

Gallagher, S., Hutto, D. (2008). Understanding others through primary interaction and narrative practice. In: Zlatev, J., Racine, T., Sinha, C., & Itkonen, E. (eds.) (2008). The shared mind: perspectives on intersubjectivity. Amsterdam: John Benjamins. pp. 17–38.

Gallese V. (2000). The inner sense of action: agency and motor representations. *Journal of Consciousness Studies: 7; 23-40.*

Gallese, V. (2001). The "Shared Manifold" Hypothesis: from mirror neurons to empathy. *Journal of Consciousness Studies*, *8*, 33–50.

Gallese, V. (2003). The Manifold Nature of Interpersonal Relations: The Quest for a Common Mechanism. *Philosophical Transactions of the Royal Society*, *358*, 517–528.

Gallese, V. (2005). Embodied simulation: From neuron to phenomenal experience*. Phenomenology and the Cognitive Sciences*, *4*, 23–48.

Gallese V. (2007). The "conscious" dorsal stream: Embodiend simulation and its role in space and action conscious awareness. *PSYCHE*: http://psyche.cs.monash.edu.au/

Gallese, V. (2012). Aby Warburg and the dialogue among aesthetics, biology and physiology. *Phenomenology*, *2*, 48–62.

Gallese, V. & Di Dio, C. (2012). Neuroesthetics: The Body in Esthetic Experience. In: Encyclopedia of Human Behavior, Vol. 2, 687–693.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593–609.

Gallese, V., & Guerra, M., (2012). Embodying Movies: Embodied Simulation and Film Studies. *Cinema: Journal of Philosophy and the Moving Image*, *3*, 183–210.

Gallese, V., & Sinigaglia, C. (2010). Understanding action with the motor system. *Behavioral Brain Sciences*, *37*, 199–200.

Gallese V., & Sinigaglia C. (2011). What is so special with Embodied Simulation. *Trends in Cognitive Sciences*, *15*, 512–519.

Gazzola, V., Rizzolatti, G., Wicker, B., & Keysers, C. (2007). The anthropomorphic brain: The mirror neuron system responds to human and robotic actions. *NeuroImage*, *35*, 1674–1684.

Geiger, S., & Reeves, B. (1993). The effects of scene changes and semantic relatedness on attention to television. *Communication Research*, *20*, 155–175.

Gerlach, C., Law, I., Gade, A., & Paulson, O.B. (2002). The role of action knowledge in the comprehension of artifacts – a PET study. *Neuroimage*, *15*, 143–152.

Gianetti, L.D. (1976). Understanding Movies. Englewood Cliffs, N.J.: Prentice-Hall.

Gibbs, R.W. Jr. (2006). Embodiment and Cognitive Science. New York: Cambridge Univ. Press.

Gibson, J.J. (1979). The Ecological Approach to Visual Perception. Boston: Houghton Mifflin.

Glassman, A., McCarthy, T., & Samuels, S. (dir.) (1992). Visions of Light. AFI/NHK. A Documentary on Cinematography.

Goldman, A. (2006). Simulating minds: the philosophy, psychology, and neuroscience of mindreading. Oxford: Oxford University Press.

Goodman, E. [1945]. Movement in the movies. In: Jacobs, L. (ed.) (1970). The Movies as a Medium. New York: Farrar, Straus & Giroux. pp. 76–84.

Goodman, N. (1995). Sprachen der Kunst. Frankfurt a. M.: Suhrkamp Verlag.

Gombrich, E.H. (1960). Art and Illusion. Princeton: Princeton University Press.

Gorki, M. [1996] Vos nerfs se tendent. In: Banda, D., & Moure, J. (eds.) (2008). Le Cinéma: naissance d'un art. 1895-1920 Paris: Flammarion. 47–57.

Grafton, S.T., Fadiga, L., Arbib, M.A., & Rizzolatti, G. (1997). Premotor cortex activation during observation and naming of familiar tools. *Neuroimage, 6*, 231–236.

Greenberg, C. (1961) Art and Culture, Boston: Beacon Press.

Grèzes, J., Armony, J.L, Rowe, J., & Passingham, R.E. (2003). Activations related to "mirror" and "canonical" neurons in the human brain: an fMRI study. *NeuroImage*, *18*, 928-937.

Grieveson, L., & Wasson, H. (2008). Inventing Film Studies. New York: Duke University Press Books.

Grill-Spector, K., & Malach, R. (2001). fMR-adaption: A tool for studying the functional properties of human cortical neurons. *Acta Pyschologica*, *107*, 293–321.

Grodal, T., (1999). Moving Pictures. A New Theory of Film Genre, Feelings and Cognition. Oxford: Oxford University Press.

Grodal, T. (2009). Embodied Visions – Evolution, Emotion, Culture and Film. Oxford: Oxford University Press.

Grodzinsky, Y., & Santi, A. (2008). The battle for Broca's region. *Trends in Cognitive Sciences, 12*, 474–480.

Grueny, C. (unpublished manuscript). Mimesis at a distance – Music as resonance. https//www.academia.edu/452760/Mimesis at a Distance Music as resonance

Grueny, C. (2014). Kunst des Übergangs. Philosophische Konstellationen zur Musik. Weilerswist: Velbrück Wissenschaft.

Hahne, A., & Friederici, A.D. (1999). Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, *11*, 194–205.

Hanich, J., (2013). Wut im Kino. Anmerkungen zu einem vernachlässigten Alltagsphänomen. *Augenblick. Konstanzer Hefte zur Medienwissenschaft*, *56/57*, 72–93.

Hari, R., Forss, N., Avikainen, S., Kirveskari, S., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proceedings of the National Academy of Sciences USA*, *95*. 15061–15065.

Hasson, U., Furman, O., Clark, D., Dudai, Y., & Davachi L. (2008a). Enhanced Intersubject Correlations During Movie Viewing Correlate with Successful Episodic Encoding. *Neuron*, *57*, 452–462.

Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., & Heeger, D.J. (2008b). Neurocinematics: The Neuroscience of Film. *Projections*, *2*, 1–26.

Haugeland J. (1991). Representational genera. In: Ramsey, W., Stitch, S.P., Rumelhart, D.E. (eds.) (1991). Philosophy and Connectionist Theory. Hillsdale, NJ: Erlbaum. pp. 61–89.

Heimann, K. (unpublished Masterthesis). Von den Bedingungen der Moeglichkeit von Aehnlichkeit. – Why Philosophy needs Embodied Cognitive Sciences. University of Witten/Herdecke.

Heimann, K., & Uithol, S. (in print). Participation over observation. The roots of action understanding in attunement. *Journal of Consciousness Studies*.

Heimann, K., Umiltà, M. A., & Gallese, V. (2013). How the motor-cortex distinguishes among letters, unknown symbols and scribbles. A high density EEG study. *Neuropsychologia*, *51*, 2833–2840.

Heimann, K., Umiltà, M.A., Guerra, M., & Gallese, V. (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*, 2087–2101.

Hickok, G. (2014). The Myth of Mirror Neurons: The Real Neuroscience of Communication and Cognition. New York/London: W. W. Norton & Company, Inc.

Hochberg, J., & Brooks, V. (1978). The Perception of Motion Pictures. In: Carterette, E.C., & Friedmann, M. P. (eds.). Handbook of Perception, Vol. X . New York: Academic Press.

Horst, S., "The Computational Theory of Mind". In: Zalta E.N. (ed.). The Stanford Encyclopedia of Philosophy (Spring 2011 Edition):
URL = <http://plato.stanford.edu/archives/spr2011/entries/computational-mind/>.

Hulme, O.J., & Zeki, S. (2006). The sightless view: neural correlates of occluded objects. *Cerebral Cortex,* doi: 10.1093/cercor/bhl031

Hutto, D. (2013). Action understanding: how low can you go? *Consciousness and Cognition*, 22, 1142–1151.

Hutto, D., & Myin, E. (2012). Radicalizing Enactivism. Cambridge/London: MIT Press.

Iacoboni, M. (2009). Imitation, Empathy and Mirror Neurons. *Annual Review of Psychology*, *60*, 653–670.

Iacoboni, M., Woods, R.P., Brass, M., Bekkering, H., Mazziotta, J.C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, *286*, 2526–2528.

Iwase, M., Ouchi, Y., Okada, H., Yokoyama, C., Nobezawa, S., Yoshikawa, E., Tsukada, H., Takeda, M., Yamashita, K., Takeda, M., Yamaguti, K., Kuratsune, H., Shimuzu, A., & Watanabe, Y. (2002). Neural Substrates of Human Facial Expression of Pleasant Emotions induced by Comic Films: A PET Study. *NeuroImage*, *17*, 758–768.

Jacob, P., & Jeannerod, M. (2003). Ways of Seeing. The Scope and Limits of Visual Cognition. Oxford: Oxford University Press.

Johnson-Frey, S., Maloof, F., Newman-Norlund, R., Farrer, C., Inati, S., & Grafton, S. (2003). Actions or hand-object interactions? Human inferior frontal cortex and action observation. *Neuron*, *39*, 1053–1058.

Kanakogi, Y., Okumura, Y., Inoue, Y., Kitazaki, M., & Itakura, S. (2013). Rudimentary Sympathy in Preverbal Infants: Preference for Others in Distress. *PLoS ONE*, *8(6)*: e65292. doi:10.1371/journal.pone.0065292

Keysers, C., & Gazzola, V. (2006). Towards a unifying neural theory of social cognition. *Progress in Brain Research*, *156*, 379–401.

Keysers, C., Kohler, E., Umiltà, M.A., Fogassi, L., Rizzolatti, G., & Gallese, V. (2003). Audiovisual mirror neurones and action recognition. *Experimental Brain Research*, *153*, 628–636.

Keysers, C., Wickers, B., Gazzola, V., Anton, J.-L., Fogassi, L., & Gallese, V. (2004). A touching sight: SII/PV activation during the observation and experience of touch. *Neuron*, *42*, 1–20.

Klee, P. (1920). Schöpferische Konfession. In: Edschmid, K., (1920). Tribüne der Kunst und der Zeit. Eine Schriftensammlung, Band XIII. Berlin: Reiß.

Kochanska, G., Coy, K. C., & Murray, K. T. (2001). The development of self-regulation in the first four years of life. *Child Development*, *72*, 1091–1111.

Koelewijn, T., van Schie, H.T., Bekkering, H., Oostenveld, R., & Jensen, O. (2008). Motor-cortical beta oscillations are modulated by correctness of observed action. *NeuroImage*, *40*, 767–775.

Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, *297*, 846–848.

Koivisto, M., & Revonsuo, A. (2003). An ERP study of change detection, change blindness, and visual awareness. *Psychophysiology*, *40*, 423–429.

Kosslyn, S.M. (1980). Image and Mind. Cambridge, MA: Harvard University Press.

Kracauer, S., (1960). Theory of Film. The Redemption of Physical Reality, New York: Oxford University Press.

Kraft, R. N. (1987). Rules and strategies of visual narratives. *Perceptual and Motor Skills*, *64*, 3-14.

Koelsch, S. (2011). Toward a neural basis of music perception - a review and updated model. *Frontiers in Psychology*, *2*, 110.

Kutas, M., & Federmeier, K. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*, 463–470.

Kutas, M., & Hillyard, S. (1980). Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity. *Science*, *207*, 203–205.

Lakoff, G., & Johnson, M. (1980). Metaphors we live by. Chicago: University of Chicago Press.

Lang, A. (1991). Emotion, formal features, and memory for televised political advertisements. In: Biocca, F. (ed.). Television and political advertising, Vol. I: Psychological processes. Hillsdale, NJ: Erlbaum.

Lang, A., Geiger, S., Strickwerda, M., & Sumner, J. (1993). The effects of related and unrelated cuts on viewers memory for television: A limited capacity theory of television viewing. *Communication Research*, *20*, 4–29.

Langer, S. K. (1953). Feelings and Form. A theory of art developed from Philosophy in a New Key. New York: Routledge & Kegan Paul.

Langer, S. K. (1957). Problems of Art. New York: MacMillan USA.

Langer, S. K. (1984). Philosophie auf neuem Wege. Das Symbol im Denken, im Ritus und in der Kunst, Frankfurt am Main: Fischer Wissenschafts Verlag.

Lau, E.F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Review Neuroscience*, *9*, 920–933.

Leocani, L., Toro, C., Manganotti, P., Zhuang, P., & Hallet, M. (1997). Event-related coherence and event-related desynchronization/synchronization in the 10 Hz and 20 Hz EEG during self-paced movements. *Electroencephalography and Clinical Neurophysiology*, *104*, 199-206.

Lewis, J.W. (2006). Cortical networks related to human use of tools. *Neuroscientist*, *12*, 211–231.

Longcamp, M., Boucard, C., Gilhodes, J.-C., Anton, J.-L., Roth, M., Nazarian, B., & Velay, J.-L. (2008). Learning through Hand- or Typewriting Influences. Visual Recognition of New Graphic Shapes: Behavioral and Functional Imaging Evidence. *Journal of Cognitive Neuroscience*, *20*, 802–815.

Longcamp, M., Tanskanen, T., & Hari, R. (2006). The imprint of action: Motor cortex involvement in visual perception of handwritten letters. *Neuroimage*, *33*, 681–688.

Maffongelli, L., Bartoli, E., Sammler, D., Koelsch, S., Campus, C., Ovivier, E., Fadiga, L., & D'Ausilio, A. (submitted). Distinct brain signatures of content and structure violation during action sequence observation.

Magliano, J.P., & Zacks, J.M. (2011). The Impact of Continuity Editing in Narrative Film on Event Segmentation. *Cognitive Science – A Multidisciplinary Journal*, *35*, 1–29.

Marr, D. (1982). Vision. New York: W.H. Freeman & Sons.

Marraffa, M. (2011). Theory of Mind. Internet Encyclopedia of Philosophy: http://www.iep.utm.edu/theomind/

Martin, A., Wiggs, C.L., Ungerleider, L.G., & Haxby, J.V. (1996). Neural correlates of category-specific knowledge. *Nature*, *379*, 649–652.

Masserman, J.H., Wechkin, S., & Terris, W. (1964). “Altruistic” Behavior in Rhesus Monkeys. *The American Journal of Psychiatry*, *121*, 584–585.

Mast, G. (1977). Film/Cinema/Movie. New York: Harper & Row.

Maunsell, J., & Treue, S., (2006). Feature-based attention invisual cortex. *Trends in Neurosciences*, *29*, 317–322.

May, J., Dean, M. P., & Barnard, P. J. (2003). Using Film Cutting Techniques in Interface Design. *Human-Computer Interactions*, 18, 325–372.

McFarland, D.J., Miner, L.A., Vaughan, T.M., & Wolpaw, J.R. (2000). Mu and beta rhythm topographies during motor imagery and actual movements. *Brain Topography*, *12*, 177–186.

Mercado, G., (2011). The Filmmaker’s Eyes. Learning (and Breaking) the Rule of Cinematic Composition. Boston: Focal Press.

Merleau-Ponty, M. [1961]. Eye and Mind. In: Johnson, G. (Ed.). (1993). The Merleau-Ponty Aesthetics Reader: Philosophy and Painting. Evanston, Illinois: Northwestern University Press. pp. 121–149.

Merleau-Ponty, M. (1964). The Film and the New Psychology. In: Merleau-Ponty, M. Sense and Non-Sense, Evanston: Northwestern University Press. Pp. 48–59.

Messaris, P. (1994). Visual Literacy: Image, Mind, and Reality. Boulder: Westview Press.

Milner, D., & Goodale, M.A. (1995). The Visual Brain in Action. Oxford: Oxford University Press.

Münsterberg, H. (1916). The Photoplay. A psychological study. New York & London: Appleton.

Mukamel, R., Ekstrom, A.D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology*, *20*, 750–756.

Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., & Rizzolatti, G. (1997). Object representation in the ventral premotor cortex (area f5) of the monkey. *Journal of Neurophysiology*, *78*, 2226–2230.

Murch, W. (2001). In the Blink of an Eye – A Perspective on Film Editing. Los Angeles: Silman James Press.

Murzyn, E. (2008). Do we only dream in colour? A comparison of reported dream colour in younger and older adults with different experiences of black and white media. *Consciousness and Cognition*, *17*, 1228–1237.

Muthukumaraswamy, S.D., & Johnson, B.W. (2004). Primary motor cortex activation during action observation revealed by wavelet analysis of the EEG. *Clinical Neurophysiology*, *115*, 1760–1766.

Muthukumaraswamy, S.D., Johnson, B.W., & McNair, N.A. (2004). Mu rhythm modulation during observation of an object-directed grasp. *Cognitive Brain Research*, *19*, 195–201.

Nielson, J.I. (2008). Camera Movement in Narrative Cinema - Towards a Taxonomy of Functions. http://pure.au.dk//portal/files/52113417/Camera_Movement_0910.pdf

Noë, A. (2004). Action in Perception. Cambridge/London: MIT Press.

Noë, A., Pessoa, L., & Thompson, E. (2000). Beyond the grand illusion: What change blindness really teaches us about vision. *Visual Cognition*, *7*, 93–106.

Oberman, L.M., McCleery, J.P., Ramachandran, V.S., & Pineda, J.A. (2007a). EEG evidence for mirror neuron activity during the observation of human and robot actions: Toward an analysis of the human qualities of interactive robots. *Neurocomputing*, *70*, 2194–2203.

Oberman, L.M., Pineda, J.A., Ramachandran, V.S. (2007b). The Human Mirror Neuron System: A Link Between Action Observation and Social Skills. *Social Cognitive and Affective Neuroscience*, *2*, 62-66.

Oldfield, R.C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, *9*, 97–113.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). Field Trip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Computational Intelligence and Neuroscience. doi:10.1155/2011/156869.

O’Regan, J.K. (1992). Solving the “real” mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, *46*, 461–488.

O'Regan, J.K, & Noë, A. (2001). A sensorimotor account of vision and visual conscious-ness, *Behavioral and Brain Sciences*, *25*, 883–975.

O’Regan J.K., Deubel H., Clark J.J., & Rensink R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition*, *7*, 191–212.

O’Regan, J.K., Rensink, R.A., & Clark, J.J. (1996). “Mud splashes” render picture changes invisible. *Investigative Ophthalmology and Visual Science*, *37*, 213.

O’Regan, J.K., Rensink, R.A., & Clark, J.J. (1997). Picture changes during blinks: Not seeing where you look and seeing where you don’t look. *Investigative Ophthalmology and Visual Science*, *38*, 707.

O’Regan, J.K., Rensink, R.A., & Clark, J.J. (1999). Change-blindness as a result of “mudsplashes”. *Nature*, *398*, 34.

Paas-Zeidler, S. (1978). Goya: Radierungen. Stuttgart: Hatje Cantz.

Patel, S.H., & Azzam, P.N. (2005). Characterization of N200 and P300: Selected Studies of the Event Related Potential. *International Journal of Medical Sciences*, *2*, 147–154.

Perani, D., Schnur, T., Tettamanti, M., Gorno-Tempini, M., Cappa, S.F., & Fazio, F. (1999). Word and picture matching: a PET study of semantic category effects. *Neuropsychologia*, *37*, 293–306.

Perry, A., & Bentin, S., (2009). Mirror activity in the human brain while observing hand movements: a comparison between EEG desynchronization in the mu-range and previous fMRI results. *Brain Research*, *1282*, 126–132.

Pessoa, L., Thompson, E., & Noë, A. (1998). Finding out about filling in: A guide to perceptual completion for visual science and the philosophy of perception. *Behavioral and Brain Sciences*, *21*, 723–802.

Pfurtscheller, G., & Aranibar, A. (1979). Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movements. *Electroencephalography and Clinical Neurophysiology*, *46*, 138–146.

Pfurtscheller, G., & Berghold, A. (1989). Patterns of cortical activation during planning of voluntary movement. *Electroencephalography and Clinical Neurophysiology*, *72*, 250–258.

Pfurtscheller, G., & Lopes da Silva, F.H. (1999). Event-related desynchronization and related oscillatory phenomena of the brain. Handbook of electroencephalography and clinical neurophysiology, Vol 6, Amsterdam: Elsevier.

Pfurtscheller, G., Neuper, C., Andrew, C., & Edlinger, G. (1997). Foot and hand area mu rhythms. *International Journal of Psychophysiology*, *26*, 121–135.

Pfurtscheller, G., Neuper, C., Brunner, C., & Lopes da Silva, F. (2005). Beta rebound after different types of motor imagery in man. *Neuroscience Letters*, *378*, 156–159.

Pineda, J. A. (2005). The functional significance of mu rhythms: Translating "seeing" and "hearing" into "doing". *Brain Research Reviews*, *50*, 57–68.

Platinga, C. (1999). Passionate Views. Film, Cognition and Emotion. Baltimore: John Hopkins University Press.

Platinga, C. (2009). Moving Viewers. American Film and the Spectator's Experience. Ewing: University of California Press.

Platon (2007). Sophistes – mit einem Kommentar von Christian Iber, Frankfurt a. M.: Reclam Verlag.

Poggio, T., Torre, V., & Koch, C. (1985). Computational vision and regularization theory. *Nature*, *317*, 314–319.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*, 2128–2148.

Ponseti, J., Bosinsky, H.A., Wolff, S., Peller, M., Jansen, O., Mehdom, H.M., Büchel, C., & Siebner, H.R. (2006). A functional endophenotype for sexual orientation in humans. *Neuroimage*, *33*, 825–833.

Pourtois, G., De Pretto, M., Hauert, C.A., & Vuilleumier, P. (2006). Time Course of Brain Activity During Change Blindness and Change Awareness: Performance is Predicted by Neural Events Before Change Onset. *Journal of Cognitive Neuroscience*, *18*, 2108–2129.

Proverbio, A.M., & Riva, F. (2009). RP and N400 ERP components reflect semantic violations in visual processing of human actions. *Neuroscience Letters*, *459*, 142–146.

Proverbio, A.M., Riva, F., & Zani, A. (2009). Observation of Static Pictures of Dynamic Actions Activates Movement-Related Brain Areas. *PLoS ONE, 4(5)*: e5389. doi:10.1371/journal.pone.0005389.

Pylyshyn, Z.W. (1973). What the mind's eye tells the mind's brain: a critique of mental imagery. *Psychological Bulletin*, *80*, 1–24.

Railo, H., Koivisto, M., & Revonsuo, A. (2011). Tracking the processes behind conscious perception: a review of event-related potential correlates of visual consciousness. *Consciousness and Cognition*, *20*, 972-983.

Ramachandran, V.S., & Oberman, L. (2006). Broken Mirrors – A theory of autism. *American Scientific*, *11*, 63–69.

Reddy, V. (1991). Playing with others' expectations: Teasing and mucking about in the first year. In: Whiten, A. (ed.). Natural Theories of Mind. Oxford: Basil Blackwell. pp. 143–158.

Reddy, V. (1998). ESRC report on person-directed play. Unpublished manuscript, University of Portsmouth.

Reddy, V. (2001). Positively shy! Developmental continuities in the expression of shyness, coyness and embarrassment. In: Crozier, W.R., & Alden, L.E. (eds.). International handbook of social anxiety: Concepts, research and interventions relating to the self and shyness. New York: Wiley. pp. 77–99.

Reddy, V. (2008). How infants know minds. Cambridge: Harvard University Press.

Reddy, V. (in print). Joining intentions in infancy. *Journal of Consciousness Studies*.

Reddy, V., Liebal, K., Hicks, K., Jonnalagadda, S., & Chintalapuri, B. (2012). The emergent practice of compliance: An exploration in two cultures. *Developmental Psychology*, *49*, 1754–1762.

Reddy, V., Markova, G., & Wallot, S. (2013). Anticipatory adjustments to being picked up in infancy. *PLoS ONE, 8(6)*: e65289. doi:10.1371/journal.pone.0065289

Reeves, B., Thorson, E., Rothschild, M., McDonald, D., Hirsch, J., & Goldstein, R. (1985). Attention to television: Intra stimulus effects of movement and scene changes on alpha variation over time. *International Journal of Neuroscience*, *25*, 241–255.

Reisz, K., & Millar, G. (1953). Technique of Film Editing. London: Focal Press.

Rensink, R.A. (2000). The dynamic representation of scenes. *Visual Cognition*, *7*, 17–42.

Rensink, R.A., O'Regan, K., & Clark, J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*, 368–373.

Rizzolatti, G. (2005). The mirror neuron system and its function in humans. *Anatomical Embryology*, *210*, 419–421.

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131–141.

Rizzolatti, G., Fogassi, L. & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Neuroscience Reviews*, *2*, 661–670.

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parietofrontal mirror circuit: interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*, 264–274.

Rochat, M.J., Caruana, F., Jezzini, A., Escola, L, Intskirveli, I., Grammont, F., Gallese, V., Rizzolatti, G., & Umiltà, M.A. (2010). Responses of mirror neurons in area F5 to hand and tool grasping observation. *Experimental Brain Research*, *204*, 605-666.

Rutherford, A. (2011). What Makes a Film Tick? Cinematic Affect, Materiality and Mimetic Innervation. Bern/ Berlin/Bruxelles/Frankfurt a. M./New York/Oxford/Wien: Peter Lang.

Sauer, M. (1999). Cezanne, Monet, Van Gogh – Genese der Abstraktion. http://archiv.ub.uni-heidelberg.de/artdok/volltexte//2573

Sauer, M. (2006). Faszination und Schrecken. Wahrnehmungsvorgang und Entscheidungsprozess im Werk Anselm Kiefers. *Zeitschrift für Ästhetik und allgemeine Kunstwissenschaft, 51,* 183–210.

Sauer, M. (2012). Zur Handlungsrelevanz ästhetischer Erfahrung anhand Anselm Kiefers Deutschland-bilder. http://archiv.ub.uni-heidelberg.de/artdoc/volltexte/2012/1851/

Sbriscia-Fioretti, B., Berchio, C., Freedberg, D., Gallese, V., & Umiltà, M.A. (2013). ERP modulation during observation of abstract paintings by Franz Kline. *PloS ONE, 8(10)*: e75241. doi:10.1371/journal.pone.0075241.

Schnall, S., Harber, K.D., Stefanucci, J.K., & Proffitt, D.R. (2008). Social Support and the Perception of Geographical Slant. *Journal of Experimental Social Psychology*, *44*, 1246–1255.

Schubotz, R., & Cramon, D.Y. (2004). Sequences of Abstract Nonbiological Stimuli Share Ventral Premotor Cortex with Action Observation and Imagery. *The Journal of Neuroscience*, *24*, 5467–5474.

Schröder, J. (1990). Die psychologische Realität von Prinzipien des Continuity Cinema. In: Schumm, G., & Wulff, H.G. (eds.). Film und Psychologie I. Kognition-Rezeption-Perzeption. Münster: Maks Publikationen. pp. 109–142.

Schwan, S., & Ildirar, S. (2010). Watching film for the first time: How adult viewers interpret perceptual discontinuities in film. *Psychological Science*, *21*, 1–7.

Shaviro, S. (1993). The Cinematic Body. Minneapolis: University of Minnesota Press.

Shepard, R.N. & Metzler, J. (1971): Mental rotation of three-dimensional objects. *Science*, *171*, 701–703.

Shmuelof, L., & Zohary, E. (2005). Dissociation between central and dorsal fMRI activation during object and action recognition. *Neuron*, *47*, 457–470.

Sierek, K. (2007). Foto, Kino und Computer. Aby Warburg als Medientheoretiker. Hamburg: Philo Fine Arts.

Simons, D.J., & Levins, D. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*, 261–267.

Simons, D.J., & Levins, D. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, *5*, 644–649.

Skinner, B.F. [1957] (1991). Verbal Behavior. Acton: Copley Publishing Group.

Smith, T. (2005). An Attentional Theory of Continuity Editing. PhD University of Edinburgh https://www.era.lib.ed.ac.uk/bitstream/1842/1076/1/smith_ATOCE_0506.pdf.

Sobchack, V. (1982). Toward inhabited space: The semiotic structure of camera movement in the cinema. *Semiotica*, *41*, 1–4.

Sobchack, V. (2004). Carnal Thoughts: Embodiment and Moving Image Culture. Berkeley: University of California Press.

Sontag, S. (1977). On Photography. London: Penguin.

Stancak, A., & Pfurtscheller, G., (1996). Event-related desynchronisation of central beta rhythms in brisk and slow self-paced finger movements. *Cognitive Brain Research*, *4*, 171–184.

Steinbeis, N., & Koelsch, S. (2008). Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cerebral Cortex*, *18*, 1169–1178.

Steinhauer, K., & Drury, J.E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain & Language*, *120*, 135–162.

Stern, D. (2000). Die Lebenserfahrung des Säuglings. Stuttgart: Klett-Cotta.

Stern, D. (2010). Forms of Vitality – Exploring Dynamic Experience in Psychology, the Arts, Psychotherapy and Development. Oxford: Oxford University Press.

Streltsova, A., Berchio, C., Gallese V., & Umiltà, M.A. (2010). Time course and specificity of sensory-motor alpha modulation during the observation of hand motor acts and gestures: a high density EEG study. *Experimental Brain Research*, *205*, 363–373.

Tan, E. S. (1995). Constraint and Convention in Psychological Film Aesthetics: On the Psychological Basis of Cinematic Conventions. In: Müller, J. (ed.). Towards a pragmatics of the audiovisual Münster: Nodus. pp. 67–86.

Toulouse, M., & Mourgue R. (1920). Des réactions respiratoires au cours de projections cinémato-graphiques. Online via http://1895.revues.org/3874#tocto2n1

Toro, C., Deuschl, G., Thatcher, R., Sato, S., Kufta, C., & Hallett, M. (1994). Event related desynchroniza-tion and movement-related cortical potentials on the ECoG and EEG. *Electroencephalography and Clinical Neurophysiology*, *93*, 380–389.

Umiltà, M.A., Berchio, C., Sestito, M., Freedberg, D., & Gallese V. (2012). Abstract art and cortical motor activation: An EEG study. *Frontiers in Human Neuroscience*, *6*, 311.

Umiltà, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti G. (2001). 'I know what you are doing': a neurophysiological study. *Neuron*, *32*, 91–101.

Ungerleider, L.G., & Mishkin, M. (1982). Two cortical visual systems, in: Ingle, D., Goodale, M., & Mansfield, R (eds.). Analysis of Visual Behaviour. Cambridge/London: MIT Press. pp. 549–586.

Urgesi, C., Moro, V., Candidi, M., & Aglioti, S.M. (2006). Mapping implied body actions in the human motor system. *Journal of Neuroscience*, *26*, 7942–7949.

Varela, F., Thompson, E., & Rosch, E. (1991). The Embodied Mind: Cognitive Science and Human Experience. Cambridge/London: MIT Press.

Vischer, R. (1874). Der aesthetische Akt und die reine Form. In: Vischer, R. (1927). Drei Schriften zum aesthetischen Formproblem. Halle. pp. 1–44.

Voss, C. (2011). Film Experience and the Formation of Illusion: the Spectator as ‘Surrogate Body’ for the cinema. *Cinema Journal*, *50*, 136–150.

Voss, C. (2013). Der Leihkörper. Erkenntnis und Ästhetik der Illusion. München: Fink Verlag.

Warburg, A. (1999). The Renewal of Pagan Antiquity. The Getty Research Institute.

Ward, D., & Stapleton, M. (2012). Es are good: Cognition as Enacted, Embodied, Embedded, Affective and Extended. *Advances in Consciousness Research*, *86*, 89–104.

Wicker, B., Keysers, C., Plailly, J., Royet, J.P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron*, *40*, 655–664.

Williams, P., & Simons, D. J. (2000). Detecting changes in novel, complex three-dimensional objects. *Visual Cognition*, *7*, 297–322.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomical Bulletin Reviews*, *9*, 625–636.

Wittgenstein, L. [1922]. Tractatus logica-philosophicus. In: Wittgenstein, L. (2003) Schriften, Band 1, Frankfurt: Suhrkamp Verlag.

Wölfflin, H. (1946). Prolegomena zu einer Psychologie der Architektur. In: Wölfflin H. Kleine Schriften. Basel. pp. 13–47.

Wölfflin, H. (1888). Renaissance und Barock. Eine Untersuchung über Wesen und Entstehung des Barockstils in Italien. München.

Worth, S., & Adair, J. (1972). Through Navajo Eyes: An Exploration in Film Communication and Anthropology. Bloomington: Indiana University Press.

Zacks, J.M., Speer, N.K., & Reynolds, J.R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General, 138*, 307–327.

Zahavi, D. (2007). Expression and empathy. In: Ratcliffe, M., Hutto, D. (eds.). Folk-psychology re-assessed. Dordrecht: Springer Verlag. pp 25–40.

Zahavi, D. (2008). Simulation, projection and empathy. *Consciousness and Cognition,* 17, 514–522.

Zahavi, D. (2010). Empathy, embodiment and interpersonal understanding: from Lipps to Schutz. *Inquiry*, 53, 285–306.

Zahavi, D. (2011). Empathy and direct social perception: a phenomenological proposal. *Philosophical Psycholoy Reviews, 2*, 541–558.

Zahavi, D. (2012). Empathy and mirroring: Husserl and Gallese. In: Breeur, R., Melle, U. (eds.) Life, subjectivity and art. Essays in honor of Rudolf Bernet. Dordrecht: Springer. pp 217–254.

Zahavi, D. (2012). Basic Empathy and Complex Empathy. *Emotion Review*, *4*, 81–82.

Acknowledgements

In the last three years, I have been crucially supported and want to thank my supervisor Vittorio Gallese, who initiated this research and then granted me immense trust while supporting me on all my scientific adventures; as well as Alessandra Umiltà who often enough had to stand in if I still felt the lack of help.

In the same vein, I want to thank my friends and colleagues Sebo Uithol and Martha Calbi, who showed me how scientific cooperation should be, as well as my "scientific background crew" including Michele Guerra, Daniel Hutto, Vasu Reddy, Martina Sauer, Joerg Fingerhut, Andreas Roepsdorff, Luciano Fadiga and many others who fostered and helped me immensely in all this time (and who I hope to keep for the future).

This all said, this thesis is undoubtedly dedicated to my wonderful mom, Angelika Heimann, who, in her silent and most respectful way, did not only turn these last four weeks of intense writing, from the sleepless and hopeless nightmare anticipated, to a luxurious stay in an oasis of peace miraculously providing delicious food, clean clothes and endless love and support (even in some rather nightmary nights), but who has also always before supported me in all possible ways and most importantly served as my biggest idol on my (continuing) way to grow into an adult.

Furthermore, I am sending all my love to those other steady rocks in my calm or troubled seas, that are the never changing though always challenging five men (plus two woman and three kids) of my life: my best friends Christian Grüny and Levien van Zon, my dad Klaus Heimann, my brother Felix Heimann, my "second mom and dad" Gisela und Günter Keller and the rest of my Wittener family: Miriam Grüny-Boy and Nellie, Jaro and Iva Grüny.

Also: my time in Parma would not have been the same without my colleagues and girl friends Marianna Ambrosecchia, Antonella Tramacere, Pia de Stefano, Miriam Kyselo and soulsister Evi Collins and – most of all – my "parmigianian" family, Antonella, Gianni, Giulia and Maria Stella Cantini – plus Giulio Fabri Poncemi – who gave me a house and a home in Italy, that I will always keep in my heart.

At the very end, the biggest hug to the magnific Anna Strasser who I apparently met in just the right moment and to whom I and this thesis owe a lot!