ESTHER EIJLERS

Emotional Experience and Advertising Effectiveness

On the use of EEG in marketing

Emotional Experience and Advertising Effectiveness

On the use of EEG in marketing

Emotional Experience and Advertising Effectiveness: On the use of EEG in marketing

Emotionele ervaringen en de effectiviteit van reclame:
Over de toepassing van EEG in marketing

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof. dr. R.C.M.E. Engels

and in accordance with the decision of the Doctorate Board.

The public defence shall be held on
Thursday, January 30, 2020 at 13.30 hours

by

Esther Eijlers
born in Vlissingen.

Erasmus University Rotterdam

Doctoral Committee

Promotor:
Prof. dr. ir. A. Smidts

Other members:
Prof. dr. ir. G.H. van Bruggen
Dr. D.J. Levy
Prof. dr. J.W. van Strien

Co-promotor:
Dr. M.A.S. Boksem

Erasmus Research Institute of Management – ERIM
The joint research institute of the Rotterdam School of Management (RSM)
and the Erasmus School of Economics (ESE) at the Erasmus University Rotterdam
Internet: www.erim.eur.nl

ERIM Electronic Series Portal: repub.eur.nl/

ERIM PhD Series in Research in Management, 487
ERIM reference number: EPS-2020-487- MKT
ISBN 978-90-5892-565-7

Design: PanArt, www.panart.nl
Cover: Illustration by Artemenko Valentyn, Shutterstock (www.shutterstock.com)

This publication (cover and interior) is printed by Tuijtel on recycled paper, BalanceSilk®
The ink used is produced from renewable resources and alcohol free fountain solution.
Certifications for the paper and the printing production process: Recycle, EU Ecolabel, FSC®, ISO14001.
More info: www.tuijtel.com

MIX
Papier van verantwoorde herkomst
FSC® C007225

Meaning of cover

In this dissertation, I investigate the relationship between brain activity that is measured using electroencephalography (EEG) in response to marketing stimuli - such as advertisements-, and the effectiveness of these advertisements. In addition, I study the underlying emotional processes that are associated with evaluating the advertisements. The image on the cover represents this in the following way:

We encounter concrete entities, or “physical things” in the world, which also include **advertisements**. These are represented in the picture by the realistic lady bug, butterfly, and flower around the head.

When the world presents itself to us, the brain reacts in a highly complex manner. This has its effect throughout the body and altogether this determines how we eventually **respond**. The gear mechanism refers to this cascade.

The brain and the rest of the body instigate, amongst others, experiences that we could categorize using emotion labels (such as “happy”). **Emotions** are often visualized by bursts of various colors (i.e., hues), as is the case in this picture. Notably, some blurring effects and circular lines are added here, implying movement. Interesting detail: *to move* can be applied in a physical as well as emotional context, see also the Latin verb movēre.

I used **EEG** to study the brain’s response. In order to record EEG, a cap - with electrodes fixed in it - is mounted on a person’s head. This is additionally represented in the picture by the colorful area, that is cap-like in its entirety, on top of the head.

Possibly, the image on this cover elicited “a positive feeling” when reflecting on it. This may be related to the use of bright colors and rustic elements (such as butterflies, flowers).

Thereby, the cover is an illustration of how even an image can implicitly convey or evoke more than one may presume it does.

Esther

Table of contents

Chapter 1

General introduction

Identifying and meeting human and social needs is at the core of marketing (Kotler & Keller, 2007). By understanding consumers, value can be added through the creation, delivery, and communication of products and services. Applying neuroscience methods and insights to the field of marketing theory and practice has become more and more popular for this purpose, and is referred to as neuromarketing. The first mentions of this term date back to the year 2002, in which the term was coined both in an academic context (Smidts, 2002; inaugural address about the prospects and opportunities of neuromarketing) and in a press release about the creation of a neuromarketing company called the "BrightHouse Institute" (Levallois, Smidts, & Wouters, 2019). Although some researchers have detected a gap between neuromarketing in practice versus neuromarketing in academia (Ariely & Berns, 2010), it is important to realize that in fact the integration between the two has been important to the emergence and successful development of the field. The neuromarketing field emerged from the cooperation between neuroscientists from academia, and entrepreneurs and businessmen (with academic credentials) from industry (Levallois, Smidts, & Wouters, 2019).

Two methods are particularly relevant to measure brain processes underlying consumer responses to marketing stimuli: electroencephalography (EEG) and functional magnetic resonance imaging (fMRI). The application of fMRI is most popular in academic neuromarketing research. Due to its high spatial resolution, it provides insight into where in the brain activity occurs in response to marketing stimuli; it is thus very informative on the decision processes underlying consumer behavior. Hereby, it is a valuable method to help generate, validate or extend theories in marketing (Plassmann, Venkatraman, Huettel, & Yoon, 2015). EEG is currently the most popular neuroimaging method in neuromarketing practice, particularly in ad testing (Smidts et al., 2014), mainly due to its affordability and actionability of results. In comparison to fMRI, the costs of EEG are relatively low, and its temporal resolution is high. The high temporal resolution is especially advantageous for marketing, because of the dynamic nature of marketing stimuli such as TV commercials. In addition, it enables monitoring consumer experience in dynamic contexts such as when watching a movie, playing an online video game, or when going through on online buying process. This high temporal resolution of

EEG thus provides a granular diagnosis of the dynamic stimulus, increasing the actionability of the metrics. For example, it enables assessing which scenes in an ad are crucial with regard to a memory-associated metric, and which scenes are not very impactful and thus could be removed from the ad.

To record EEG, electrodes are fixed in an elastic cap that is mounted on the head. The number of electrodes is typically 32 or 64 in the current context to cover the whole head, and openly study activity recorded at all sites (i.e., without prior assumptions on the location of activity). It thereby measures the potential for electrical current to pass between different sites at the scalp, expressed in voltage (Luck, 2005). Communication in the brain occurs through billions of interconnected cells, also called neurons, and the electrical activity picked up by EEG reflects the summed activity of millions of these neurons at the surface of the brain (Stern, Ray, Quigley, 2001).

Since the introduction of EEG by Hans Berger in 1929, it has been noted that changes in an individuals' engagement of an activity, co-occur with changes in frequency (i.e., the number of oscillations per second) and amplitude of the EEG signal. This resulted in a convention to define oscillations in the EEG signal in terms of different frequency bands (with specific topography) that are typically associated with psychological constructs such as attention. (In Chapter 4 I will more elaborately discuss the link between frequency bands and proposed underlying constructs).

In this dissertation, I investigate the brain's response elicited by marketing related stimuli using EEG, and the extent to which such a response is associated with advertising effectiveness. This research extends existing knowledge by elucidating two proposed aims of neuromarketing (Smidts, 2002). The first aim concerns neuromarketing offering additional insight into (ongoing) implicit psychological processes. The second aim concerns contributing to predictions of (market-level) behavioral responses or "advertising effectiveness". With respect to the first aim regarding implicit processes, I will specifically focus on emotional experience (see *Figure 1.1* for a schematic overview of the focus on topics within the specific chapters of this dissertation).

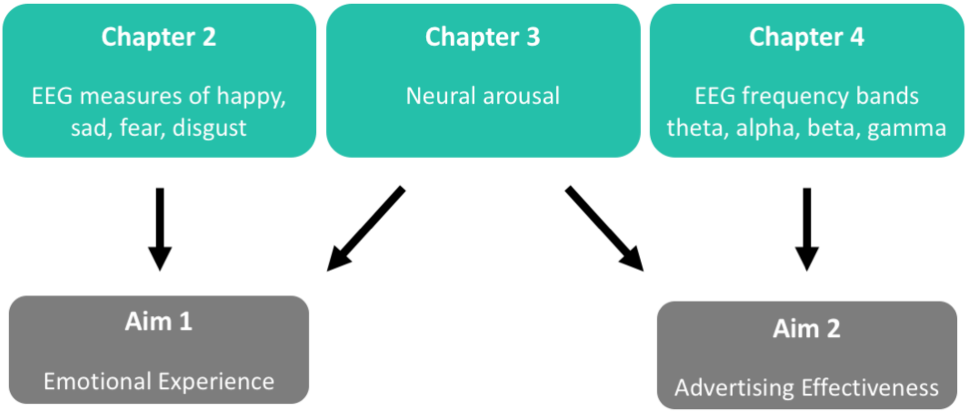


Figure 1.1. Schematic overview of focus on topics within chapters of the dissertation

Emotions are fundamental in guiding our behavior (Dolan, 2002), and as a consequence, they have been studied extensively in marketing (e.g., Bagozzi, Gopinath, & Nyer, 1999; Folkes, Koletsky, & Graham, 1987; Schmitt, 1999), and more specifically, in advertising (e.g., Burke and Edell 1989; Holbrook & Batra 1987; Pham, Geuens, & De Pelsmacker, 2013). Focusing on how a given marketing stimulus is processed and (emotionally) responded to, it is very insightful for managers to gain knowledge on the customer experience.

Emotions, however, are short-lived experiences that change over time and are not necessarily experienced consciously (Winkielman & Berridge, 2004). Yet, in previous marketing research, such emotional responses to advertisements have been mainly measured using self-reports. In a study by Baumgartner, Sujan, & Padgett (1997) for example, respondents were asked to indicate on a moment-by-moment basis how positive or negative they felt in response to advertisements. Although certainly not without merit, such reports are potentially biased by social desirability, and by the fact that they entail a cognitive interpretation during the emotional experience itself, rather than a direct measurement of the emotional response that is subject to change over time. Recording brain activity underlying both conscious and unconscious processes with high temporal resolution using EEG, could offer a solution here.

Outline of the Dissertation

Because it has proven difficult to measure emotional experiences unobtrusively, particularly for dynamic stimuli, the research in Chapter 2 ("Implicit measurement of emotional experience and its dynamics") focused on distinguishing different emotional experiences elicited by audiovisual stimuli that were designed to evoke particularly happy, sad, fear and disgust responses. I use a multivariate approach, and base supervised classification of the emotion categories on EEG activity that distinguishes between these four emotion categories. Unique features of the study are the absence of a priori assumptions about which features of the signal (e.g., frequency bands or scalp topography) would be predictive of distinctions between emotional experiences, and the retention of this kind of information in the data to be used for classifying emotional experiences. The advantage of this method is that it enables interpreting the observed differences between emotional experiences, based on the patterns of frequencies and their topography, in terms of underlying component psychological processes that are associated with these activation patterns. In addition, an illustrative application demonstrates how this method of classifying emotional experiences can be used on a moment-by-moment basis in order to track dynamic changes in the emotional response over time. Being able to monitor these specific emotional experiences "in real time" would be of great managerial relevance.

In Chapter 3 ("Arousal and advertising success: Neural measures suggest that arousing ads stand out more but are liked less"), the emphasis is shifted from understanding the stimulus itself, to additionally understanding the effect of the stimulus. Here, I investigate the extent to which an emotional neural response to advertisements is associated with evaluation of and behavior towards the advertisement at the population level. The emotional response I focus on is arousal, an important aspect of ad-evoked feelings (e.g., Holbrook & Batra, 1987). Arousal is a fundamental aspect of emotion in general, and is defined as the intensity or level of activation of one's (emotional) response (Lang & Bradley, 2010). The chapter takes a more theoretical approach by exploring the reverse inference problem (Poldrack, 2006), which is a common concern for most of the techniques used in neuromarketing practice today. Reverse inference refers to the validity of inferring that a particular psychological process is engaged (e.g., arousal) from the presence of a specific type of brain activity (e.g., reduced alpha band activity: reduced oscillations between 8-12 Hz). This deduction does not have to be valid, because not only arousal has been found to result in reduced alpha band activity (e.g., also attention, memory demands and general alertness suppress alpha oscillations, Klimesch, 2012), and also other frequency bands have been related to arousal (e.g., increased gamma band activity, 30-65 Hz, has been observed in response to emotional arousing pictures compared to neutral pictures, Keil et al., 2001; Muller et al., 1999).

In Chapter 3, I therefore diminish the reverse inference problem by first estimating how arousal is represented in the brain via a separate task, and thereafter use this representation to measure arousal in response to advertisements. Next, I estimate the relationship between this a priori identified process (arousal as measured by EEG) and external measures of ad effectiveness in the population at large (as measured by notability, attitude toward the ad, and choice, respectively) across two studies. Chapter 3 thereby also adds to elucidating a second proposed aim of neuromarketing: contributing to improving predictions of (market-level) behavioral responses or advertising effectiveness.

In the past decade, researchers in neuromarketing have started to investigate the possibility of using neural data collected in response to marketing stimuli in a relatively small group of people (denoted a 'neural focus group', Falk, Berkman, & Lieberman, 2012), to predict (the consequence of the stimulus on future) behavior of a larger group of people, or even real-world market level success. Chapter 4 ("EEG metrics relating to population-wide commercial success of movies: A meta-analysis") extends knowledge on this second aim of neuromarketing. In addition, it also contributes methodologically to the field by conducting a meta-analysis to explore whether predictive effects found in several single studies, are generalizable across studies and stimuli and also hold in a much larger sample. Indeed, many neuroimaging studies contain a relatively small number of participants and/or stimuli. In addition, notable differences between studies exist in pre-processing of the EEG data (e.g., presence of correction for eye movements), but also in the specific neural activity that is extracted, and in the measures of market level success concerning the stimulus.

In Chapter 4, I therefore combine data from four studies in which neural activity in response to a similar stimulus (here, movie trailers) was investigated using EEG, for a systematic re-analysis of all data (covering a total data set of n = 130 participants and k = 145 stimuli). I examined for five metrics (or frequency bands, that have been studied extensively: theta, alpha, alpha asymmetry, beta, and gamma, respectively) extracted from the EEG signal, whether the metrics are predictive of population-wide success of the corresponding movies expressed in terms of U.S. box office. Importantly, I test whether the neural activity measured in response to the movie trailers contributes above and beyond traditionally available information such as genre of the movie and self-reported liking measures. In Chapter 5, I conclude with a discussion of the research presented in the previous chapters and present suggestions for further research.

In sum, this dissertation extends existing knowledge by elucidating two proposed aims of neuromarketing: offering additional insight into emotional experience and contributing to predicting advertising effectiveness through measuring brain activity in response to marketing stimuli. In addition, the dissertation contributes methodologically to the field of neuromarketing in multiple ways. In Chapter 2 I investigate processing of the stimulus in an open, data driven manner using multivariate pattern analysis, in Chapter 3 I explore the problem of reverse inference, and in Chapter 4 I address the issue of neuroimaging studies using small samples. The managerial contribution of the present research consists of showing a novel way of measuring customer experience, in addition to demonstrating associations between EEG metrics measured in response to marketing stimuli and their effects at the population level.

Declaration of Contribution

For Chapter 2, I (EE) formulated the research question in collaboration with my daily supervisor Maarten Boksem (MASB) and promotor Ale Smidts (AS). Yuhee Kim helped with construction of the stimulus set and collection of the data. EE conducted the data analysis with input from MASB and AS. EE wrote the manuscript and implemented feedback from MASB and AS.

For Chapter 3, EE formulated the research question and conceptualized the studies in collaboration with MASB and AS. Nancy Detrixhe en Dennis Hoogervorst (Magazines.nl) assisted with the stimulus generation of the print ads and data collection of the population sample for Study 1. For Study 2, the same stimuli and population sample data were used as in the paper published by Linda Couwenberg, MASB, Roeland Dietvorst, Loek Worm, Willem Verbeke, and AS (see Couwenberg et al., 2017). EE collected the EEG data of Study 1 with assistance of Pauldy Otermans, and Study 2 with the help of Jia (Phyliss) Gai. EE conducted the data analysis with input from MASB and AS. EE wrote the manuscript and implemented feedback from MASB and AS.

For Chapter 4, EE formulated the research question in collaboration with MASB and AS. The data collection for this meta-analysis was executed by different parties depending on the source of the data. The data of Study 1, Study 2, and Study 3 have been collected at the Erasmus Behavioral Lab by Nigel Pouw, Jia (Phyliss) Gai, and Suvi Väisänen, respectively. The data for Study 4 was collected by the lab of Christoforos Christoforou (see Christoforou et al., 2017). EE conducted the data analysis with input from MASB and AS. EE wrote the manuscript and implemented feedback from MASB and AS.

EE wrote Chapter 1 and Chapter 5, and implemented feedback from MASB and AS.

Chapter 2

Implicit measurement of emotional experience and its dynamics[1]

Introduction

Emotions are fundamental in guiding our behavior; they are indices of events that we value or desire to different extents in our everyday lives (Dolan, 2002). Numerous studies have shown that emotions have a profound impact on cognition: emotions modulate attention (e.g. Armony & Dolan, 2002; Ohman, Flykt, & Esteves, 2001) and enhance memory for valuable events (e.g. Dolan, 2002; Phelps, 2004) in order to better predict occurrences of such events in the future. Emotions also influence social and economic decision-making (Elster, 1998; Loewenstein, 2000; Peters, Vastfjall, Garling, & Slovic, 2006) by acting as a motivator (Chen & Bargh, 1999; Peters et al., 2006), by providing information (Peters et al., 2006; Slovic, Finucane, Peters, & MacGregor, 2007), and by influencing the way we interact with others (Van 't Wout, Chang, & Sanfey, 2010).

However, the actual measurement of emotions has proved to be challenging. Emotion ratings acquired through self-report can potentially be distorted because of social desirability concerns (Fisher, 1993). That is, people may *not want to* express exactly how they feel. Even in the absence of these factors, it has been shown that people are very limited in their ability to reflect on their internal mental processes and to accurately report on these processes (Nisbett & Wilson, 1977). That is, they may *not even be able to* put their feelings into words accurately. Indeed, as affective processes largely occur outside our awareness (Zajonc, 1980), emotions do not have to be experienced consciously to influence judgement and behavior (Winkielman & Berridge, 2004). In addition, the task of consciously reporting on one's (unconscious) emotional state may actually change this state, potentially changing the relationship between emotion and subsequent behavior (Dholakia & Morwitz, 2002; Feldman & Lynch, 1988).

---

[1] This chapter is based on Eijlers, E., Smidts, A., & Boksem, M.A.S. (2019). Implicit measurement of emotional experience and its dynamics. *PLoS One, 14*: e0211496.

Neuroimaging methods may provide a solution to this problem by recording brain activity underlying both conscious and unconscious processes, without the need to consciously and cognitively reflect on them (Plassmann, Venkatraman, Huettel, & Yoon, 2015). Functional magnetic resonance imaging (fMRI) has been employed successfully to localize neural networks involved in many cognitive processes such as working memory or valuation of choice alternatives, while participants perform a task that engages one of these specific processes implicitly (see Wager & Smith, 2003; Bartra, McGuire, & Kable, 2013 respectively for meta-analyses). There also have been several fMRI studies in which the neural correlates of emotions are explored (see Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012 for a meta-analytic review). However, multiple meta-analyses have shown that there is little evidence for activity in any single brain region to be consistently and specifically associated with a specific emotion. This is why a multivariate pattern analysis (MVPA) approach has been suggested to be more appropriate for investigating emotions; to allow for the search of neural activation patterns that occur distributed (but simultaneously) across the brain (Kassam, Markey, Cherkassky, Loewenstein, & Just, 2013, but see Kragel & LaBar, 2016).

While the core advantage of fMRI is providing insight into the particular brain structures involved, it is less useful for gaining insight into how these neural processes evolve over time. Emotions are transient experiences (Fredrickson & Branigan, 2005), and people's (intensity of their) experienced emotions are subject to change under the influence of the external environment (Gross & Levenson, 1995). The dynamics of the emotional experiences have a critical impact on the subsequent (behavioral) response: People do not assess an affective experience based on the average experience, but instead rely heavily on the intensity of peak and final moments, also referred to as the peak-end rule (e.g., Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993; Do, Rupert, & Wolford, 2008). It would therefore be highly valuable to be able to decode and monitor discrete emotional responses relatively unobtrusively on a moment-by-moment basis. Being able to accurately measure the dynamics of emotions would serve many practical purposes in contexts such as media consumption, gaming, online buying, and other aspects of consumer experience, but also in clinical settings in which one is concerned with changes of the patient's emotions over time.

Electroencephalography (EEG) is a suitable alternative to fMRI against this background, with a lower spatial resolution but with a much higher temporal resolution. With EEG, the fluctuations in voltage that are measured by electrodes at the scalp reflect the summed activity of large, synchronously active, populations of neurons at the surface of the brain (Coles & Rugg, 1995). This (in combination with volume conduction) precludes accurate localization of the source of the measured activity. However, because electrical activity is measured directly (as opposed to via the hemodynamic response as with fMRI), the temporal resolution is retained.

Studies in the past decades have shown that oscillations in different frequency ranges or so-called frequency bands of the EEG signal, relate to specific psychological processes in the brain (Basar, Basar-Eroglu, Karakas, & Schurmann, 1999; see Knyazev, 2007 for review). With regard to emotions, early EEG studies (e.g., Perria, Rosadini, Rossi, 1961) have investigated positive versus negative affective experiences using the asymmetry in oscillatory activity between hemispheres. Although the initial studies suggested that greater left than right frontal activity was associated with the experience of positive affect, and greater right than left frontal activity with the experience of negative affect (Davidson & Fox, 1982), later studies revealed that the underlying factor was motivational direction (i.e., approach and withdrawal rather than positive and negative affect, respectively) (see Harmon-Jones, Gable, & Peterson, 2010 for review).

Going beyond emotional valence, measuring more specific emotions would provide more detailed information regarding an elicited response and its potential behavioral consequences. However, clear EEG correlates of specific emotional experiences have so far not been conclusively shown. As with fMRI, it is unlikely that specific emotions are associated with each their own particular EEG component. The aim of our study is therefore to use a multivariate approach in order to search for patterns of frequency distributions in the EEG data that distinguish different emotional experiences. In the current study, these experiences were elicited by audiovisual stimuli designed to evoke particularly happy, sad, fear and disgust emotions. It should be noted that we not necessarily measure the specific emotions happy, sad, fear, and disgust (if they exist), but rather *representations of emotional experiences*, as elicited by audio-visual stimuli, that can be grouped together and labeled as such. Thus, we use happy, sad, fear and disgust merely as descriptive labels for particular experiences as elicited by audio-visual stimuli.

In our multivariate approach, we based supervised classification of the emotion categories on activity that distinguishes between emotion categories. Importantly, we do not make a priori assumptions about which features of the signal (frequency bands or scalp topography) would be predictive of distinctions between emotional experiences. The advantage of this method is that we will be able to interpret the observed differences between emotional experiences, based on the patterns of frequencies and their topography, in terms of underlying processes that are known to be associated with these activation patterns.

We elicited the specific emotional experiences by displaying short videos that we selected for this purpose, as dynamic multimodal audiovisual stimulation represents the best and most natural way of eliciting emotions (Gross & Levenson, 1995; Baumgartner, Esslen, & Jäncke, 2006). Participants viewed five short clips for each of the four emotions under investigation, while their EEG was recorded. We then classified the emotional content of these clips based on the features (frequency and topography) of the EEG signal. Finally, we illustrate that the method we applied to classify emotional experiences can be used to track dynamic changes in the emotional response over time.

Methods

Participants

We recruited 40 students from the university population. They all had normal or corrected-to-normal vision and had no history of neurological illness. Before the experiment, written informed consent was obtained, and participants received 25 euro for their participation. Three participants were excluded from the analysis because of excessive artefacts in the reference channels and/ or channels recording the eye movements, precluding appropriate pre-processing of the data. The final sample therefore consisted of 37 participants (24 female) between 18 and 28 years (*M* = 22.2, *SD* = 2.6) of age.

Stimuli

We selected videos that would elicit a strong emotional response in the participants according to an expert panel. The content of the videos consisted of scenes from movies or documentaries and were selected to elicit one specific emotional experience (see Appendix 2.A and 2.B for details). The length of the video clips ranged from 22 seconds to 200 seconds (*M*= 96.2 s, *SD*= 36.9 s). More specifically, the happy videos had a mean duration of 112.2s (SD = 54.6), the sad videos 118.0s (SD = 14.3), the fear videos 82.0s (SD = 32.6), the disgust videos 58.0s (SD = 23.2). The videos eliciting happy and sad responses were relatively longer in duration than the videos eliciting fear and disgust, because eliciting happiness or sadness requires in general more time to build up in a context, whereas disgust and fear responses are more immediate without much need for context (see Appendix

2.D for robustness check 1 in which the analyzed segments have equal durations across emotion conditions).

We included a video clip from the beginning of the animated movie *Up* specifically to illustrate tracking of the emotional response over time, because this clip comprises a complete storyline (i.e., a summary of the lives of a man and woman that get together). In the first and main part of this video the content is predominantly happy, but at a certain point in the video the happy content clearly ceases to dominate while the sad content increases, allowing us to demonstrate content validity of our method when we track the emotional response over time.

Procedure

The Erasmus Research Institute of Management (ERIM) Internal Review Board granted approval to conduct the experiment (2016/04/26-44486mvb). The participants received written and verbal instructions on the task that they were going to perform upon arrival at the lab. Participants were unaware of the purpose of the study, but they were made aware that the videos that they were going to watch included content from the genres action, comedy, crime, horror, thriller, romance, drama, mystery and musical. We asked the participants to empathize with the people in the videos as much as possible, stay attentive and enjoy watching the videos. We notified participants beforehand of the presence of some intense scenes from movies and TV series. We did not mention that we would ask them to complete a questionnaire about the videos after the EEG recording.

During the EEG data collection, participants were seated in a slightly reclining chair positioned in front of a 19-inch PC monitor in a sound-attenuated, electrically shielded, dimly lit room. After showing the instructions again on the screen, the videos were presented in blocks, with each block consisting of five videos belonging to one of the four emotion categories happy, sad, fear or disgust. We reasoned that a block design was the best approach in order to induce and maintain the emotional experience optimally, rather than a design with rapid and constant switching between emotions. We randomized the order of the blocks as well as the videos within blocks, across participants. Between each block, a neutral video that contained part of a documentary was presented in order to return to a neutral or baseline emotional state. The videos were presented at a resolution of 1280 x 720, and the inter-stimulus interval, consisting of a black screen, was three seconds.

To verify the videos' effectiveness in eliciting the specific emotional responses in our participants, we asked participants to complete a questionnaire about the previously viewed videos after we finished the EEG data collection. Participants had to indicate for each video the extent to which they, respectively, felt happy, sad, fear, and disgust during the video on a scale from one (felt not at all e.g., happy) to five (felt extremely e.g., happy). In order to aid the recollection of (the experience of) the video, we provided a screenshot of a characteristic scene from that video before the question. The video screenshots and questions about the videos were presented in random order (i.e., not in blocks per emotion).

EEG recording and analysis

The EEG data was acquired using the BioSemi Active Two system with 64 active Ag-AgCl electrodes. Additional flat type electrodes were placed on the right and left mastoid, and in the eye region in order to record eye movements or electro-oculograms (EOGs): Electrodes were placed below and above the left eye in line with the pupil to record vertical EOGs, and at the outer canthi of both eyes to record horizontal EOGs. The EEG and EOG signals were sampled at a rate of 512 Hz. All preprocessing was done in Brain Vision Analyzer software (BVA; Brain Products). The data was first down-sampled to 256 Hz, then re-referenced to the averaged mastoids, and filtered with a low cutoff filter of 1 Hz with a slope of 48 dB/octave and a notch filter of 50 Hz. Thereafter, the data was segmented into 25 segments (one for each video), with segments lasting from the beginning to the end of the video. We then split the segments further into 50% overlapping segments of 256 data points. We applied Gratton and Coles ocular correction as implemented in BVA, and standard artifact detection and rejection criteria where segments were rejected that contained jumps larger than 30μV/ms, amplitude differences exceeding 150μV/ 200ms, and amplitude differences below 0.5 μV/ 100 ms. Note that only the channels that contained artifacts were deleted within the given segment, and not the entire segment. Then, data was decomposed into different frequencies (1-128 Hz) using a Fast Fourier Transform (FFT, using a 100% hanning window). Finally, we averaged the frequency data across all segments for each video, and for each participant separately. The resulting frequency data was exported to Matlab (Mathworks). Note that this results for each video in averaged frequency data across the entire video duration (but for each electrode, for each frequency), since our DV (emotion category labels) is also at the video level.

For the initial phase of the analysis described below, we only used the first part of the *Up* video for the representation of a happy emotional response. Based on the predominantly happy content of the first part of the video, we averaged frequency data across the first 200 seconds of the video. In order to track the emotional response over time during the complete video, we additionally exported the non-averaged frequency-domain data for the entire *Up* video per second (261 seconds in total).

Statistical analyses

After transforming the EEG data obtained during viewing of the videos to the frequency domain, we standardized (i.e., z-transformed) the data for every participant, electrode, and frequency across all videos. Further analyses consisted of two parts (one for classifying the emotional experiences happy, sad, fear, and disgust that were elicited by viewing videos, based on the patterns of frequency distributions observed in the EEG data, and one for illustration of tracking of the emotional response over time), each with multiple stages (see Appendix 2.C for more details on the statistical analyses).

For the first part, classifying the emotional experiences, we started with feature selection: Using a subset of the observations (i.e., a subset of the videos), we selected features (i.e., electrode-frequency combinations) that were most informative in distinguishing the specific emotions in order to reduce the dimensionality of the data. Per participant, electrode, and frequency, each emotion was contrasted with the average of the other three emotions. For each emotion then, one-sample *t*-tests across participants were applied to determine the 10% most informative features to use for classification (see Appendix 2.E for similar results with 5% and 20% features: robustness check 2). Thereafter, we proceeded with training and testing the classifiers. That is, with the remaining observations (i.e., those not used for feature selection), we trained support-vector machines (SVMs; six two-class classification models for the six combinations of four emotions, and also a multi-class model) on the selected features to generalize the distinction between emotional responses to new data, using cross-validation. We repeated feature selection and classifier training and testing 500 times, with for each repetition a different random subset of observations used for feature selection and thus also for the training and testing stage in order to rule out a selection bias as explanation of our results (see Appendix 2.G for robustness checks regarding the number of repetitions).

For the second part of the analysis, focused on tracking the emotional response over time, we applied a newly trained classifier to the complete video from the animated movie *Up*. We first performed feature selection and training of a classifier on happy, sad, fear and disgust emotional experiences elicited by the videos (i.e., computed a multi-class model), but this time we excluded the happy video *Up*, as well as for each other emotion the video with the lowest average rating on the emotional response it should have elicited. The rest of the analysis was similar to the analysis in part one, except for the final testing stage that was now replaced by a prediction stage. In the prediction stage, we used the trained classifier to compute the posterior probabilities that the emotional response was happy, sad, fear or disgust for every second of the *Up* video, averaged across participants. Since the content of the video becomes less happy over time (after approximately 200 seconds), and the reverse holds for the sad content, we show the contrast between the probability that the response is classified as happy versus sad.

Results

Participants indicated for each video the extent to which they felt happy, sad, fear, and disgust during viewing the video, after we finished the EEG data collection. This enabled us to verify the videos' effectiveness in eliciting the specific emotional responses in our participants (i.e., manipulation check). Based on the results of the manipulation check (see Appendix 2.C for details on statistical analyses and Appendix 2.H for results), we concluded that the emotional responses that the videos targeted to elicit, are indeed the emotions that the participants predominantly experienced during viewing of the videos. These results suggest that the EEG activity averaged across the duration of the videos, is representative of a happy, sad, fear, and disgust response, respectively, and that we can use this data to functionally localize specific emotion-related activity patterns.

Classifying the emotional experiences based on EEG data

Feature selection. The magnitude and sign of the *t*-values, averaged across the 500 repetitions, indicate how a specific emotional response differentiated from the other emotional responses at the different electrode-frequency combinations (see *Figures 2.1-2.4*). Note that we did not group the data into corresponding frequency bands in any of the analysis stages, but we merely did so here to provide an interpretable structure to the figures and results.

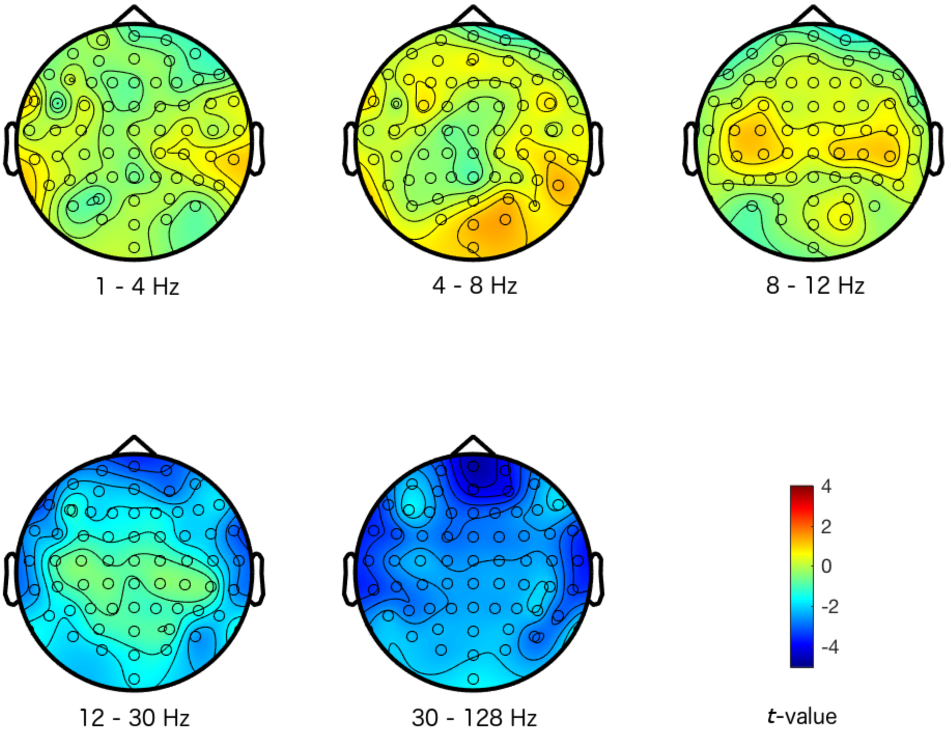


Figure 2.1. Maps of the difference between a happy response and the other emotional responses. The colors represent *t*-values. The different scalp maps show the contrast (expressed in *t*-values) between activity representing a happy response, and activity representing the other emotional responses for the specific frequencies that are indicated below the maps (the delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz), and gamma (30-128 Hz) frequency range respectively), and across the head for the 64 electrodes.

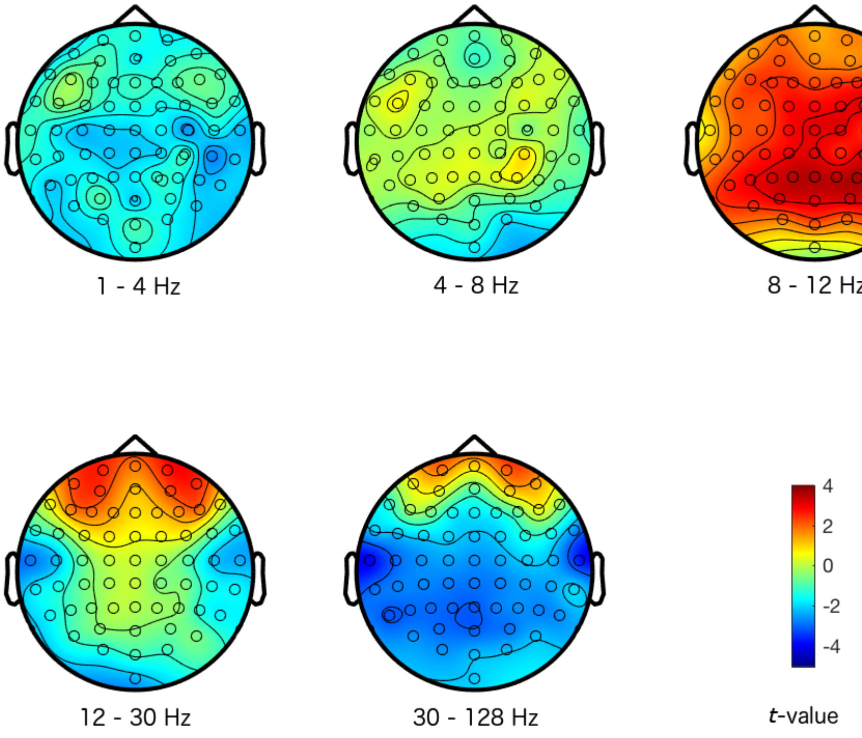


Figure 2.2. Maps of the difference between a sad response and the other emotional responses. The colors represent *t*-values. The different scalp maps show the contrast (expressed in *t*-values) between activity representing a sad response, and activity representing the other emotional responses for the specific frequencies that are indicated below the maps (the delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz), and gamma (30-128 Hz) frequency range respectively), and across the head for the 64 electrodes.

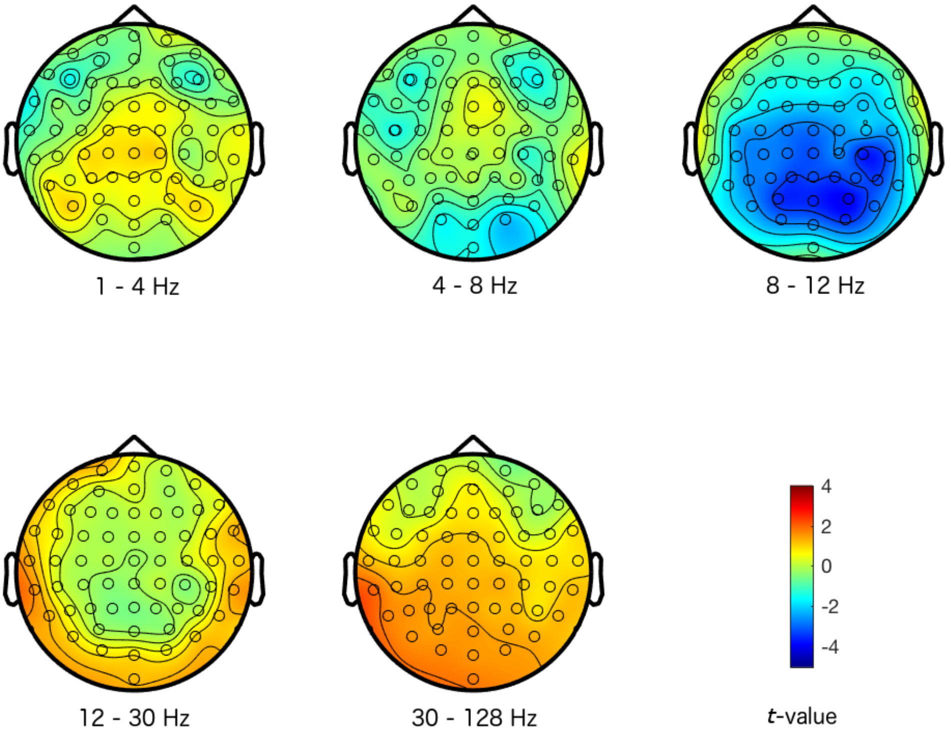


Figure 2.3. Maps of the difference between a fear response and the other emotional responses. The colors represent *t*-values. The different scalp maps show the contrast (expressed in *t*-values) between activity representing a fear response, and activity representing the other emotional responses for the specific frequencies that are indicated below the maps (the delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz), and gamma (30-128 Hz) frequency range respectively), and across the head for the 64 electrodes.

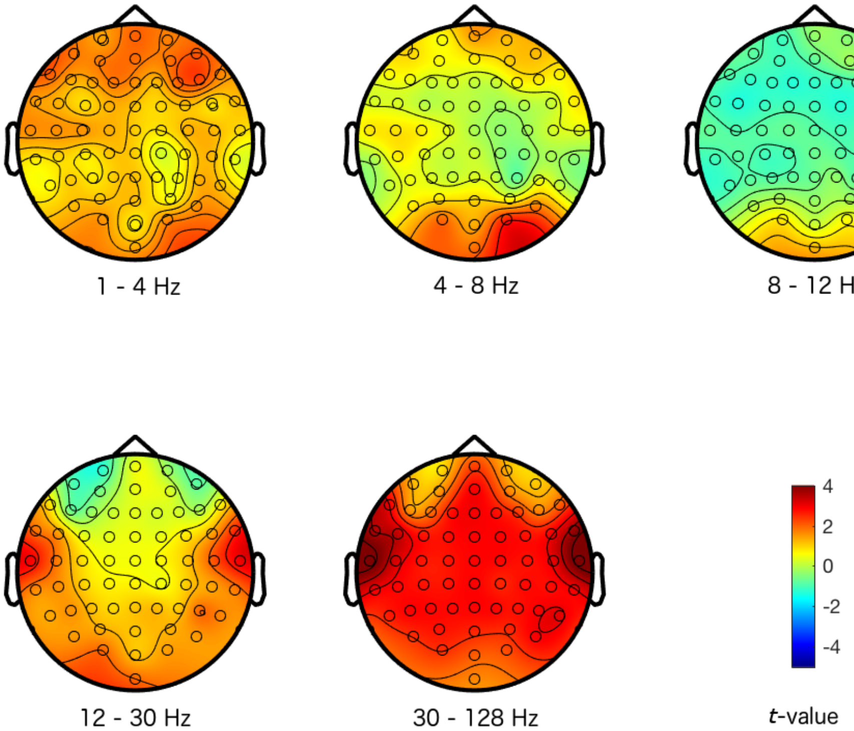


Figure 2.4. Maps of the difference between a disgust response and the other emotional responses. The colors represent *t*-values. The different scalp maps show the contrast (expressed in *t*-values) between activity representing a disgust response, and activity representing the other emotional responses for the specific frequencies that are indicated below the maps (the delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz), and gamma (30-128 Hz) frequency range respectively), and across the head for the 64 electrodes.

Inspecting *Figures 2.1-2.4*, suggests that the happy and disgust response differentiated most strongly from the other emotional responses in the higher frequency ranges. Happy was mostly associated with decreased gamma activity at frontal and temporal sites, and disgust was associated with increased gamma at temporal areas. The sad response most strongly differentiated from other emotional responses with more alpha activity present across the scalp. Finally, the fear response differentiated most clearly from other emotional responses in the alpha frequency range, with reduced alpha predominantly at centro-posterior sites (see *Figures 2.1-2.4* for more detailed differences between the emotions).

Across the 500 repetitions of the complete analysis, slightly different features were selected, and thus also used in the stages of training and testing the classifiers (see Appendix 2.I for how often the specific features were selected across the repetitions).

Classifier training and testing. Computing the out-of-sample generalization accuracy for all 500 repetitions, resulted in a distribution of accuracies indicating generalizability of the distinction between emotions to new data, for the seven classifiers (see Table 2.1).

Table 2.1. *Mean and percentiles for the distributions of out of sample generalization accuracies across 500 repetitions*

Classifier		Mean	Min.	2.5%	25%	50%	75%	97.5%	Max.
Fear	Disgust	71.37	60.36	64.86	69.37	71.62	73.42	77.48	79.73
Sad	Disgust	81.54	73.87	77.03	80.18	81.53	83.33	86.04	88.74
Sad	Fear	74.56	66.67	68.47	72.52	74.32	76.58	80.18	83.78
Happy	Disgust	77.05	68.92	71.62	75.23	77.03	79.28	81.98	85.59
Happy	Fear	76.14	67.57	70.72	74.32	76.13	77.93	81.53	85.14
Happy	Sad	78.18	70.27	72.52	76.13	77.93	80.18	83.33	86.49
Multi-class (all 4 emotions)		57.54	49.55	52.48	55.86	57.66	59.23	61.94	64.86

Since the models were trained on an equal number of category members (videos per emotion), theoretical chance level accuracy was 50% for the two-class models, and 25% for the multiclass models. The ability of the classifiers to generalize the distinction between emotions to new data was well above chance level, with the fear and disgust response being the most difficult to distinguish (median 71.62% out of sample generalization accuracy) and the sad and disgust response being the easiest to distinguish (median 81.53% out of sample generalization accuracy). The multi-class classifier was also well able to generalize the distinction between all four emotions to new data with a median accuracy of 57.66% compared to chance level of 25% (see Appendix 2.J for significant differences from benchmarks created by permuting emotion labels, which were approximately similar to theoretical chance levels). Although we did not intend to classify a 'neutral response' from the neutral videos that were presented between emotion blocks, including the neutral category in the classifiers yielded similar results (see Appendix 2.F for robustness check 3).

Illustration of tracking the emotional response over time

The first and main part of the *Up* video contains predominantly happy content, and at a certain point in the video, the content becomes clearly less happy, while the level of sad content increases. This allowed us to demonstrate content validity of our method, by applying a classifier that was trained on the four emotion categories to the EEG data obtained during viewing the *Up* video, as well as to illustrate the application of the method with high temporal resolution. *Figure 2.5* shows the posterior probabilities that the emotional response was happy or sad for every second of the *Up* video, averaged across participants and 500 repetitions (for illustrative purposes, we do not show the fear and disgust time courses in *Figure 2.5*; for the probabilities of the response being classified as each of the four emotions see Appendix 2.K. Classification of the emotional response is based on the same multi-class model in both figures, hence the only difference between the figures is the visibility of the fear and disgust time courses).

Figure 2.5 shows that although the emotional response of participants was mainly happy throughout the video as reflected by the estimated posterior probabilities, this response clearly decreased in the middle and also towards the end (while sadness is showing the opposite pattern), tracking the main ups and downs in the narrative.

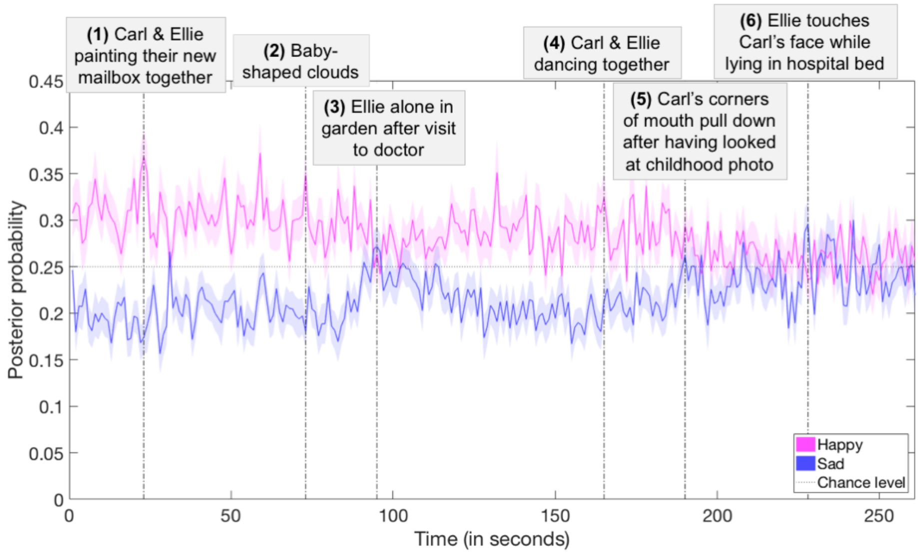


Figure 2.5. Dynamics of posterior probabilities for *Up*. Averaged across participants and 500 repetitions, with different observations used in the feature selection and classifier training stages across repetitions. The shaded areas indicate the standard deviation across repetitions. The vertical lines illustrate six examples of scenes at different moments in time, with moments (3), (5) and (6) indicating parts of the video that contain relatively more sad content. (1) Carl and Ellie just got married and are renovating the house: the posterior probabilities for a happy response are high. Once in a while they go on a picnic and look at the sky full of clouds. First, they see a cloud turn into an animal, then they see a cloud turn into a baby, and eventually at (2) all the clouds start to look like babies. (3) The sad part in de middle is elicited by the moment Ellie "gets told" in the hospital that she cannot have a baby: the posterior probabilities for a sad response rise briefly above chance level. After a short while, the couple picks up where they left off, and the distinction between the probabilities of a happy and sad response increase again. Over time, Carl and Ellie grow old and although they are still very happy with their lives together (4; Carl and Ellie dance together-scene), Carl realizes after having looked at an old photo (5) that much time has passed and their lives have not turned out the way they had hoped for. Eventually, we see Ellie in a hospital bed and at (6) she has just given back to Carl the book in which they had saved all their planned adventures. The posterior probabilities of a sad response rise above chance level from time to time and the happy response does not clearly dominate anymore.

Discussion

Although many studies revealed that emotions and their dynamics have a profound impact on cognition and behavior, it has proven difficult to unobtrusively measure these emotions at a high temporal resolution. In the current study, our objective was to distinguish between emotional experiences using EEG in order to be able to continuously track dynamic changes in the emotional response over time. We investigated how accurately we could classify the experiences labelled as happy, sad, fear, and disgust which were naturally elicited by viewing various videos, based on the distinct patterns of frequency distributions observed in the EEG data. In addition, we illustrated how this method of classifying emotions can be applied on a moment-by-moment basis in order to track the emotional response elicited by viewing a movie clip dynamically over time.

The results showed that the classifiers were able to generalize the distinctions between emotions to new data well above chance level. We obtained a mean accuracy of 58% (vs. 25% chance level) when differentiating between all four emotions, and between 71 and 82% (vs. 50% chance level) when differentiating pairwise between specific emotions. Fear and disgust were shown to be the most difficult to distinguish based on the mean attained accuracy of just under 72%. This is in agreement with the self-report ratings of the videos showing that the videos which were meant to elicit fear, also elicited disgust to some extent; more so than videos which were meant to elicit the other emotions. Nevertheless, presenting our set of videos appeared to be a natural and reliable way to elicit specific emotional experiences consistently among participants. This is demonstrated by the ratings being specifically increased for the emotion that the videos targeted to elicit, and the high agreement between ratings reflecting participants' similar feelings during viewing the multimodal dynamic stimuli.

Hence, we have demonstrated that we can distinguish between the specific emotional experiences happy, sad, fear, and disgust, and validated that the specific emotions that the videos were meant to elicit, corresponded with what the participants described to have experienced during viewing the videos. Thus, even though we cannot answer fundamental questions about the existence of (basic or specific) emotions in the brain (Shackman & Wager, 2019), the results do suggest that representations of these emotional experiences, described as happy, sad, fear and disgust by our participants, can be distinguished in EEG data.

One potential issue is that, even though we targeted specific emotional experiences, the more general underlying dimensions of valence and arousal may (partly) also underlie our results. Since the four emotions in the current design do not sample the affective space of valence (from negative to positive) and arousal (from calm to excited) evenly, it is not possible to examine whether valence and arousal processes (partly) drive our results. Nevertheless, the data suggests that valence and arousal cannot completely account for our results either: the emotions fear and disgust should be assigned to a very similar position, in the same quadrant (i.e., middle/ high arousal and negative valence) within the valence-arousal space, yet the classification model is well able to distinguish fear from disgust.

With the current results, we were able to inspect how the emotional responses actually differed from each other in terms of patterns of frequency distributions and topography, upon which the classification is based. Importantly, the current approach allowed us to speculate about the interpretation of the differences in neural activity between emotional experiences, in terms of more general underlying processes that are known to be associated with these activation patterns (e.g., Barrett & Wager, 2006). Alpha band activity (8-12 Hz), for example, has traditionally been related to the inverse of cortical activity. In a study combining EEG and fMRI registration in awake subjects at rest, alpha power correlated negatively with brain activity in parietal and lateral frontal cortices that are known to support attentional processes (Laufs et al., 2003). Indeed, many studies have shown that alpha activity is negatively associated with attention and task demands in general. More specifically, several EEG studies have shown that a decrease in posterior alpha power was associated with an increase in emotional arousal (e.g., DeCesarei & Codispoti, 2011; Simons, Detenber, Cuthbert, Schwartz, & Reiss, 2003, but see Aftanas, Varlamov, Pavlov, Makhnev, & Reva, 2002; Uusberg, Uibo, Kreegipuu, & Allik, 2013).

Our results seem to be in line with these observations: we found that activity in the alpha frequency band was increased for the sad response, but reduced for the fear response predominantly at centro-posterior sites, in comparison to the other emotional responses. These activity patterns potentially reflect that attention and arousal are more strongly engaged for fear, but that they are attenuated for sad responses.

While alpha band activity mainly distinguished between the sad and fear response, activity in the higher frequency ranges distinguished happy and disgust from other emotions. Inspecting the scalp topography of the distinctions in these higher frequencies for happy and disgust, showed that the differences were rather local (instead of widespread) and peripheral. This suggests that these distinctions between emotions may in fact reflect muscle activity. Although we did not record electromyogram (EMG) from the relevant facial muscles, activity from the temporal and frontal muscles represents the most common form of EMG activity that is picked up by EEG (mainly in the higher frequency bands). Contraction of these muscles is produced by jaw clenching and raising eyebrows respectively, which the EEG picks up near the active muscles at the periphery of the scalp (Goncharova, McFarland, Vaughan, & Wolpaw, 2003). We could therefore speculate that increased high frequency activity at temporal sites elicited by disgust videos could have been caused by clenching the jaws, whereas reduced activity at temporal and frontal sites for the happy response may reflect reduced tension in the jaws and less frowning (as sad videos also appeared to elicit more frontal high frequency activity, potentially related to more frowning).

A side effect of having used multimodal stimuli to elicit emotions naturally could have been that participants displayed facial expressions corresponding with the elicited emotions (they were not asked to actively supress them). This, however, should not pose a problem, and may even work to our advantage if this kind of muscle activity from the facial expressions naturally occurs with the elicited emotions, and thus can be used in combination with brain activity to distinguish between emotional responses. That is, for purposes of decoding, it does not matter very much whether the signals that are used originate from the brain, the face, or from elsewhere within the body.

These findings demonstrate the value of the approach of retaining the frequencies that are present in the data as features to base classification of emotional experiences on. Particularly, when classification of emotions would be investigated in future studies with other (audio-visual) stimuli, differences in classification accuracies may occur because of the specific stimuli that are used. In other words, the use of different stimuli across studies very likely will result in different classification accuracies across studies that may not be generalizable. However, with the current approach we could still assess the overlap between studies in terms of emotion-specific patterns of frequency distributions in the EEG data and their topography, reflecting interacting component psychological processes, despite any differences in stimuli and classification accuracies. That is, the emotion-specific EEG patterns on which the classification is based, may nevertheless be generalizable. Beyond offering insight into the processes underlying the differences between specific emotional experiences, this will ultimately aid generalization of the results and enable application of monitoring emotions over time in practice.

There have been some earlier attempts to classify specific emotions using EEG, but these studies have taken a different approach. Murugappan, Nagarajan, and Yaacob (2011) aimed at a maximal classification of the emotions happy, surprise, fear, disgust, and neutral based on statistical features that were extracted from the EEG signal. Their entropy measure performed well at emotion classification, but leaves the distinctions between emotions uninterpretable. Using different stimuli, Lin and colleagues (2010) presented their participants music to elicit emotions differing in valence and arousal. They averaged the frequencies present in the EEG data into five frequency bands (delta, theta, alpha, beta, gamma) and found that, based on asymmetries in oscillations, frontal and parietal electrode pairs were the most relevant in attaining the maximum classification accuracy. However, these authors did not attempt to classify specific emotions per se but rather the putative underlying dimensions of valence and arousal. Moreover, neither of these previous studies has shown whether it is possible to track emotions dynamically over time with their classification approach.

A unique feature of our study is the inclusion of an illustration of how this method of classifying emotions can actually be applied to relatively unobtrusively monitor emotional responses on a moment-by-moment basis. This tracking of the emotional experience is important, because emotions are, in essence, momentary experiences (Fredrickson & Branigan, 2005). In the current study, participants viewed a movie clip from the animated movie *Up* that was especially included because of its complete story arc, in order to track dynamic changes in the emotional response elicited by viewing the movie clip, over time. We used a classifier that was trained on videos with happy, sad, fear and disgust content to estimate the average happy and sad response across participants, second-by-second during the movie clip. It appeared that the emotional response, which was estimated based on the EEG data, was able to accurately track the main ups and downs of the narrative, demonstrating content validity. In other words, this illustrates that our classification approach could be generalized to other videos that are not limited to eliciting mainly one emotion to an extreme extent, at a high temporal resolution. Further research is however needed in order to confirm the differentiation of emotional responses over time for a more diverse set of dynamic stimuli.

With the methodology advanced here, future research could address how the evolving emotional experience over time, as measured using EEG, relates to subsequent cognition and behavior without interfering with or disrupting the emotional experience under investigation. The implications of being able to implicitly measure people's emotional response are numerous, and valuable in many contexts in which one is concerned with how a given stimulus is experienced over time. The user experience can provide information about attractiveness and appreciation in a variety of contexts ranging from clinical settings to consumer settings such as the consumption of digital media (such as movies, TV shows, broadcasted sports events), gaming, and online shopping.

To summarize, in the current study we elicited the emotional experiences happy, sad, fear, and disgust, and demonstrated that we could classify these using a multivariate approach. We retained all the frequencies that are present in the data, which allowed us to interpret the differences between emotions in terms of component psychological processes such as attention and arousal that are known to be associated with these activation patterns. The advantage of this approach is that it enables assessing the overlap between similar studies in terms of emotion specific patterns of frequency distributions in the EEG data and their topography. Additionally, we illustrated how this method of classifying emotional experiences can be applied on a moment-by-moment basis in order to relatively unobtrusively monitor dynamic changes in the emotional response as elicited by viewing a movie clip, over time.

Appendix

Appendix 2.A Details video clips

Supplementary Table 2.1. *Selected video clips*

Emotion category	Video content from	Duration (in s)	
Happy			
	500 Days of Summer	112	
	About Time	88	
	Love Actually	52	
	The Holiday	109	
	*UP**	200	
	Category mean	112.2	(54.6)
Sad			
	The Green Mile	135	
	The Help	106	
	Marley & Me	132	
	The Champ	110	
	The NeverEnding Story	107	
	Category mean	118	(14.3)
Fear			
	Anaconda	43	
	Annabelle	55	
	Friday the 13th - Part 2	96	
	Maze Runner – The Scorch Trials	123	
	The Ring	93	
	Category mean	82	(32.6)
Disgust			
	BuzzFeed Food	75	
	Fear Factor	80	
	Mr. Creostote (Monthy Python)	50	
	Pitch Perfect	22	
	Trainspotting	63	
	Category mean	58	(23.2)
Neutral (documentaries)			
	Wild Namibia	112	
	Modern Masters - Andy Warhol	100	
	The Archers of Butan	105	
	China's High-Speed Train	106	
	Megastructures - Burj Khalifa Dubai	130	
	Category mean	110.6	(11.7)

Note: *The complete video duration of Up used for the illustration of tracking the emotional response over time was 261 seconds

Video content description

Happy

500 Days of Summer (112 seconds)

You make my dreams come true scene. From the moment Tom steps outside, walks through the door, until he steps into the elevator and the doors close.

About Time (88 seconds)

Subway scene (How long will I love you). From the moment they close the door, walk down into the subway, until they take the escalator back up.

Love Actually (52 seconds)

Love actually is all around scene at the airport (Love Actually intro). From the moment two women hug each other just before the voice starts speaking "whenever I get gloomy with the state of the world", until just the words 'love actually' are visible.

The Holiday (109 seconds)

Arthur's award ceremony scene. From the moment Arthur and Iris step inside the building, until Miles runs into the room and says "the man is a rock star".

UP (200; 261 seconds)

Intro scene. From the moment the camera flashes to capture their wedding, until 200 seconds later when Carl prepared a picnic and Ellie starts walking up the hill (for classifying happy), and until Carl steps inside his house again and closes the door.

Sad

The Green Mile (135 seconds)

John Coffey "I'm tired" scene. From the moment Paul says "John, I have to ask you something very important now", until John asks if Paul can understand and Paul says "Yes John, I think I can".

The Help (106 seconds)

Bathroom/ burying scene. From the moment Minny finds Celia at the floor, until the moment Celia puts a plant in the ground.

Marley & Me (132 seconds)

"You're a great dog, Marley" scene. From the moment John hugs Marley and says "It's OK", until we have seen the video tape and Jenny for the second time.

The Champ (110 seconds)

Death, ending scene. From the moment TJ cries "Champ" and starts walking towards him, until he says to his father that he is not gone.

The NeverEnding Story (107 seconds)

Swamp of sadness scene. From the moment Bastian walks into the swamp with the horse and says "everyone knew that whoever let the sadness overtake him, would sink into the swamp", until he screams "I won't give up, don't quit".

Fear

Anaconda (43 seconds)

At the waterfall scene. From the moment Danny is struggling to get out of the water and we see from a different shot the snake disappearing and a body floating in the water, until the snake gets hold of Danny and we see Danny reaching out (the clip ends before Terri shoots).

Annabelle (55 seconds)

Little girl ghost scene. From the moment Mia lies on the ground and something hits her as she is getting up, until the ghost runs towards her, screaming, and Mia stands in the corner of the room.

Friday the 13th - Part 2 (96 seconds)

Surprise for Vickie scene. From the moment Vickie closes her umbrella and steps inside, until Vickie is standing against the wall and the knife is stabbed.

Maze Runner – The Scorch Trials (123 seconds)

Underground zombies (cranks) scene. From the moment they shine the flashlight onto the walls saying "Over here, look at this", until they run away, stand outside and stop running just in time.

The Ring (93 seconds)

Girl comes out of TV scene. From the moment the TV is being zoomed out of, and the phone is ringing, until Noah crawls away and starts turning around on his back.

Disgust

BuzzFeed Food (75 seconds)

The raw brains taste test. From the moment the presenter says “Hi, today I’m gonna be eating a variety of brains”, until he puts the fork with a piece of goat brain in his mouth and says “Ahh”.

Fear Factor (80 seconds)

Roach coach - eating rat hair chips, blood covered maggots and roaches. From the moment the Madagascar hissing cockroaches are released from a box and added to the rat hair chips, until there is 12 minutes and 35 seconds left on the clock, and the presenter screams “there you go Bobby, that’s what I’m talking about. Get crazy”.

Mr. Creostote (Monthy Python) (50 seconds)

Vomiting in restaurant scene. From the moment a waiter comes running with a bucket towards the table of Mr. Creostote, until the moment the waiter presents him the menu, and asks if he would care for an aperitif or if he prefers to order straight away.

Pitch Perfect (22 seconds)

Aubrey’s vomit scene. From the moment we see Cynthia Rose looking difficult and we hear a retching sound, appearing to come from Aubrey who starts to vomit, until Chloe starts screaming that they could have been champions.

Trainspotting (63 seconds)

The toilet scene. From the moment we see the dirty toilet pot and Renton decides to enter the little room anyway, until he dives into the toilet pot and he lifts his feet.

Appendix 2.B Confirmation of elicited emotions by videos

In order to show that the videos we selected indeed predominantly elicited the experience of happy, sad, fear, and disgust (and not predominantly other emotions we did not inquire about), we ran an MTurk study with a free response format. 262 MTurkers each viewed four videos, with each video belonging to one of the four emotion categories (in total 20 videos with number of views per video varying from 50 to 56). Per video, we asked the participants which emotion they predominantly experienced during viewing the video. Before analyzing the results, two independent raters classified the MTurk responses as being valid responses or not, since some responses were clearly not valid (e.g., 'big snake', 'black guy speaking', 'teeth'). The proportion of agreement on validity of responses was .59, as expressed by Cohen's Kappa (good interrater agreement). The raters *agreed* that 54 out of the 273 unique responses should be considered *invalid*. Only the responses agreed on as invalid were deleted. Examples of disagreement across the two raters on validity of responses were 'crying', 'comradery', and 'ugly'. One participant was removed because he/she did not comment on any video, two participants were removed because they replied with 'bored' in response to all videos, and 23 because they replied with only one valid response across the four videos they watched (thus questioning the motivation of the responder). Single invalid responses were removed as well. The final analysis consisted of 933 responses (mean N across videos = 47, SD = 2.9, minimum N = 39 (*The Green Mile*), maximum N = 51 (*Trainspotting*)).

In order to analyze how many respondents answered with a label that corresponded to the hypothesized emotion the video predominantly elicited, we searched for the presence of the following words or word stems in the responses (see Supplementary Table 2.2 and 2.3).

Supplementary Table 2.2. *Responses with synonyms for emotion category labels*

Emotion category label	Corresponding words and word stems
Happy	happ*, joy*
Sad	sad*, grief, tearful
Fear	fear, afraid, fright*, scar*
Disgust	disgust

Supplementary Table 2.3. *Responses with broader synonyms for emotion category labels*

Emotion category label	Corresponding words and word stems
Happy	happ*, joy*, amuse*, excit*, cheer*, content, glad, glee, good, pleasant
Sad	sad*, depress*, upset, downcast, grief, tearful
Fear	fear, afraid, anxi*, creepy, fright*, scar*, spooked, terr*
Disgust	disgust, gross*, nausea, sick

In Supplementary Table 2.4, we present the percentage of respondents that answered with (synonyms for) the emotion category labels happy, sad, fear, and disgust, and in Table 2.5 we included the more broader synonyms for the emotion category labels. We additionally included 'love' as category label (word stems lov*, roman*), since there were a significant number of responses corresponding to this category (for the happy videos; *Love Actually* and *About Time*, 22% and 24% respectively, still considerably less than the happy responses, even when classified according to strict synonyms, 46% and 54% respectively). All other responses were categorized as 'other'. These other responses were quite diverse, ranging from labels belonging to the same 'emotion family' (most frequently), to labels describing a different feeling (more rarely). Examples of responses categorized as 'other' for the happy videos were: warm, hope*, touched, moved, inspir*, encouraged, sympathy, connected, confident, energetic. Examples of responses categorized as 'other' for the sad videos were: pity, sorrow, compassion, pain, remorse, empathy, sympathy, loss, impressed, cry* (cry* was not taken into account as sad response because it is a description of what occurs in some videos). Examples of responses categorized as 'other' for the fear videos were: thrill*, anticipation, nervous, unease, suspense, crying, curiosity, excitement, stress, panic, horr*, intrigued. Examples of responses categorized as 'other' for the disgust videos were: repulsed, unpleasant, annoy*, irritat*, nasty, ugly.

Supplementary Table 2.4. *Percentage of responses for different emotion categories (strict synonyms)*

Video category/ Response	Happy	Sad	Fear	Disgust	Love	Other
Happy	55.98	4.27	0.43	0.43	11.11	27.78
Sad	0.43	73.39	2.58	0.00	0.86	22.75
Fear	0.00	2.12	69.07	1.69	0.00	27.12
Disgust	3.04	3.04	2.61	67.83	0.00	23.48

Supplementary Table 2.5. *Percentage of responses for different emotion categories (including broader synonyms)*

Video category/ Response	Happy	Sad	Fear	Disgust	Love	Other
Happy	61.54	4.27	0.43	0.43	11.11	22.22
Sad	0.86	74.25	3.00	0.00	0.86	21.03
Fear	1.69	2.12	76.27	1.69	0.00	18.22
Disgust	5.65	3.04	2.61	72.17	0.00	16.52

The results in the tables show that even in a free response format, the participants accurately label the predominantly experienced emotions as happy, sad, fear and disgust (corresponding to the hypothesized elicited emotion). The tables show the average percentage of responses across videos within an emotion category, but this was also true for every single video. Thus, we show that asking the participants to indicate the extent to which they felt happy, sad, fear, and disgust on a five-point scale, produces highly similar results to an open response format question.

Appendix 2.C Detailed statistical analyses

Manipulation check. We first checked whether the videos were indeed effective in eliciting the targeted emotional response in the participants based on self-reported ratings of the emotions during viewing the videos (see Appendix 2.H for results of the manipulation check and Supplementary Table 2.10). We started with verifying whether video category label had an effect on ratings of the different experienced emotions using a MANOVA. This main test was followed up by performing a repeated measures ANOVA for each of the categories of experienced emotions, to investigate whether the ratings of one specific experienced emotion were indeed higher for the videos that targeted the corresponding emotional response, compared to videos that targeted other emotional responses. Video category label served as IV (happy, sad, fear, disgust, neutral), and rating averaged across the five videos belonging to the category label served as DV. In addition to testing the effect of video label on the extent to which participants reported experiencing a particular emotion, we tested whether the four experienced emotions were rated differently within one video category. We checked whether the experienced emotion that was targeted by the videos, received a higher rating compared to the experienced emotions that were not targeted, by performing repeated measures ANOVAs for each of the video category labels (except for neutral), with experienced emotion as IV (happy, sad, fear, disgust), and rating averaged across the five videos belonging to the category label as DV. In order to check whether all participants experienced the emotions to a similar extent during viewing of the videos, we computed the intraclass correlation coefficient (ICC) for each of the experienced emotion categories that were rated using a two-way mixed-effects model where we were interested in the absolute agreement between raters.

Classification analyses. After transforming the EEG data obtained during viewing of the videos to the frequency domain, we standardized (i.e., z-transformed) the data for every participant, electrode, and frequency across all videos. Further analyses consisted of two parts with multiple stages.

In the first part of the analysis we investigated how accurately we could classify the emotional experience that was elicited by viewing a variety of videos, based on the patterns of frequency distributions in the EEG data. This part of the analysis consisted of three stages.

The first stage was feature selection. Using a subset of the observations (all the videos for all participants), we selected features (i.e., electrodes, frequencies) that were the most informative in distinguishing the emotions, in order to reduce dimensionality of the data and provide insight into the differentiation of emotional responses. We started with randomly selecting two videos out of the five videos, for each of the emotions, for each of the participants. Using these two videos for each emotion, we created contrasts between emotions to find the most informative features. For each participant, and for each electrode and frequency, we determined the mean activity across the two videos for each emotion. From this mean, for each emotion, we subtracted the mean of the other three emotions, combined from six videos. This results in a value that indicates how distinctive that electrode and frequency for a participant is regarding one emotional response versus the average of the other three emotional responses. In order to detect similarities in these distinctions across participants, we computed for every emotion one-sample t-tests across participants, for every electrode and frequency. Finally, in order to discover the most distinctive and thus informative features, we selected for every emotion the electrodes and frequencies with the 10% highest t-values (see Appendix 2.E for similar results with 5% and 20% features: robustness check 2). We used the most informative features from all emotions combined in the subsequent stages.

Second, in order to associate the emotion labels with patterns in the EEG data, we used classification models in the form of SVMs with a linear kernel function. Having detected the most informative features, we only used these electrodes and frequencies from the EEG data as features in the SVM models for the remainder of the observations that were not used for feature selection (i.e., three videos for each emotion and participant). After having divided the remaining observations into 10 different folds with approximately equal representations of the emotions, the training stage employed 9 out of the 10 folds, referring to 9/10 of these remaining observations. Since we were not only interested in how well the emotions can be classified in general, but also in a comparison of the specific emotions one-by-one, we computed six two-class classification models for the six combinations of four emotions (happy-sad, happy-fear, happy-disgust, sad-fear, sad-disgust, fear-disgust). In addition, a multi-class classification model was estimated containing all four emotion categories. We approached the multi-class classification problem using the Error-Correcting Output Codes (ECOC) framework (Dietterich & Bakiri, 1995), in which multiclass learning is reduced to multiple (SVM) binary learners. We used an Error-Correcting Output Codes (ECOC) classifier as implemented by Matlab (R2016b; Statistics and Machine Learning Toolbox) with a one-versus-all coding design in the case of classifying all four emotions.

All this resulted in training seven classifiers to associate the emotion labels with patterns in the EEG data. We did not intend to classify a 'neutral response' from the neutral videos that were presented between emotion blocks, but including the neutral category in the classifiers yielded similar results (see Appendix 2.F for robustness check 3).

Third, we tested the trained classifiers on the remaining data. We iterated stage two and three ten times for all possibilities of leaving out one of the ten folds. We then evaluated the ability of the classification models to generalize the distinction between emotions to new data by calculating the percentage of accurately predicted emotion labels across observations from all ten folds (i.e., the out of sample generalization accuracy).

An important quality of the analysis design is that the same data is not used twice, for reduction of the data and training/ testing of the models. This means however, that the specific subset of observations that we used in the different analysis stages, are of influence on the features that are selected as the most informative ones in the first stage, and also ultimately on the accuracy of the SVM models in the final stage. Since we did not want to select the most informative features based on specific videos, we decided to randomly select a subset of videos (two per emotion and per participant) for feature selection, leaving the remainder of the observations for the train and test stage. In the end, we repeated all of the stages described above 500 times, with for each repetition a different random selection of observations for the feature selection and thus the train and test stage (see Appendix 2.G for robustness checks regarding the number of repetitions). These 500 repetitions resulted in a distribution that approximates all the possible values that our measures of interest (the generalization accuracies in the final stage) can adopt, and thus rules out a selection bias as explanation of our results.

In the second part of the analysis, we classified happy, sad, fear, and disgust emotional experiences from the happy, sad, fear, and disgust videos, and thereafter applied the trained classifier to EEG data obtained during viewing of the complete *Up* video on a moment-by-moment basis, in order to track the probability that the emotional response is happy or sad over the time course of the video. In this analysis, we first removed the happy video *Up* from the data, as well as the sad video *The Help* which had the lowest average rating for a sad response, the video *Anaconda* which had the lowest average rating for a fear response, and the video *Pitch Perfect* which had the lowest average rating for a disgust response (see Appendix 2.H), in order to keep the number of videos equal across emotion conditions and to prevent training a classification model biased towards one emotional response. With the remaining four videos for happy, sad, fear, and disgust, we proceeded through stage one of feature selection and stage two of training a classifier similarly to in the first part of the analysis (except that the contrast in feature selection was now based on one video per emotion per participant instead of on two). Instead of testing the classifier at the third stage, we used the classifier to compute the posterior probabilities that the emotional response was happy or sad for every second of the *Up* video, and we averaged these probabilities across participants. We repeated all of the stages 500 times also for this second part of the analysis, with for each repetition different observations used for the feature selection and training (and prediction) stage. This again resulted in a distribution that approximates all the possible values that our measure of interest (here, posterior probabilities) can adopt, and thus rules out a selection bias as explanation of the results. For the probabilities of the response being classified as each of the four emotions see Appendix 2.K.

Appendix 2.D Robustness check 1: Similar duration of analyzed segments across emotions

In order to empirically test whether classification was potentially biased due to the difference in duration of the videos (thus, possibly a systematically different signal to noise ratio in the EEG data) across emotion conditions, we conducted a robustness check. We repeated the complete analysis with only the data in response to the final 22 seconds of each video (i.e., the duration of the shortest video) in order to eliminate the effect of duration difference across emotion conditions. We find that the resulting classifiers were still able to generalize the distinction between all four emotional experiences to new data well above chance level, even when merely the final 22 seconds of the data were taken into account for all videos. Hence, we can exclude that different durations of the videos across emotions biased the classification results. We can also conclude from these results that there is indeed information specific to the emotions in the extra length of the videos from the sad and happy categories, since especially this distinction is less accurately predicted by the final 22 seconds compared to the complete data (i.e., approximately a 10 percentage-point decrease).

Supplementary Table 2.6. *Mean and percentiles for the distributions of out of sample generalization accuracies across 500 repetitions, using only the final 22 seconds of each video*

Classifier		Mean	Min.	2.5%	25%	50%	75%	97.5%	Max.
Fear	Disgust	64.91	53.15	58.56	62.16	64.86	67.34	71.62	76.58
Sad	Disgust	76.03	66.22	70.27	73.87	76.13	78.15	81.98	84.68
Sad	Fear	70.60	61.26	64.41	68.47	70.72	72.52	76.58	81.08
Happy	Disgust	73.02	63.06	66.67	70.72	72.97	75.23	79.73	81.53
Happy	Fear	69.95	62.16	63.51	67.57	69.82	72.52	76.13	77.93
Happy	Sad	67.50	59.01	60.81	65.32	67.57	69.82	73.87	75.68
Multi-class (all 4 emotions)		46.10	38.51	40.99	43.92	46.17	47.97	51.58	54.28

Appendix 2.E Robustness check 2: Changing the number of features that are used for classification

In order to explore the influence of the number of features used in the classifiers, we varied the percentage of selected features between 5%, 10%, and 20% of the highest *t*-values. The results in the tables below show that this only had a marginal impact on the ultimate out-of-sample accuracies.

Supplementary Table 2.7. *Mean and percentiles for the distributions of out of sample generalization accuracies across 500 repetitions, using the 5 percent most informative features*

Classifier		Mean	Min.	2.5%	25%	50%	75%	97.5%	Max.
Fear	Disgust	70.30	57.21	63.06	68.02	70.27	72.52	77.48	80.18
Sad	Disgust	81.14	73.42	76.13	79.28	81.08	82.88	86.04	87.84
Sad	Fear	74.16	65.32	68.47	72.07	74.10	76.13	80.18	83.33
Happy	Disgust	76.44	67.57	71.17	74.32	76.58	78.38	81.98	85.59
Happy	Fear	75.41	67.57	70.27	73.42	75.23	77.48	81.08	83.33
Happy	Sad	78.49	69.37	72.97	76.58	78.38	80.63	83.78	85.59
Multi-class (all 4 emotions)		56.37	48.65	51.35	54.73	56.31	58.11	61.26	64.19

Supplementary Table 2.8. *Mean and percentiles for the distributions of out of sample generalization accuracies across 500 repetitions, using the 20 percent most informative features*

Classifier		Mean	Min.	2.5%	25%	50%	75%	97.5%	Max.
Fear	Disgust	71.82	61.71	65.77	69.82	72.07	73.87	77.03	79.73
Sad	Disgust	81.67	74.77	77.03	80.18	81.53	83.33	86.49	88.74
Sad	Fear	74.97	66.22	69.37	72.97	75.23	77.03	80.63	83.33
Happy	Disgust	77.32	69.37	71.62	75.23	77.48	79.73	82.88	85.59
Happy	Fear	76.92	68.92	71.17	74.77	76.58	78.83	81.98	84.23
Happy	Sad	77.90	67.12	72.07	76.13	77.93	79.73	82.88	87.84
Multi-class (all 4 emotions)		58.05	50.68	53.60	56.53	58.11	59.68	62.61	64.86

Appendix 2.F Robustness check 3: Classifying the neutral videos viewed in-between emotion blocks

Including 'neutral' as an emotion category yielded out of sample generalization accuracies for the classifiers that were in a similar range to the classifiers currently described in the main manuscript. This can be regarded as another robustness check since the neutral trials were not presented sequentially in one block to participants, but instead presented between the other emotion blocks, yet the classifiers were able to recognize these separate neutral trials as a category distinct from the trials belonging to the other emotion categories. Supplementary Table 2.9 shows that happy versus neutral was the most difficult to distinguish based on the classifier's attained accuracy. This is in agreement with the ratings of the videos showing that the videos which were meant to be neutral, clearly elicited more of a happy feeling (mean rating across neutral videos 2.43), than a sad, fear, and disgust feeling (mean rating across neutral videos respectively 1.34, 1.20, 1.12) (see Appendix 2.H Results manipulation check, Supplementary Table 10).

Supplementary Table 2.9. *Mean and percentiles for the distributions of out of sample generalization accuracies across 500 repetitions, including neutral as category*

Classifier		Mean	Min.	2.5%	25%	50%	75%	97.5%	Max.
Disgust	Neutral	77.69	68.02	72.52	75.68	77.93	79.28	83.33	86.49
Fear	Neutral	74.09	65.32	68.92	72.07	74.32	76.13	79.73	83.33
Fear	Disgust	72.08	61.26	65.77	69.82	72.07	74.32	78.38	80.63
Sad	Neutral	74.75	66.22	69.82	72.97	74.77	76.58	79.73	81.53
Sad	Disgust	81.98	74.32	77.03	80.18	81.98	83.78	86.49	88.74
Sad	Fear	75.32	65.77	69.82	72.97	75.23	77.48	81.08	83.33
Happy	Neutral	63.84	54.50	57.21	61.71	63.96	66.22	70.27	72.07
Happy	Disgust	77.53	70.27	72.52	75.68	77.48	79.28	82.43	84.23
Happy	Fear	77.13	70.72	72.07	75.23	77.03	78.83	82.43	85.14
Happy	Sad	78.53	70.72	72.97	77.03	78.83	80.18	83.33	87.84
Multi-class (all 5 emotions)		49.84	44.14	45.23	48.29	49.91	51.35	54.59	56.58

Appendix 2.G Repetition of the analysis with different observations in the different stages to rule out a selection bias

An important quality of the analysis design is that the same data is not used twice in the different analysis stages, but this also means that the specific subset of observations that we used in the stages are of influence on the outcome. In the first stage, a specific subset of observations influences the features that are selected as the most informative ones, and in the final stage it influences the out of sample generalization accuracy of the SVM models. Since we did not want to select the most informative features based on specific videos, we decided to randomly select a subset of videos for feature selection. For each of the 37 participants, we selected for each of the 4 emotions, randomly 2 videos out of the 5 videos ('ABCDE'), which leaves us with 10 possible selections of videos per emotion per participant (AB, AC, AD, AE, BC, BD, BE, CD, CE, DE) for feature selection. This means that there are 10^(4x37) possibilities available to split the data for feature selection versus training/ test set, rendering an extensive computation of accuracies for all possible video selections unfeasible. We therefore repeated the complete analysis including the three stages multiple times in order to investigate the influence of using a specific subset of observations in the first stage, leaving the remainder of the observations for the final stages. For every repetition of the analysis, we selected a different random subset of observations in the first stage for feature selection (keeping the numbers across participants and emotions constant), thus a different subset of observations remained for the second and third training and testing stage. We compared the distributions of accuracies for 10 repetitions, 50, 100, 500, 1000 and 5000 repetitions of the complete analysis. Although the distributions of generalization accuracies (from the final stage) did not change drastically from 10 to 5000 repetitions, at 500 repetitions, the distribution was virtually indistinguishable from 5000 repetitions. Therefore, we repeated the complete analysis with all of the stages 500 times, each time with a different selection of observations for the stages, to be able to create a distribution that approximates all the possible values that our measures of interest (i.e., *t*-values in the feature selection stage, generalization accuracy in the final testing stage) can adopt, and thus to rule out a selection bias.

Appendix 2.H Results manipulation check

Because we wanted to verify the videos' effectiveness in eliciting the specific emotional responses in our participants (i.e., manipulation check), we asked participants to complete a questionnaire about the previously viewed videos after we finished the EEG data collection. Participants had to indicate for each video the extent to which they felt happy, sad, fear, and disgust during the video on a scale from one (not felt at all e.g., happy) to five (felt extremely e.g., happy).

As expected, the MANOVA indicated an interaction between video label (i.e., the target emotion) and the emotion participants reported they actually experienced, using Pillai's trace ($V = 0.98$, $F(9, 28) = 129.89$, $p < .001$). The follow-up ANOVAs showed that the ratings of a specific experienced emotion were higher for the videos that targeted this particular emotional response, compared to videos that did not target this emotional response ($F_{\text{Happy ratings}}(4,144) = 158.72$, $F_{\text{Sad ratings}}(4,144) = 121.32$, $F_{\text{Fear ratings}}(4,144) = 78.90$, $F_{\text{Disgust ratings}}(4,144) = 186.71$, all $p < .001$, with follow-up Bonferroni corrected pairwise comparisons also $p <.001$ for the videos targeting the specific emotional response compared to the videos targeting the other emotional responses). That is, we checked that participants reported to feel happier during viewing happy videos compared to during viewing sad, fear, disgust, or neutral videos (and more sad during sad videos, compared to happy, fear disgust, and neutral videos, etc.). The *F*-values are presented in the Supplementary Table 2.10 together with ratings of the experienced emotions for the corresponding videos. The second set of ANOVAs revealed that the experienced emotion that was targeted, received higher ratings compared to the experienced emotions that were not targeted ($F_{\text{Happy videos}}(3,108) = 341.85$, $F_{\text{Sad videos}}(3,108) = 145.19$, $F_{\text{Fear videos}}(3,108) = 67.51$, $F_{\text{Disgust videos}}(3,108) = 153.58$, all $p <.001$, with follow-up Bonferroni corrected pairwise comparisons also $p <.001$ for the targeted experienced emotion compared to the three other non-targeted emotions). That is, we also checked that participants reported to feel happier during viewing happy videos than they reported to feel sad, fear and disgusted (and feel more sad during sad videos, than to feel happy, fear, and disgusted, etc.). The intraclass correlation coefficient (ICC) of each of the experienced emotion categories for the mean of multiple ratings was $ICC_{\text{Happy}} = .98$, $ICC_{\text{Sad}} = .98$, $ICC_{\text{Fear}} = .97$, $ICC_{\text{Disgust}} = .99$.

Based on the results of the ANOVAs and the ICCs, we conclude that the emotional response that the video targeted to elicit, is indeed the emotion that the participants predominantly experienced during viewing of the videos. These results suggest that the EEG activity averaged across the duration of the videos, is representative of a happy, sad, fear, and disgust response, respectively, and that we can use this data to functionally localize specific emotion-related activity patterns.

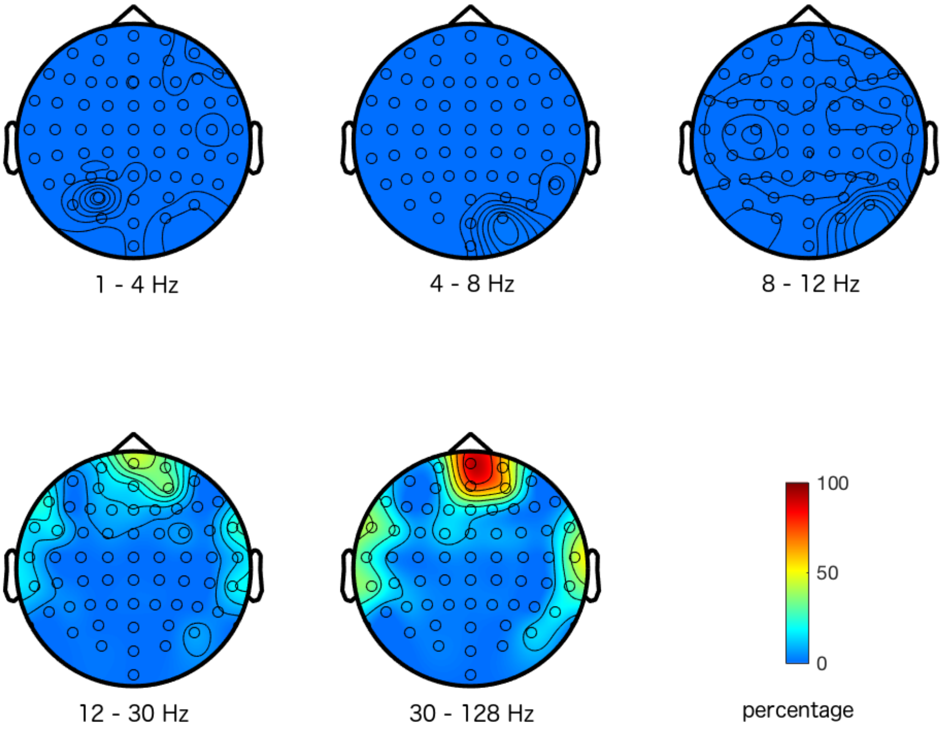
Supplementary Table 2.10. *Mean (SD) ratings of videos across the 37 participants*

	Happy		Sad		Fear		Disgust		RM ANOVA $F(3,108)$
Happy videos									
500 days of summer	4.11	(0.81)	1.05	(0.33)	1.00	(0.00)	1.00	(0.00)	
About time 2	3.73	(1.07)	1.46	(0.69)	1.16	(0.55)	1.11	(0.46)	
Love actually	3.97	(0.99)	1.95	(1.05)	1.16	(0.60)	1.05	(0.33)	
The holiday	3.76	(1.01)	1.38	(0.86)	1.19	(0.46)	1.05	(0.23)	
Up	4.11	(0.81)	2.22	(1.42)	1.22	(0.67)	1.00	(0.00)	
Category mean	3.94	(0.94)	1.61	(0.87)	1.15	(0.46)	1.04	(0.20)	341.85*
Sad videos									
Green mile	1.24	(0.55)	3.54	(1.12)	2.08	(1.23)	1.38	(0.86)	
Marley and me	1.19	(0.46)	3.86	(1.23)	1.70	(1.02)	1.22	(0.63)	
The champ	1.11	(0.31)	3.89	(1.02)	1.84	(1.09)	1.22	(0.67)	
Help	1.24	(0.49)	2.95	(1.31)	1.41	(0.90)	1.22	(0.53)	
The neverending story	1.19	(0.46)	3.70	(1.22)	2.49	(1.35)	1.30	(0.57)	
Category mean	1.19	(0.46)	3.59	(1.18)	1.90	(1.12)	1.26	(0.65)	145.19*
Fear videos									
The ring	1.22	(0.53)	1.38	(0.72)	3.68	(1.38)	2.57	(1.30)	
Friday the 13th	1.22	(0.58)	1.65	(1.18)	3.38	(1.34)	2.08	(1.23)	
Anaconda	1.30	(0.62)	1.86	(0.98)	2.62	(1.32)	2.05	(1.27)	
Anabelle	1.11	(0.39)	1.43	(0.93)	3.38	(1.40)	1.30	(0.66)	
Mazerunner	1.16	(0.50)	1.22	(0.63)	3.32	(1.36)	2.59	(1.40)	
Category mean	1.20	(0.53)	1.51	(0.89)	3.28	(1.36)	2.12	(1.17)	67.51*
Disgust videos									
Buzzfeedfood	1.81	(1.05)	1.24	(0.55)	1.49	(0.69)	3.46	(1.37)	
Fear factor	1.62	(0.95)	1.54	(0.90)	2.00	(1.20)	4.54	(0.77)	
Mr. Creostote	1.86	(1.11)	1.19	(0.52)	1.14	(0.48)	3.57	(1.19)	
Pitch perfect	2.19	(1.41)	1.11	(0.39)	1.08	(0.28)	3.16	(1.17)	
Trainspotting	1.43	(0.96)	1.24	(0.55)	1.43	(0.83)	4.43	(0.87)	
Category mean	1.78	(1.10)	1.26	(0.58)	1.43	(0.70)	3.83	(1.07)	153.58*
Neutral videos									
Namibia (wild life)	2.49	(0.99)	1.68	(1.11)	1.43	(0.77)	1.19	(0.57)	
Andy Warhol	2.30	(1.15)	1.30	(0.66)	1.11	(0.39)	1.19	(0.52)	
Butan (archery)	2.62	(1.11)	1.43	(0.69)	1.11	(0.39)	1.11	(0.39)	
China (Maglev)	2.49	(1.19)	1.22	(0.58)	1.16	(0.44)	1.05	(0.33)	
Dubai (elevators)	2.24	(1.32)	1.05	(0.23)	1.19	(0.46)	1.05	(0.23)	
Category mean	2.43	(1.15)	1.34	(0.65)	1.20	(0.49)	1.12	(0.41)	/
RM ANOVA $F(4,144)$	158.72*		121.32*		78.90*		186.71*		

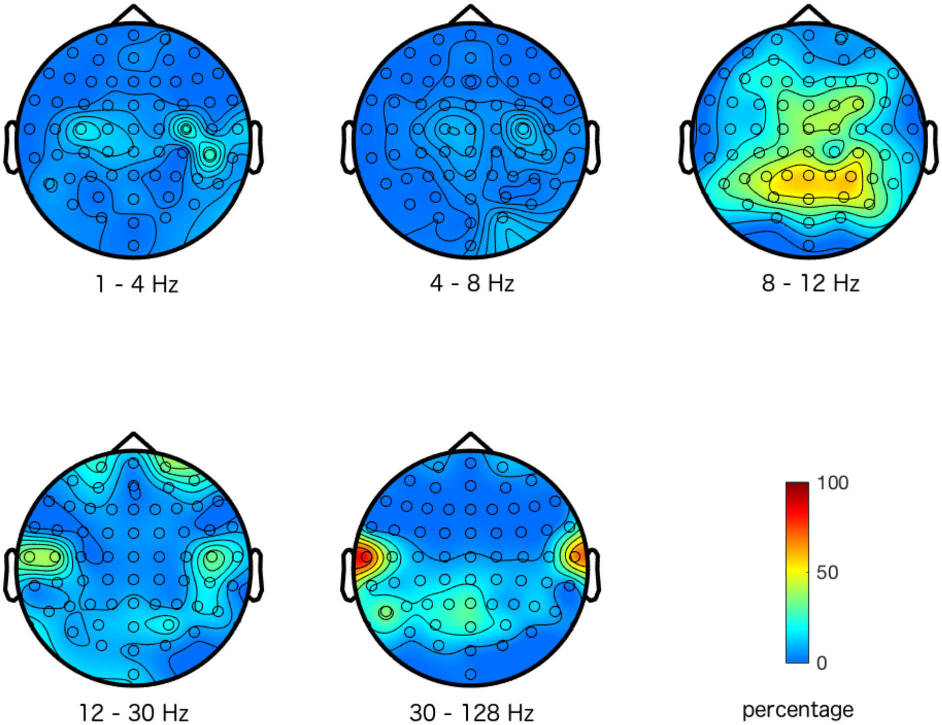
Note: * $p < .001$

Appendix 2.I Selection of the most distinctive features for classification

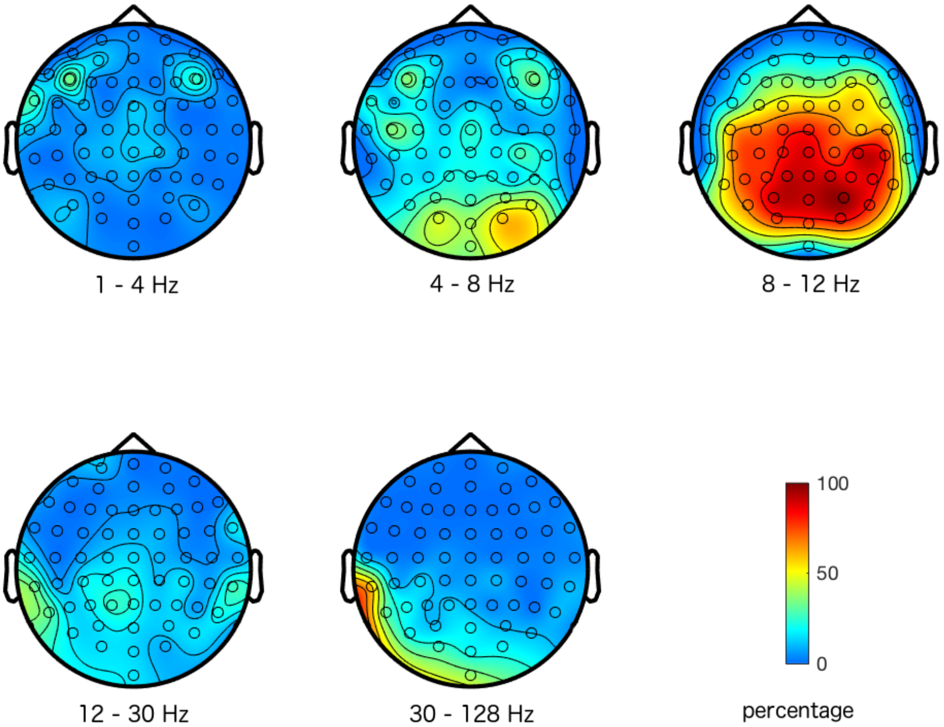
In the subsequent figures, we present how often the features were selected, and thus were used in training and testing the classifiers taking as criterion the 10% highest *t*-values per emotion, expressed as a percentage across the 500 repetitions.



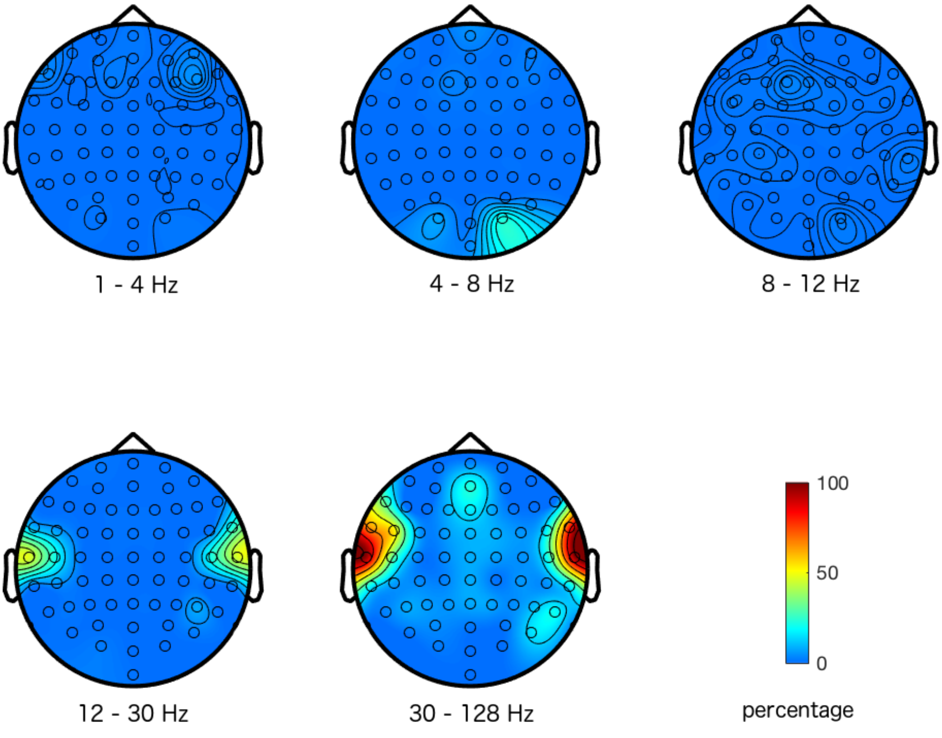
Supplementary Figure 2.1. Color maps of how often the features are selected as most distinctive for the happy response and thus used in training and testing the classifiers. The colors represent the percentage of repetitions that the specific frequencies (mentioned below the different scalp maps), and electrodes (across the heads) were selected as the features with 10% highest *t*-values for the happy contrast.



Supplementary Figure 2.2. Color maps of how often the features are selected as most distinctive for the sad response and thus used in training and testing the classifiers. The colors represent the percentage of repetitions that the specific frequencies (mentioned below the different scalp maps), and electrodes (across the heads) were selected as the features with 10% highest *t*-values for the sad contrast.



Supplementary Figure 2.3. Color maps of how often the features are selected as most distinctive for the fear response and thus used in training and testing the classifiers. The colors represent the percentage of repetitions that the specific frequencies (mentioned below the different scalp maps), and electrodes (across the heads) were selected as the features with 10% highest *t*-values for the fear contrast.



Supplementary Figure 2.4. Color maps of how often the features are selected as most distinctive for the disgust response and thus used in training and testing the classifiers. The colors represent the percentage of repetitions that the specific frequencies (mentioned below the different scalp maps), and electrodes (across the heads) were selected as the features with 10% highest *t*-values for the disgust contrast.

Appendix 2.J Permutation-based benchmark for models' generalization accuracies

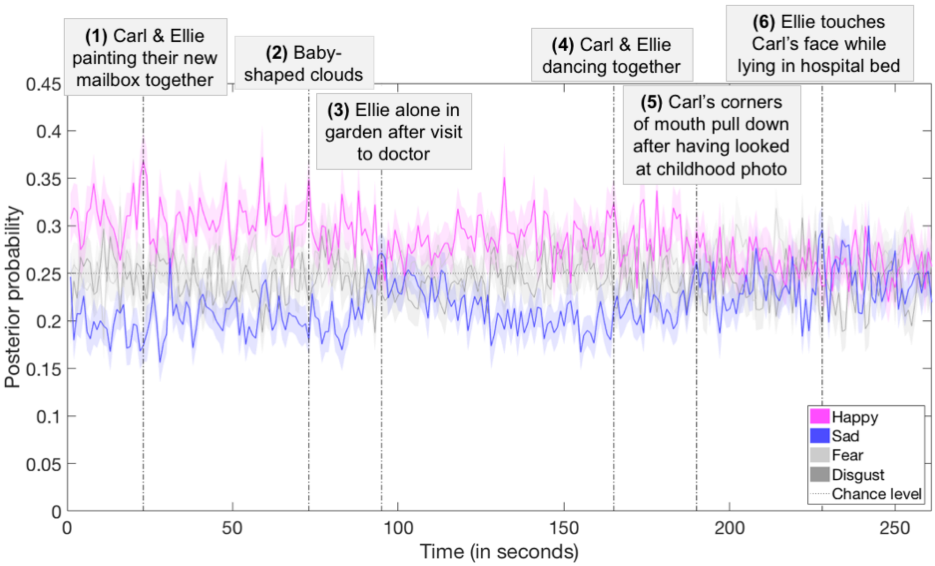
In order to determine whether there is a significant amount of information present in the EEG data that is emotion-specific, we created a permutation-based benchmark to compare the performance of the models against, instead of simply testing the models' performance to the theoretical (a priori) chance level. We repeated the training stage and the testing stage, but with random permutation of the emotion labels for the training set, and with testing of the models' performance on the true labels, keeping the folds similar to the original classification procedure. We repeated this procedure ten times with different random permutations of the training data, and averaged the generalization accuracy across these ten repetitions. We repeated this analysis also 500 times in a similar fashion to our original analysis (i.e., each repetition we used different cases for the feature selection and training/ testing stages), resulting in similar distributions of generalization accuracies as for the original data. Computing the out of sample generalization accuracy for the 500 repetitions using a model trained on random data showed that the median of the distribution was around theoretical chance level, and the maximum achieved accuracy (i.e., 54.46%) was still clearly lower than the minimum of the models trained on original data (not taking into account the multi-class model, for which the randomly trained model's accuracy achieved maximally 26.98%) (see Supplementary Table 2.11). This means that the two distributions with generalization accuracies from the original models and with generalization accuracies from the random models did not overlap. Non-parametric and distribution free Kolmogorov-Smirnov tests confirmed this (all $D=1.00$, $p<.00001$), hence the EEG data contained a significant amount of information that is class-specific and could be used in order to predict the emotional responses from participants based on their EEG activity.

Supplementary Table 2.11. *Mean and percentiles for the distributions of out of sample generalization accuracies across 500 repetitions, using permuted emotion labels*

Classifier		Mean	Min.	2.5%	25%	50%	75%	97.5%	Max.
Fear	Disgust	49.93	47.16	47.88	49.19	49.91	50.70	51.98	52.48
Sad	Disgust	49.96	46.67	47.79	49.21	49.95	50.72	52.12	53.06
Sad	Fear	49.90	47.16	47.57	49.17	49.82	50.68	52.12	53.78
Happy	Disgust	49.92	47.21	47.88	49.19	49.86	50.63	52.07	53.29
Happy	Fear	50.02	46.80	47.84	49.35	50.05	50.72	52.07	54.46
Happy	Sad	49.94	46.85	47.66	49.14	49.98	50.77	51.98	53.24
Multi-class (all 4 emotions)		25.01	23.40	23.76	24.55	25.02	25.41	26.33	26.98

Appendix 2.K Emotional response classification for *Up* as happy, sad, fear or disgust

Supplementary Figure 2.5 shows the posterior probabilities that the emotional response was classified as happy, sad, fear or disgust for every second of the *Up* video, averaged across participants and 500 repetitions (for illustrative purposes, we did not show the fear and disgust time courses in *Figure 2.5* in the main paper). Classification of the emotional response is based on the same multi-class model in both figures, hence the only difference between the figures is the visibility of the probability that the response is classified as fear and disgust.



Supplementary Figure 2.5. Dynamics of posterior probabilities for *Up* with four emotions. Averaged across participants and 500 repetitions, with different observations used in the feature selection and classifier training stages across repetitions. The shaded areas indicate the standard deviation across repetitions.

Chapter 3

Arousal and advertising success: Neural measures suggest that arousing ads stand out more but are liked less[2]

Introduction

Marketing, and more specifically advertising, is an important means for companies to convey to consumers what they have to offer and thereby bring together supply and demand. Moreover, the successfulness of advertising is an important factor in increasing the sales of products and services. Multiple marketing research techniques are being applied to develop effective ads and to assess their effectiveness. Currently, these methods range from self-report measures and focus groups to the more innovative autonomic measures and brain imaging techniques in small so-called 'neural focus groups' (e.g. Ariely & Berns, 2010; Poels & DeWitte, 2006). In the current study, we specifically investigate the impact of arousal in response to advertisements as measured with EEG and assess the relationship between this 'neural arousal' measure and measures of advertising effectiveness in an independent larger sample of the population.

Theoretical framework

Ad testing refers to assessing the consumers' response to advertisements before spending money on broadcasting of an advertisement. These responses include consumers' judgements or cognitions about advertisements, but, critically, also consumers' first reactions and 'feelings' in response to ads are deemed to be important (e.g. Burke & Edell, 1989). Abundant research has established that ad-evoked feelings are strong predictors of consumers' response to advertising (Burke & Edell, 1989; Edell & Burke, 1987; Holbrook & Batra, 1987, Morris, Woo, Geason, & Kim, 2002; Olney, Holbrook, & Batra, 1991; Pham, Geuens, & De Pelsmacker, 2013; Stayman & Aaker, 1988).

[2] This chapter is submitted for review

An important aspect of ad-evoked feelings is arousal (e.g., Holbrook & Batra, 1987). Arousal is a fundamental aspect of emotion and is defined as the intensity or level of activation of one's (emotional) response (Lang & Bradley, 2010). Over the years, multiple studies in advertising have shown a positive relationship between measures of arousal and advertising effectiveness consisting of attitude toward the ad (Holbrook & Batra, 1987; Olney et al., 1991), attitude toward the brand (Holbrook & Batra, 1987), conative attitude (i.e., purchase intention; Morris et al., 2002), and viewing time (Olney et al., 1991).

In all the above-mentioned studies, consumers' feelings in response to advertisements have been measured using self-reports. Although certainly not without merit, such reports, particularly when they involve reporting on feelings or other internal states, do have their limitations. In addition to the fact that people are limited in their ability to reflect accurately on their internal mental processes, concerns with social desirability compound this problem (Nisbett & Wilson, 1977; Wiles & Cornwell, 1990). Moreover, the ability to report on mental states requires cognitive processing, which may interfere with or even change the originally evoked feelings (Poels & Dewitte, 2006). Although the use of non-verbal graphical scales such as the well-known Self-Assessment Manikin (SAM) may reduce the amount of cognitive processing required (e.g. Morris et al., 2002), it still requires introspection and reflection on mental processes which largely occur outside our awareness (Zajonc, 1980).

Implicit measures in advertising research

Implicit techniques that measure physiological consequences (of psychological antecedents) in response to advertisements, share the advantage that they are free of the cognitive biases described above. A relatively new development in marketing is the application of brain imaging to measure the direct implicit response of the brain toward products, brands and advertisements (Plassmann, Venkatraman, Huettel, & Yoon, 2015). Electroencephalography (EEG) is the most popular of these methods in neuromarketing practice, particularly in ad testing (Smidts et al., 2014), because of its relatively low costs and high temporal resolution compared to, for example, functional magnetic resonance imaging (fMRI) (Ariely & Berns, 2010).

EEG is a non-invasive method to record brain activity by means of measuring voltage changes at the scalp. The human brain encompasses tens of billions of brain cells, or neurons, that communicate with each other, and this communication can be described electrically (Stahl, 2008). By placing electrodes at the scalp, the summed activity of synchronously active populations of neurons (at the surface of the brain) are measured as variations in voltage at the scalp (Rugg & Coles, 1995). Decades of research have taught us that oscillations in the EEG signal in certain frequency ranges can be associated with specific psychological processes in the brain (e.g. Basar, Basar-Eroglu, Karakas, & Schurmann, 1999).

One of these frequency ranges, or frequency bands, in the EEG signal is the alpha frequency band which is defined as oscillations between 7-12 Hz. Desynchronisation of the alpha band or 'alpha suppression' has been observed since the first application of EEG by Hans Berger (1929) in response to opening the eyes. EEG studies investigating arousal showed that the processing of emotional arousing stimuli is related to alpha suppression (e.g., DeCesarei & Codispoti, 2011; Simons, Detenber, Cuthbert, Schwartz, & Reiss, 2003, but see Aftanas, Varlamov, Pavlov, Makhnev, & Reva, 2002; Uusberg, Uibo, Kreegipuu, & Allik, 2013). The generally adopted procedure in experimental studies investigating emotional arousal presents participants with stimuli from the International Affective Picture System (IAPS, Lang, Bradley, & Cuthbert, 2008) (e.g., Aftanas, Varlamov, Pavlov, Makhnev, & Reva, 2002; DeCesarei & Codispoti, 2011; Keil et al., 2001; Muller, Keil, Gruber, & Elbert, 1999; Uusberg, Uibo, Kreegipuu, & Allik, 2013). These standardized and validated photographs depict various scenes, objects and people, and have been shown to evoke emotional states varying along the two dimensions of valence and arousal. Viewing arousing stimuli enhances cortical excitation compared to viewing neutral or less arousing stimuli, and thereby reduces alpha activity (DeCesarei & Codispoti, 2011).

Other psychophysiological techniques that have been applied in an advertising context to measure consumers' primary arousal response to advertising stimuli, consist of measuring the pupillary response (i.e., the changes in the size of an individual's pupil), electrodermal activity (EDA) of the human body (with 'emotional sweating' influencing the ability of the skin to conduct electrical current), and heart-rate (e.g., a change in the interval between consecutive heart beats). The responses of these measures are however, rather sluggish compared to EEG (see Poels & DeWitte, 2006; Wang & Minor, 2008 for a review on the strengths and weaknesses of the different techniques in advertising and marketing respectively). Particularly in the context of dynamic stimuli such as commercials, applying EEG would be more advantageous in order to assess (fast paced, short lived) responses toward advertisements, compared to the other techniques with a more delayed response (in the range of milliseconds compared to (tens of) seconds respectively).

Neuromarketing and the current study

Typically, in neuromarketing practice, brain activity in certain frequency bands is measured, and then specific psychological processes are inferred from this activity. As discussed previously, suppression of activity in the alpha frequency band (7 – 12 Hz) has been observed when arousal is experimentally increased for example. For this reason, in neuromarketing practice, low measured alpha activity is usually interpreted as high arousal.

Whereas measuring this type of brain activity and inferring a specific psychological process from it is simple and straightforward, it is also problematic because it involves reverse inference (Poldrack, 2006). The fact that inducing high arousal leads to observing alpha suppression (forward inference), does not warrant the reverse inference that observing alpha suppression in a particular instance must mean that arousal is high. This logic is only valid if arousal is the only process that would lead to alpha suppression, which is not the case since also attention, memory demands and general alertness suppress alpha oscillations (Klimesch, 2012). Furthermore, also other frequency bands have been related to arousal, for example increased gamma band activity (30-65 Hz) has been observed in response to emotional arousing pictures compared to neutral pictures (Keil et al., 2001; Muller et al., 1999). This so-called reverse inference problem is thus a common issue for most of the techniques used in neuromarketing practice today.

In the current study, we therefore first applied a so-called functional localizer task to search for patterns of EEG oscillations that we could objectively and independently establish as being associated with arousal. We adopted the same procedure that is generally adopted in other studies investigating emotional arousal, hence we presented the participants with IAPS stimuli. The oscillatory EEG activity evoked by the arousal dimension of these pictures will be used as measure of arousal.

In addition, although EEG can be fruitfully applied to measure the implicit arousal response to advertisements, it is crucial to show that the inferred psychological process is actually relevant for advertising. We therefore investigated whether arousal evoked in response to the ads, as measured neurally in a relatively small group of individuals, is associated with measures of advertising effectiveness in the population at large. Even though EEG is currently the main method applied in neuromarketing practice, evidence on the relationship between EEG oscillations and advertising effectiveness is limited. In six recent studies, EEG-based measures have been shown to be related to market-level success (Barnett & Cerf, 2017; Boksem & Smidts, 2014; Christoforou, Papadopoulos, Constantinidou, & Theodorou, 2017; Dmochowski et al., 2014; Guixeres et al., 2017; Venkatraman et al., 2015). The current study would add to this previous work by showing the specific role of arousal in response to advertisements, as inferred from an EEG measure estimated via a separate task in one and the same study, in relation to advertising effectiveness in an independent larger sample of the population.

Advertising effectiveness can be defined in multiple ways, since several different communication objectives exist for advertising. A first communication objective is increasing awareness in consumers, in order to be able to rely upon memory, either recall or recognition, before or at the moment of the purchase decision (Percy & Rossiter, 1992). Research on personal theories of practitioners in advertising revealed that professionals in the field see capturing consumers' attention as their main goal in order to break through 'the clutter' of other advertisements and daily life (Kover, 1995; Nyilasy, Canniford, & Kreshel, 2013).

A second objective of advertising is generating a favorable attitude toward the ad because of its influence on brand attitude (Brown & Stayman 1992; MacKenzie, Lutz, & Belch, 1986) and subsequent behavior (Ajzen & Fishbein, 1977). The third and ultimate communication objective is at the level of consumer behavior, with the consumer considering purchasing the advertised product (Percy & Rossiter, 1992). In Study 1, we obtained ratings from a large sample of the population on notability and attitude toward print advertisements as measures of advertising effectiveness covering the first two communication objectives (awareness and attitude toward the ad). In the second study, in which we used TV commercials, we obtained a choice measure of advertising effectiveness from a large sample of the population, corresponding to the third communication objective.

In sum, in the present study we first applied a localizer task in order to functionally localize arousal-related patterns of oscillatory EEG activation. After having estimated how arousal is represented in the brain via a separate task, we measured arousal evoked by both static print advertisements (Study 1) and dynamic TV commercials (Study 2), and related those neural measures of arousal to measures of ad effectiveness in the population, consisting of self-reported ratings of notability and attitude toward the ad, (Study 1), and actual click-through-rate to the vendor's website in response to the commercial in Study 2.

STUDY 1: PRINT ADVERTISEMENTS

Methods

Participants

The sample for the EEG study consisted of thirty-one students (16 female, age range = 19-27 years, M = 22.3, SD = 2.5) recruited from the university population. They all had normal or corrected-to-normal vision and had no history of neurological illness. In return for their participation they received 30 euro.

The population sample consisted of an external consumer panel of 1260 participants of the same nationality as that of the EEG participants (650 female, age range = 18 − 88 years, M = 54.1, SD = 12.4) recruited by a market research company. Their educational background covered the whole range from low (23%), to intermediate (35%), to high (42%).

Tasks

EEG functional localizer task. In this task, we showed our 31 participants 100 standardized and validated photographs from the International Affective Picture System (IAPS, Lang, Bradley, & Cuthbert, 2008) while we recorded their EEG, in order to functionally localize arousal-related patterns of activity. We selected photographs at the low and high ends of the arousal dimension, and at the negative and positive ends of the valence dimension in order to keep a balanced design and to control for valence of the stimuli. This results in four picture categories, each containing 25 photos: positive valence/ low arousal, negative valence/ low arousal, positive valence/ high arousal, and negative valence/ high arousal. Averaging the mean normative ratings on arousal and valence of all our selected pictures results in the following mean ratings per picture category (on a scale from one to nine): positive valence = 7.32/ low arousal = 3.56, negative valence =3.56/ low arousal = 3.61, positive valence = 7.12/ high arousal = 6.52, negative valence = 2.73/ high arousal = 6.75 (see Appendix 3.A). Stimuli were presented for 3000 ms, with an interstimulus interval of 1500 ms. We instructed participants to empathize with the situation depicted and try to imagine being in that situation.

Print advertisements. The stimuli in this task consisted of 150 real print advertisements from the United States (US) from the Starch database (GfK MRI USA) and Ads of the World, in order to guarantee that they were new to our

European participants and thus to control for familiarity of the ads. The ads pertained to five product categories: cars, gadgets, food, beauty, and fashion. Only women viewed the beauty and fashion advertisements, and only men the cars and gadgets advertisements. Both women and men viewed the food advertisements. During the EEG study, the advertisements were presented for 5500 ms with an interstimulus interval of 1500 ms. Brain activity was recorded while participants viewed the advertisements.

Participants from the population sample rated a random subset of 10 ads on two measures of ad effectiveness as part of an online survey: *Notability of the ad*, and *Attitude toward the ad* (nine items averaged), in order to capture effectiveness of the advertisement in the population at large (for more details on the items, see Appendix 3.B). We created one measure of *Attitude toward the ad* out of the nine items, because the items mutually correlated significantly at the level of $p < .01$ (ranging from $r = .59$ to $r = .93$), Cronbach's α was .96, and a principal component analysis (PCA) returned only one component explaining 84 % of the variance. All items were Z-transformed at the advertisement level. The descriptives of the raw scores at the advertisement level indicated that the 150 ads differed substantially both in Notability and Attitude (see Table 3.1). Since the population sample is more diverse in age and level of education than the EEG sample, we checked whether a subsample of the population sample with similar characteristics as the EEG sample (relatively young and highly educated) differed in their ratings of the print advertisements from the overall sample. Ratings appeared to be robust to differences in age and education (see Appendix 3.C).

Table 3.1. *Descriptives of raw scores on Notability and Attitude toward the ad at the advertisement level (5-point scale, with 1 = Strongly disagree and 5 = Strongly agree)*

	M	*SD*	Min.	Max.
Notability	3.51	0.32	2.77	4.11
Attitude	2.90	0.27	2.22	3.52

Statistical analyses

EEG sample data. The first steps of the analysis are concerning the functional localizer task. The EEG data obtained during viewing of the IAPS pictures was first transformed to the frequency domain with a fast Fourier transform (i.e., we decomposed the signal into its components of frequencies ranging from 1 to 128 Hz; see Appendix 3.D for details on EEG recording and analysis).

We then estimated how (at which frequencies in the signal) and where (at which electrodes on the 64-channel cap) the difference between low and high arousal is most pronounced across participants. The purpose of this first step in the analysis is thus essentially data reduction. To achieve this, we tested on the first (participant) level, whether the affective category that the presented picture belonged to could be used to predict the EEG data using a regression. To this end, we first log transformed the EEG data for normalization purposes. We then regressed the EEG data onto three independent variables (IV's): arousal, valence, and the interaction of valence and arousal, respectively. Even though our main interest is arousal, for completeness, we also checked the differentiation in brain activity for valence, and for the interaction between valence and arousal. Per participant, the regressions were performed at each electrode and each frequency. Then, at the second level (i.e., the group level), the resulting regression coefficients (i.e., one for each electrode-frequency combination) were used to test the overall arousal effect across participants. Here, we searched for a cluster of frequencies and electrodes where the regression coefficients consistently deviated from zero across participants using (one-sample) t-tests (in the form of a non-parametric cluster-based permutation test). That is, we searched for activity that represents a difference between low and high arousal states in a similar fashion across participants (Step A in *Figure 3.1*, see Appendix 3.E for more details on statistical analysis, Step A). We obtained a cluster of frequency-electrode combinations at which the EEG activity evoked by high and low arousal images differed significantly, in the alpha frequency band (7 - 12 Hz) on central electrode locations (see results section). Henceforth, we will refer to this cluster of frequency-electrode combinations where we find significant effects for low versus high arousal as the arousal Frequency by Region of Interest (F-ROI). No such robust effects of valence (main effect and interaction with arousal) were observed so the valence dimension will not be discussed further.

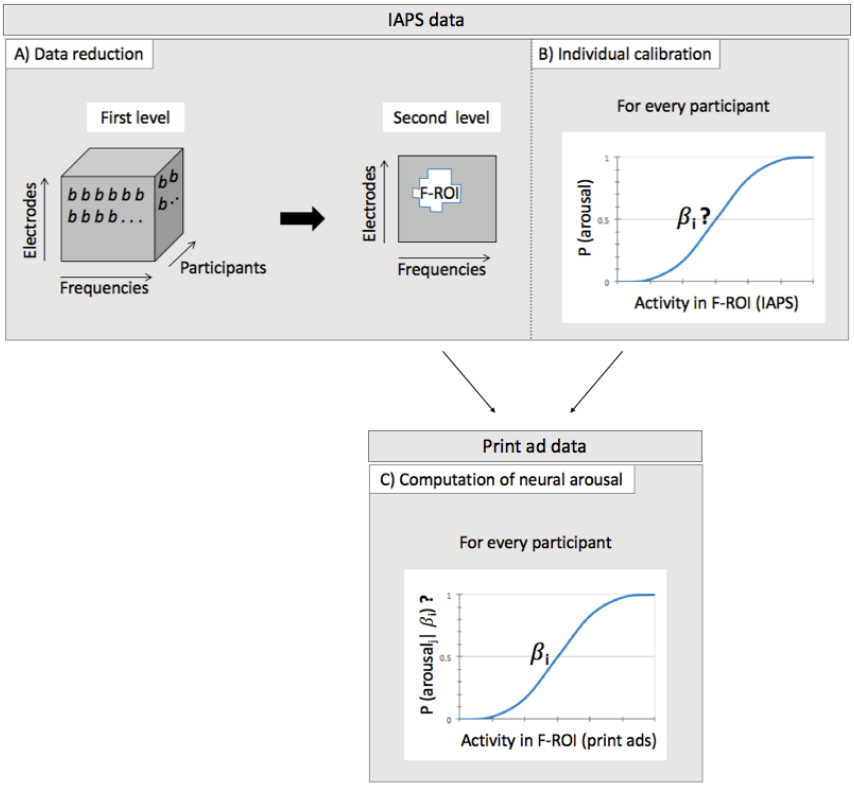


Figure 3.1. Schematic representation of the analysis. Step A and step B are performed on the IAPS data, Step C on the print ad data. A) Data reduction by searching for how and where in the brain, the difference between low and high arousal is consistently represented in the EEG data across participants. This step results in an arousal Frequency by Region of Interest (F-ROI). B) Individual calibration of the data as defined in Step A, by searching for each participant for the optimal relation between the participants' specific EEG response on the one hand (in the F-ROI) and the IAPS pictures on the other hand (DV = low vs high arousal). This step results in participant specific coefficients (indicated by β_i) that link EEG activity and arousal level. C) Computing for each participant i and print ad_j, the probability that a print advertisement evokes neural arousal by selecting activity from the arousal F-ROI that is obtained while viewing the print ads (and given β_i).

After having defined how arousal is represented in the brain across participants, we conduct individual calibration on this data to take heterogeneity of the (brain) response into account (Step B in *Figure 3.1*, see also Appendix 3.E, Step B for results on classification accuracy). Although we found a similar pattern of brain activity in response to the low versus high arousal pictures across participants in Step A, this does not necessarily mean that the specific/exact EEG response to the low and high arousal pictures is the same across participants. Differences in properties of the skull, in the orientation of the neuronal sources and coherences between the sources, in addition to differences between responses at the neuronal source, could result in heterogeneity of EEG responses across participants (Teplan, 2002). We conducted the individual calibration by performing a logistic regression for every participant. We took the (Z-transformed) EEG data from the arousal F-ROI and averaged across the selected frequencies and electrodes, to serve as independent variable. The IAPS picture category served as the dependent variable (DV: low versus high arousal). The participant specific coefficients that resulted from the logistic regressions can be used in the next step of the analysis on the EEG data obtained while viewing the print ads, in order to scale the EEG responses to the ads, to the standardized and validated levels of arousal of the IAPS pictures.

In the next step of the analysis pertaining to the print advertisements, the EEG data that was recorded while viewing the print ads was also first transformed to the frequency domain. We then selected EEG data from the arousal F-ROI as determined in the first step of the analysis, but measured while viewing the print ads, for every participant. In addition, we scaled this activity in response to the ads for every participant, to the arousal level of the standardized and validated set of IAPS pictures. We did this by using a logistic function. For every participant, the selected (Z-transformed) activity from the arousal F-ROI was averaged across the electrodes and frequencies to serve as a predictor. We then used the participant specific coefficient that was computed in the second step of the analysis and a logistic function, to estimate the probabilities that the advertisements evoked arousal for that participant, based on the EEG activity in the F-ROI (Step C in *Figure 3.1*).

Finally, per ad, the estimated arousal probabilities were averaged across participants to arrive at a neural arousal score for each print ad, thus making the advertisement our focal unit of analysis.

Importantly, it should be noted that the individual calibration was performed to take heterogeneity in the EEG response to a specific print ad across multiple participants into account. That is, the same EEG response to a specific print ad, will indicate more arousal for this ad if it originates from someone who reacts less to the high arousing IAPS pictures, than when it originates from someone who reacts more to these same high arousing IAPS pictures. When someone is less easily aroused in general, and thus responds 'with less arousal' to high arousing pictures, without individual calibration, this would lead to rather flat arousal responses in reaction to all print ads. Since our focal unit of analysis is the advertisement, we believe it therefore makes sense to scale the individual responses to these ads, to the validated set of IAPS pictures. Not conducting this individual calibration would result in less clean (more crude) relationships between the EEG activity and the arousal process that it should represent across individuals (see Appendix 3.F for results of analysis without individual calibration; effects are in the same direction but weaker).

Focal analyses on advertisement level. To test whether the neural arousal scores of the print advertisements are related to ad effectiveness in the population at large, we computed the correlation between the neural arousal score of the ads, and averaged ratings of the ads by the population sample on Notability and Attitude.

To control for other features of the ads, we performed hierarchical regressions. We controlled for low-level visual features (i.e., luminance and contrast of the ad), brand familiarity, and dummies for product category in the first block, and tested the significance of the neural arousal scores in the second block. If one of the ad features other than the neural arousal scores appeared to be significant as a predictor, we also tested the interaction term consisting of that ad feature and the neural arousal scores in the third block. Using the 'imread' function from the Image Processing Toolbox in Matlab, we extracted luminance (i.e., the average pixel intensity) and contrast (i.e., the standard deviation of pixel intensity) for each of the print ads.

Although we selected US advertisements because they were new to our participants, several brands in the advertisements may have been familiar to the participants to varying extents. In order to control for the effect of brand familiarity, we collected data on 'familiarity with the brand' with the brand logo presented and scale ranging from (1) *Very unfamiliar*, to (7) *Very familiar*, from a separate sample of 46 students, but with highly similar characteristics to the students in our EEG study (23 female, age range = 18-30 years, M = 21.0, SD = 2.6). Ratings were averaged across participants in order to arrive at one score per ad on brand familiarity (M = 4.6, SD = 2.0, range = 1.2–6.9).

Results

Functional localizer

In the data reduction step of the functional localizer task, the analysis (Step A in *Figure 3.1*) reveals a significant difference between low and high arousal IAPS pictures that was consistently represented in the EEG data across participants ($p <$ .01 FWE). The effect was widespread but mainly present at central sites of the head, and in the alpha frequency band (7-12 Hz, on electrodes F1, F3, F5, FCz, FC1, Cz, C2, C4, T8, CP2, CP3, CP4, CP6, Pz, P2, P5, P7, and POz; see *Figure 3.2*). Activity in the alpha frequency band was lower in the high arousal condition than in the low arousal condition. This cluster of activity will constitute the arousal F-ROI.

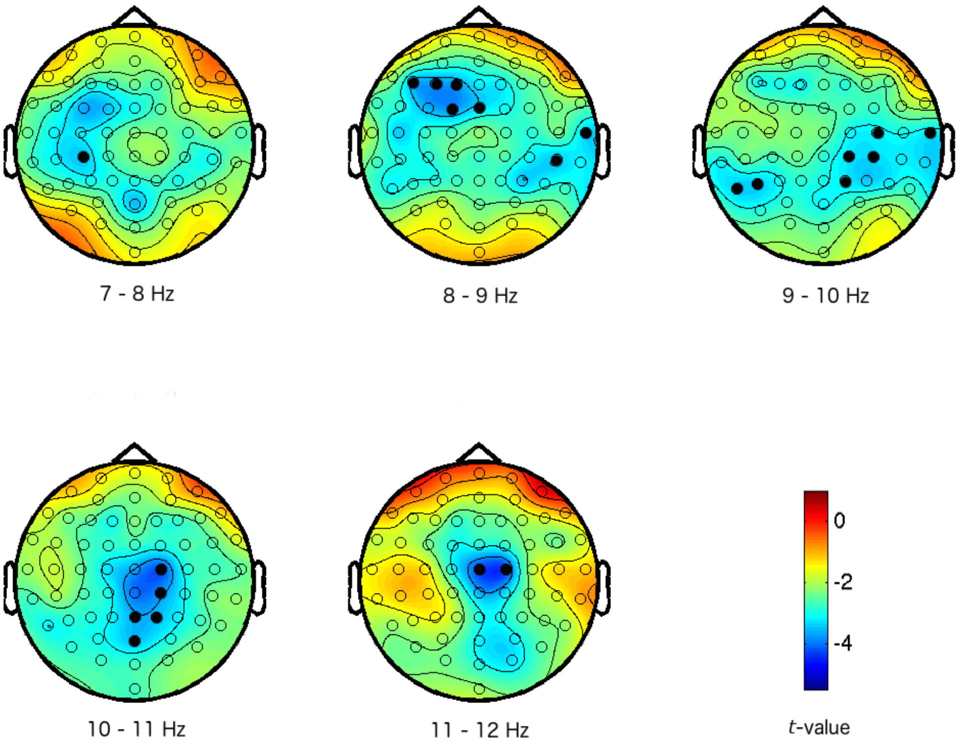


Figure 3.2. Maps of the difference between high and low arousal IAPS pictures in activity in the alpha frequency band (7 - 12 Hz). Notes: The colors represent t-values. The different scalp maps show the contrast (expressed in t-values) between high and low arousal for the specific frequencies that are mentioned below the maps, and across the head for the 64 electrodes. Electrodes with activations above the threshold are marked in black (i.e., electrodes that are part of the arousal effect/ cluster/ F-ROI: preset cluster inclusion for data points $p < .005$, cluster significance $p < .01$ FWE).

Relationship between neural arousal and population sample evaluation

EEG activity selected from the arousal F-ROI that is described above, but obtained during viewing of the print ads was then used, together with the participant specific coefficient obtained in step B to predict the probability that each specific ad evoked arousal (Step C in *Figure 3.1*). The resulting neural arousal scores of the 150 print advertisements were then correlated with the population-wide ad effectiveness. We observed a significant correlation with ad effectiveness as evaluated by the population sample: The higher the advertisements scored on neural arousal, the higher they were rated by the population sample on Notability ($r = .20$, $p < .05$), but the lower they scored on Attitude toward the ad ($r = -.26$, $p < .005$, both p-values corrected for multiple comparisons) (see *Figure 3.3*).

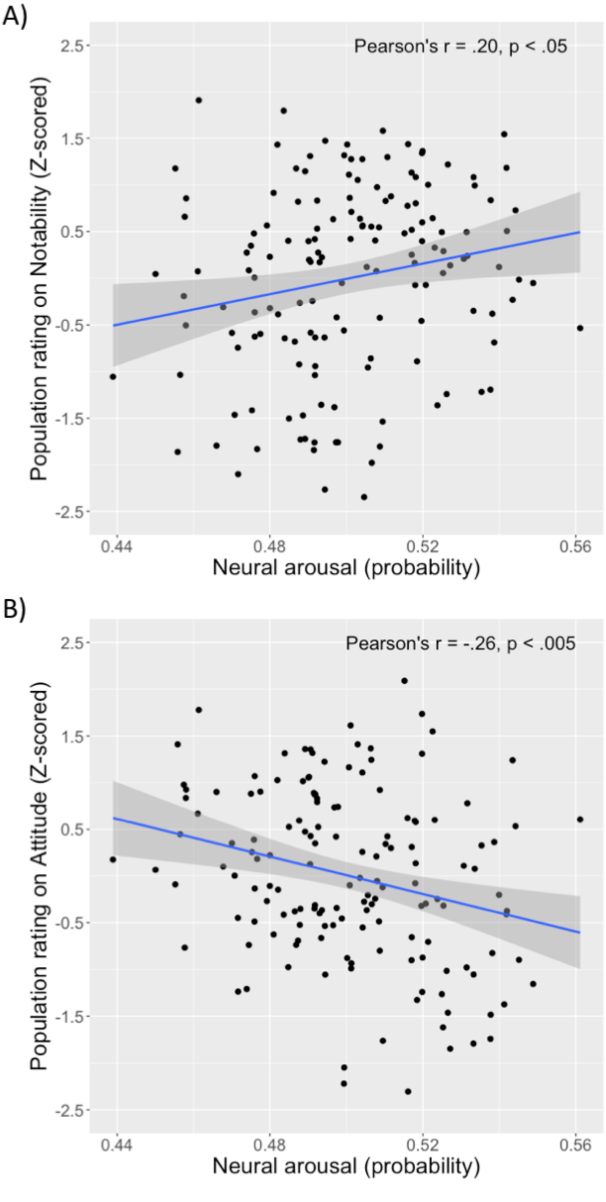


Figure 3.3. Scatterplots of the relationship between neural arousal score of the 150 print advertisements and average population sample ratings on A) Notability of the ad and B) Attitude toward the ad.

When controlling for other ad features, these direct relationships between neural arousal and its respective ad effectiveness measures remain. Table 3.2 shows that neural arousal is positively associated with notability of the ad (standardized regression coefficient β = .19, t = 2.31, p < .05) when controlling for product category, brand familiarity and luminance and contrast. Adding neural arousal to the model in step 2 significantly increased the fit of the model (R^2change = .03, F-change = 5.33, p < .05). The interaction term between the neural arousal score and contrast was not significant (p > .10) and therefore did not enter the model. Similarly, Table 3.2 shows that adding neural arousal to the model predicting attitude toward the ad, significantly increased the fit of the model (R^2change = .06, F-change = 11.27, p < .005) with neural arousal being negatively associated with attitude toward the ad (β = -.25, t = -3.36, p < .005). The interaction terms between the neural arousal score, dummy product category 1, dummy product category 2, and brand familiarity were all non-significant (p > .10) and did not enter the model.

Table 3.2. *Relationships Between Population Sample Ratings on Notability of the ad, Attitude toward the ad, and Features of the Ad*

	Notability			Attitude		
	β		t	β		t
Product category cars vs food	-.16		-1.51	-.30	***	-2.97
Product category gadgets vs food	-.17		-1.58	-.23	*	-2.32
Product category beauty vs food	-.17		-1.66	-.08		-0.85
Product category fashion vs food	-.11		-1.03	-.12		-1.29
Brand familiarity	.12		1.30	.42	***	5.10
Luminance	-.09		-1.06	-.03		-0.45
Contrast	-.24	*	-2.93	-.04		-0.51
Neural arousal	.19	*	2.31	-.25	***	-3.36
R^2	.13			.23		
R^2 change neural arousal	.03			.06		
F change neural arousal	5.33	*		11.27	***	

Notes: *p < .05, **p < .01, ***p < .005. Beta is the standardized regression coefficient.

Cross-validation analyses confirmed that neural arousal also contributed significantly to the models in predicting ratings of a holdout sample, supporting the results that neural arousal positively predicts notability of the ad, but negatively predicts attitude towards the ad, controlling for product category, brand familiarity and luminance and contrast.

Because of the null result for valence related activity (and the negative relationship between neural arousal and attitude), we wanted to rule out the possibility that our arousal measure is biased towards negative valence. Conducting an independent samples t-test on arousal activity in response to the IAPS pictures (i.e., activity that is averaged across the arousal F-ROI), for positively versus negatively valenced IAPS pictures, indicates that these pictures on average do not differ in arousal for all participants (no significant differences for any of the participants; range of *t*-values = -1.88 to 1.92, one sample *t*-test on *t*-values is n.s. $M = 0.06$, $t = 0.35$). This shows that there is no (negative) valence bias in our arousal measure.

Follow-up analysis on negative relationship between arousal and attitude

In order to shed light on why we find a negative relationship between neural arousal evoked by the print ads and ratings on attitude by the population at large, we followed up on the results with an additional exploratory analysis. Inspecting the scatterplot illustrating this relationship (*Figure 3.3.B*) suggests that the ads in the bottom right corner contribute heavily to the negative relationship between our measure of arousal and ratings on attitude: More ads are present in this corner compared to in the other corners, and they extend more to the limits of the two variables. Examining these ads for a common denominator reveals that many of those ads deal with the human body in a creative, uncommon manner or contain unusual content of a physical nature, which may not be appreciated by some of our participants and may therefore not be perceived as positive. An example of such an ad depicts one arm with hand, with a second hand pulling of the skin of the first hand, as if an invisible glove is put on, advertising for a hand sanitizer. In another such ad, the face of a person is presented with a fist appearing from the neck, punching the person in the face, referring to the (wasabi) taste of the chips. We hypothesize that the negative relationship between arousal and attitude in our study may exist because of these creative ads that are perceived as rather odd or confusing, instead of being perceived as positively creative. Searching the whole stimulus set for print ads that tick this box ($k = 22$), and excluding these from the analysis results in a non-significant correlation between arousal and attitude ($k = 128$) of -.13 (p =.15). Thus, these (mostly negatively evaluated) creative ads indeed contributed to the negative relationship between arousal and attitude, which is perhaps due to the nature of the creativity used here (with the use of the human body in an uncommon fashion).

Conclusion and Follow-Up

Our localizer task for which we used IAPS pictures to evoke an arousal response in the EEG revealed a neural correlate in the form of a decrease in alpha activity at specific electrodes at central sites of the head. We then used this arousal F-ROI to compute a neural arousal score for a variety of print advertisements. In turn, this neural arousal score contributed to explaining effectiveness of the advertisements in the population at large, even after controlling for other ad features such as brand familiarity and luminance. The neural based arousal score of print advertisements from our neural focus group correlated significantly with population sample ratings of the advertisements: positively with Notability of the ad, and negatively with Attitude toward the ad. These results suggest that although arousal-evoking ads stand out, they do not necessarily generate positive attitudes. A possible explanation for the surprising negative relationship is that a puzzling and difficult to understand creative execution of ads, may not be perceived positively.

An open question that remains is how neural arousal in response to the ads relates to an actual behavioral choice measure of ad effectiveness. Whereas consumers may not like arousing ads, such ads may still drive their choice. Liking of the advertisement is not a prerequisite for all types of advertisements to be acted upon (Percy & Rossiter, 1992). Given that the ultimate marketing communication objective is at the level of behavior, with the consumer considering purchasing the advertised product, we obtained a behavioral measure of advertising effectiveness from a large sample of the population in a second study. Additionally, in order to examine generalizability our findings, we replaced the static print advertisement stimuli by dynamic TV commercials in Study 2.

In Study 1, we first established - via a separate task- that measuring EEG activity in the alpha frequency band is (inversely) related to the experienced level of arousal across participants. Secondly, we found that this neural measure of arousal predicted measures of advertising effectiveness in the population at large. However, collecting EEG data for a functional localizer task and calibrating for individual differences is not something practitioners would be inclined to do every time they want to measure arousal as this would be too costly and time consuming. Thus, in the second study, we took a less refined approach that is close(r) to a practitioner's approach, without individual calibration: In order to measure arousal for every participant, we simply selected the EEG activity that was representative of arousal in a similar fashion across participants (as revealed in Study 1, i.e., activity in the F-ROI; 7 - 12 Hz at central electrode sites) from the EEG data that was measured while participants viewed TV commercials, and subsequently related this brain activity to population level success of the commercials. Additionally, in order to explore the possible contribution of an ineffective creative execution of an ad to neural arousal (and subsequently to ad effectiveness), a panel of experts assessed the commercials on how confusing they were. Although the TV commercials are not necessarily weird or creative in a similar manner as the print ads in Study 1 (i.e., no strange or extreme content of a physical nature), confusion can result from a creatively ineffective execution (e.g., a complex storyline).

STUDY 2: DYNAMIC COMMERCIALS

In Study 2 we assessed effectiveness of TV commercials in the population at large by measuring click-through rate to the product website of the advertised product. We measured neural arousal evoked by the TV commercial in a small group of participants using EEG and investigated its association with click-through rate (CTR) in a larger independent population sample.

Methods

Participants

Forty students recruited from the university population (21 female, age range = 17-24 years, *M* = 20.3, *SD* = 1.6) participated in this EEG study. They received 25 euro in return for their participation. Participants all had normal or corrected-to normal vision and no history of neurological illness.

The population sample consisted of 1239 participants (615 female, age range = 25 – 55 years, *M* = 41.8, *SD* = 8.8). They were randomly selected from a consumer panel to fill out an online survey that was hosted by a market research company. The expert panel consisted of 9 independent professionals who had substantial knowledge and working experience (all over 7 years) in the field of advertising, marketing or communication. For further details on this data collection, see Couwenberg et al. (2017).

Stimuli

The stimulus set consisted of eleven television commercials advertising the same brand and product (i.e., a well-known muscle and joint gel). The commercials contained a similar voice-over, but differed in executional style and were developed in the context of a competition between advertising agencies. The duration of all commercials was 20 seconds. To test for uniformity across commercials concerning low-level visual features, we used the 'imread' function from the Image Processing Toolbox in Matlab to extract luminance (i.e., average pixel intensity per frame), contrast (i.e., the standard deviation of pixel intensity per frame) and the amount of movement and cuts (i.e., the pixel-by-pixel correlation relative to the previous frame; the cross-frame correlation) for each commercial separately. All measures were computed on a frame-by-frame basis and then averaged over the commercial. One-sample Kolmogorov-Smirnov tests indicated uniformity across the commercials for luminance and contrast (both $p > .05$). Although cross-frame correlations were not uniformly distributed across commercials ($p < .05$), they did not correlate significantly with the ad effectiveness measure. These results indicate that any findings related to our ad effectiveness measure cannot be attributed to differences in low-level visual features.

Tasks

In the EEG study, the participants passively viewed all 11 commercials four times in total. In the first block, they viewed the commercials twice, in random order. Thereafter participants executed another task not related to this study. In the final block, they again viewed the 11 commercials twice in random order.

The participants from the population sample passively viewed one of the 11 randomly selected commercials in an online survey. We provided participants with the option to click through to the product website after presentation of the commercial, which is our measure of interest (i.e., 0 or 1), or to end the survey. Participants were informed of a discount for the product before they decided to click though to the website, and on the website, they could find more information about the product and potentially purchase it. The product was offered at a discount upon request by the company in order to stimulate purchase behavior on the product website. Although the discount could have sensitized the participants' click-through behavior, it would have done so equally for all commercials.

The advertising professionals from the expert panel were asked, in addition to other questions that are not of relevance to the current study, to judge how confusing the advertisements were on a scale from 1 to 4.

Statistical analyses

The EEG data obtained during passive viewing of the commercials was first transformed to the frequency domain (see Appendix 3.G for EEG recording and analysis of Study 2), and then Z-transformed per electrode, frequency, repetition of the commercial, and per participant.

As mentioned above, we estimated via a separate task in Study 1 that measuring EEG activity in the alpha frequency band is (inversely) related to the experienced level of arousal across participants. For every participant, we select this EEG activity that is representative of arousal from the data that was measured while viewing the commercials (i.e, activity from the arousal F-ROI; 7 - 12 Hz at central electrode sites), and averaged the activity across the selected electrodes and frequencies. In addition, we averaged this arousal-evoked EEG response across the four repetitions of the commercials and across participants in order for the commercial to become the final unit of analysis. This results in one neural arousal score per commercial that is expressed in (Z-transformed) alpha activity. In this study, however, we flipped the sign of the Z-scored EEG data because of the inverse relation between arousal and alpha activity as established in Study 1 (i.e., low alpha is assumed to indicate high arousal), and to avoid confusion about the direction of the relationship across studies. Thus, when we discuss arousal in Study 2, it refers to the inverse of alpha activity, instead of probabilities that the ad evokes arousal as in Study 1.

The effectiveness measure click-through rate (CTR) was computed as the percentage of participants from the population sample clicking through to the product website per commercial, and ranged from 6.19% for the 'worst' commercial to 16.35% for the 'best' commercial (M = 9.67%, SD = 2.88% across commercials). In order to test whether the neural measure of arousal evoked by the commercials was related to the effectiveness of the commercials in the population at large as measured by CTR, we computed Spearman's correlation coefficient, since the data consists of only eleven data points (for the 11 commercials) that are not normally distributed.

Results

Computing the correlation coefficient between CTR and the neural measure of arousal revealed a marginal significant negative relation (r = -.55, p = .08) as visualized in *Figure 3.4*.

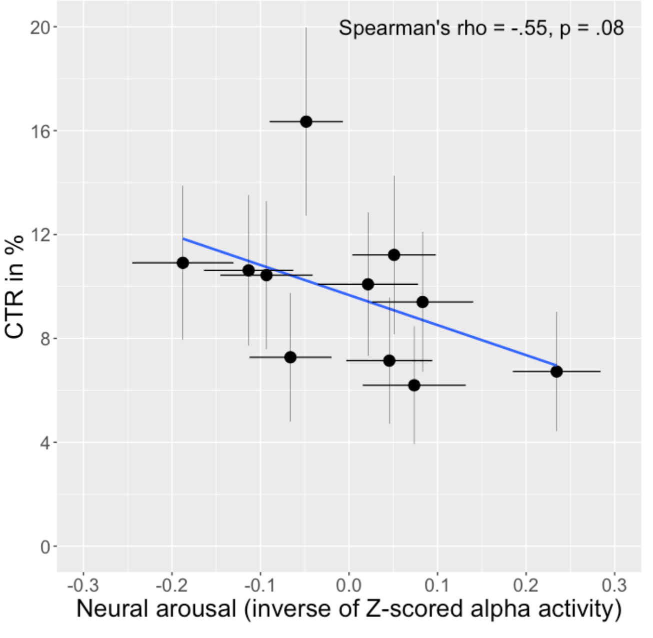


Figure 3.4. Scatterplot of the relationship between neural arousal of the 11 commercials and CTR by the population sample. Lines indicate standard errors around the mean of scores.

This correlation indicates a negative relationship between the neural measure of arousal evoked by the commercials and effectiveness of the commercials in the population at large as measured by CTR.

Additionally, in order to explore to what extent confusion affects the relationship between neural arousal and CTR we computed the correlation between confusion and arousal, and between confusion and CTR. There was a marginally significant relationship between confusion and neural arousal that was positive (Spearman's rho = .56, p = .07) and a marginally significant relationship between confusion and CTR that was negative (Spearman's rho = -.59, p = .06). Computing the partial correlation coefficient indicated that the relationship between neural arousal and CTR was indeed reduced by controlling for confusion (Spearman's rho = -.33, p = .35, compared to Spearman's rho = -.55, p = .08 before). Note that these correlation coefficients indicate large effects (> .50), despite the fact that p-values are not below .05, which is due to the small n (only 11 commercials).

Discussion

Main findings and contributions

We presented two studies in which we showed the role of arousal, as measured using EEG, in predicting advertising effectiveness. The aim of the present studies was twofold. The first aim was to measure arousal in response to advertisements using EEG, by first estimating how arousal is represented in the brain via a separate task. Our second aim was to show that this neural measure of arousal was associated with actual marketing effectiveness. We specifically investigated whether the evoked arousal in a relatively small group of individuals was predictive of both the notability of the advertisements, as well as consumers' attitudes and behavior in response to the ads in the population at large, which are considered the three primary communication objectives for advertising.

Importantly, we first administered a localizer task to functionally localize arousal-related patterns of EEG activity in order to diminish the problem of reverse inference. We found that arousal evoking IAPS pictures prompted a change in power in the alpha band in our participants, allowing us to define an EEG map of arousal. This finding is in agreement with previous literature on the neural correlates of arousal, showing a decrease in alpha activity when comparing high to low arousing stimuli (DeCesarei & Codispoti, 2011; Simons et al., 2003, but see Aftanas et al., 2002; Uusberg et al., 2013).

After defining EEG activity that is representative of arousal using the localizer task, we sampled this EEG response obtained during viewing of advertisements (i.e., from the arousal F-ROI), which allowed us to estimate the level of arousal experienced for each of the print advertisements (Study 1) and dynamic TV commercials (Study 2). In Study 1 we calibrated participants' EEG response to the ads to the standardized and validated levels of arousal of the IAPS pictures. In Study 2 we took a less refined practitioner-like approach by simply selecting this brain activity of which we estimated via a separate task that it should represent arousal across participants in a similar fashion (without individual calibration). We then related those neural measures of evoked arousal to measures of ad effectiveness in the population sample, consisting of self-reported ratings in Study 1 (notability and attitude toward the ad) and a behavioral measure in Study 2 (click-through-rate).

The results showed that the neural measure of evoked arousal in response to the ads was positively associated with notability of the print ads, but that attitude toward the print ads and also click-through-rate in response to viewing the TV commercials were negatively associated with neural measures of evoked arousal in response to the ads. The negative relationships between neural arousal and attitude, as well as between neural arousal and click-through-rate may be explained by creative executions of the ads that seem not to be perceived positively. In Study 1, the exploratory analysis suggested that the use of the human body in an unusual fashion contributed to the negative relationship between neural arousal and attitude. In Study 2 we obtained a measure of the extent to which the TV commercials generated confusion, possibly resulting from an ineffective creative execution and causing the commercials to be perceived as less fluent. We found that confusion was indeed related to both increased arousal as well as lower CTR, possibly suggesting that arousing but non-fluent ads may result in less positive evaluations of those ads.

Managerial and theoretical implications

The results of the present studies showed that the neural measure of arousal evoked in response to the ads as measured using EEG was associated with three communication objectives of advertising, even after controlling for other ad features such as contrast, luminance, product category, and brand familiarity. Using this measure of arousal in response to advertisements (and applied on a moment-to-moment basis for commercials) may provide advertisers with more concrete and actionable insights on the current mental state of the consumer. Equally as important, these results imply that the EEG measure reflecting arousal could provide advertisers with more guidance as to whether the ad design meets a particular communication objective, which may serve as input into decisions to optimize the ad. That is, advertisers need to consider the main objective of their campaign when designing ads since different measures of ad effectiveness may not align and may call for different ad designs. Ads that are neurally arousing may be effective in the sense that they are highly notable and thus enhance awareness, but may not necessarily be effective in terms of being positively evaluated or indeed acted upon.

The first communication objective of raising awareness was reflected by ratings on notability. As expected, the more arousing ads are found to be more notable as rated by the population sample. We conjecture that the downstream consequences of our 'notability measure' are comparable to the ones of 'Starch *noted* scores' as seen in older research (e.g. Lucas, 1960; Spotts, Weinberger, & Parsons, 1997). The Starch score indicates the extent to which an advertisement has been 'noted' by viewers, so both this measure and our measure assess at least initial attention and give an indication of the extent to which advertisements can break through 'the clutter' of other presented material. Whether such scores, however, actually measure recognition and thereby ad memory is debatable (Lucas, 1960; Spotts et al., 1997).

For the second and third communication objective however, we found somewhat surprising results. The more arousing ads are evaluated less positively in Study 1 and are acted less upon in Study 2. In Study 1, the exploratory analysis suggested that the (mostly negatively evaluated) creative ads contributed to the negative relationship between arousal and attitude, which is perhaps due to the nature of the creativity used (with the presentation of the human body in an uncommon fashion). Both this uncommon use of body parts in ads in Study 1, and the additional confusion assessment of the TV commercials in Study 2, tap into processing fluency. Previous literature shows that fluency is perceived positively, with fluent processing being achieved in different ways, from purely perceptually (e.g, symmetry) to semantically (e.g., priming) (Labroo, Dhar, & Schwarz, 2008; Lee & Labroo, 2004; Reber, Schwarz, & Winkielman, 2004; Reber, Winkielman, & Schwarz, 1998). This suggests that non-fluent or difficult to comprehend and confusing stimuli may very well be perceived negatively. In Study 2 we indeed find that the more confusing TV commercials evoked more arousal but were acted less upon. All in all, these exploratory follow-up analyses confirm our tentative hypothesis that an ineffective (unusual, confusing, non-fluent) execution of creativity may even result in a negative relationship between arousal and attitude as well as between arousal and actual behavior. Importantly, future studies are necessary in order to further investigate the specific process underlying the negative relationship between arousal as measured with EEG and ad effectiveness found in the current study, before this relationship can be generalized.

Nevertheless, research on personal theories of practitioners in advertising revealed that professionals in the field see attracting attention as the most important aspect of advertising (Kover, 1995; Nyilasy et al., 2013), and thereby assume that attention grabbing ads will be evaluated positively. Advertising agencies however realize that crafting a creative ad that cuts through the clutter involves taking a risk. In addition, they feel that their clients fear creative advertising and often resist creative advertising solutions (Nyilasy et al., 2013). The results of our studies suggest that the fear of clients may perhaps be justified and thus that advertisers need to have specific communication objectives from their clients. It is important to determine in advance how crucial it is, e.g. given the current status of the brand, to stand out and create brand awareness with the advertisement while taking the risk of not getting it attitudinally right.

Our findings of a negative association between the neural measure of arousal and two out of the three advertising objectives, are not in line with previous studies in marketing which in general show a positive relationship between arousal as measured by self-report, but also as measured with other techniques reflecting autonomic nervous system activity (e.g., electrodermal activity), and ad effectiveness. These measures of effectiveness consisted of attitude toward the ad (Holbrook & Batra, 1987; Olney et al., 1991), attitude toward the brand (Holbrook & Batra, 1987), conative attitude (i.e., purchase intention; Morris et al., 2002), viewing time (Olney et al., 1991), and sales (LaBarbera & Tucciarone, 1995).

One possible explanation for our contrasting findings of a negative relationship, is a difference between stimuli in our study and in other studies. An additional explanation for our findings in contrast to studies using self-reports, is that the neural measure of arousal reflects the 'level of activation' in a different way compared to self-reports of arousal. Self-reports require cognitive processing with conscious retrospective reconstruction of one's feelings. In contrast, our neural measure of arousal consists of direct measurement in real time of power in the alpha frequency band, across the whole duration of the commercial (or across a period of 2.5 seconds for the print ads). Such instantaneous cumulative responses may therefore easily differ from a retrospective account of the experience which is ultimately expressed by self-reports. It is thereby plausible that self-reports of arousal underestimate or largely disregard attentional processes active during evaluation of advertising. The alpha frequency band has also been associated with attention (see review Klimesch 2012), which highlights the challenge to infer a specific psychological process (e.g., arousal) from certain brain activity, even when it is defined using a functional localizer task, and to distinguish the specific process from other ongoing processes (such as attention).

Limitations and future research directions

We acknowledge several limitations that should be considered in the interpretation and generalization of our findings. As mentioned above, the negative relationships between arousal and attitude, and arousal and behavior may be specific to the stimulus sets that were used here that contained some creative executions. It would therefore be valuable to study creativity in more depth, in order to search for patterns in executions of creativity that are perceived as positive versus negative. This knowledge will aid in the prediction of the actual tone of interest toward the ad (i.e., pleasurable or unpleasurable), in addition to the mere intensity of the interest response in the form of arousal. In addition, it would be interesting to further examine if the level or extent of creativity can be measured (on a moment to moment basis) with the neural measure of arousal.

While our measures of advertising effectiveness are important ones, more research is necessary to understand the specific downstream consequences of these measures. For example, even if more notable ads cause better memory of the ad, the memory for the brand or product will depend on many factors such as focus of the ad (on product or brand versus not), consistency between existing knowledge of the brand and the execution of the ad, familiarity of the brand, and possibly an interaction between these factors. In addition, even though click-through rate as a behavioral measure of advertising effectiveness can be regarded as a precursor of purchase behavior, the question remains how all the previous steps (e.g., notability, attitude) are weighted and thereby influence the ultimate purchase of the product.

Finally, although our neural measure successfully predicted effectiveness of the advertisements at the population level, we did not examine if its contribution is above and beyond that of self-reports. Though this was not the aim of the current study, it is important to emphasize because the question remains whether applying neuroimaging is more beneficial in predicting commercial success than other measurements such as questionnaires and to what extent this may be the case (Venkatraman et al., 2015). Nevertheless, as we have shown here, neural measures may reveal important insight into the underlying processes (such as arousal) that may lead to attitudes and behavior towards advertising messages.

Final Conclusion

In summary, the current studies demonstrate how to measure arousal evoked in response to advertisements using EEG. Additionally, we show that this neural measure of evoked arousal as obtained in a relatively small group of individuals and estimated via a separate task in one and the same study may be useful for marketing purposes. While controlling for other ad features, it contributes uniquely to explaining different measures of advertising effectiveness in the population at large that cover three communication objectives of advertising. This evidence is important because of the increasing implementation of small so called 'neural focus groups' in marketing practice, aimed at providing indices of advertising effectiveness at the population level. The findings imply that different levels of alpha activity evoked in response to ads can not automatically assumed to be associated with ad effectiveness in general. Although advertisements that evoke more arousal, or less alpha activity, will likely be more noted, they are not necessarily perceived positively and acted upon. Nevertheless, using this EEG measure and taking the discussion points into consideration potentially allows advertisers to improve on assessing the effectiveness of an ad design at an early stage, and even apply this measure on a moment-to-moment basis. Although further research is needed, because of the relative inexpensiveness of EEG, the benefits of more and concrete improvements on ad designs, and the unique insights into the underlying psychological processes may indeed outweigh the costs.

Appendix

Appendix 3.A Presented IAPS pictures

Positive valence/ high arousal 1650, 5621, 5626, 5629, 8030, 8034, 8080, 8161, 8170, 8178, 8179, 8180, 8185, 8186, 8190, 8191, 8200, 8251, 8300, 8341, 8370, 8400, 8470, 8490, 8501; positive valence/ low arousal 1604, 1610, 1620, 2299, 2304, 2360, 2370, 2387, 2388, 2530, 2540, 5000, 5001, 5010, 5200, 5201, 5220, 5551, 5611, 5760, 5779, 5780, 5811, 5891, 7325; negative valence/ high arousal 1050, 1052, 1120, 1300, 1525, 1931, 1932, 2730, 3500, 3530, 5971, 6230, 6250, 6260, 6300, 6313, 6350, 6510, 6540, 6550, 6560, 8485, 9600, 9800, 9810; negative valence/ low arousal 2221, 2399, 2440, 2490, 2491, 2590, 2722, 2750, 2753, 5120, 5130, 7060, 7224, 7234, 7700, 9000, 9001, 9046, 9090, 9220, 9280, 9290, 9330, 9331, 9360.

Appendix 3.B Panel data ratings of print advertisements

Each of the 1260 participants from the external panel (650 females) rated 10 ads, which results in 12600 ratings in total.

Items panel data (translated into English):

S01 Is notable
S02 Arouses my interest in the brand or product
S03 Stimulates me to respond or go to a store
S04 Is confusing and hard to understand (R)
S05 Makes me feel good
S06 Moves me
S07 Is joyful to watch
S08 Irritates me (R)
S09 Is believable

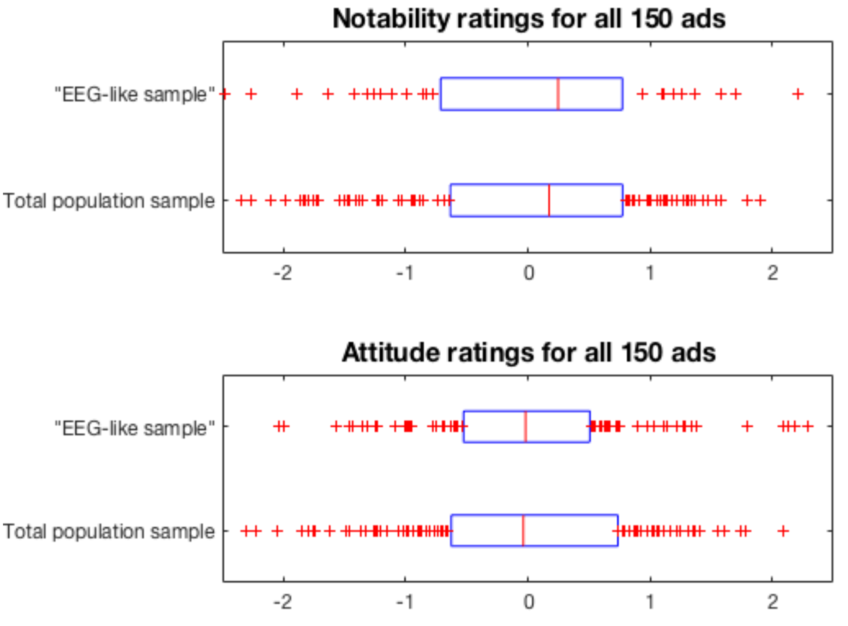
1 = Strongly disagree
5 = Strongly agree

Overall rating:
If you would have to rate the overall impression of the advertisement on a scale from 1 to 10, which score would you give? Scale: (1) Very bad to (10) Excellent.

Performing a PCA on the ten items resulted in two extracted components with eigenvalues over Kaiser's criterion of 1, with only Notability loading on the second component. S02, S03, S04, S05, S06, S07, S08, S09, and the overall score were taken together to form the effectiveness measure of Attitude toward the ad. S01 is the measure of Notability of the ad. The final measures Notability and Attitude do not correlate (rho = -.11, p-value = .19).

Appendix 3.C Similarity of ratings of print ads across different subgroups

Our population sample consists of a random sample from the general population, which can be regarded as a strength of the study. However, the EEG data that is measured in response to the print ads is measured in a more homogeneous sample of university students. In order to establish that our results are not an artefact of the differences between the EEG sample and population sample we checked whether different subgroups in the population sample rated the ads in a similar fashion. Specifically, we inspected the ratings from a subset of the population sample with similar characteristics as the EEG sample with respect to age and education (n=62 out of the 1260, education: at least high school or higher and age: 18-35 years). We label this subgroup of the population the 'EEG-like sample'. *Supplementary Figure 3.1* shows the average ratings from the total population sample and from the EEG-like sample on Notability and on Attitude.



Supplementary Figure 3.1. Averaged ratings of the ads on Notability and Attitude by a subsample of the population with the same characteristics as the EEG sample, and by the total population sample. The central mark represents the median, with the edges of the box representing the 25th and 75th percentiles. The points outside the 25th and 75th percentiles are indicated as 'outliers'.

The distributions of ratings on Notability and Attitude appear to be very similar for the two samples. Independent samples t-tests revealed no differences between ratings from the two different samples on Notability or Attitude ($p = 1.0$ for both measures).

In addition, we computed the correlations between the ratings on notability and attitude from the EEG-like sample and the total sample in order to check whether the same ads are rated as low and high on these measures. The correlations were highly significant ($p < .001$) with Pearson´s rho = .55 for Notability and rho =.49 for Attitude. Visual inspection of ratings per age group and per education type revealed similar patterns for all different subgroups. We therefore conclude that the findings in our study are not artefacts and not caused by differences between the two samples.

Appendix 3.D EEG recording and analysis

The EEG data was acquired using the BioSemi Active Two system with 64 active Ag-AgCl electrodes. Additional flat type electrodes were placed on the right and left mastoid, and in the eye region in order to record eye movements or electro oculograms (EOGs): Electrodes were placed below and above the left eye in line with the pupil to record vertical EOGs, and at the outer canthi of both eyes to record horizontal EOGs. The EEG and EOG signals were sampled at a rate of 512 Hz, and digitally low-pass filtered with a 128 Hz cut-off (3 dB). All preprocessing was done similarly for the two tasks in Brain Vision Analyzer software (BVA; Brain Products). The data was first down-sampled to 256 Hz, then re-referenced to the averaged mastoids, and filtered with a low cutoff filter of 1 Hz and a notch filter of 50 Hz with a slope of 48 dB/octave. We split the continuous data into 100 segments (one for each IAPS picture) for the functional localizer task, and 90 segments (one for each ad) for the print ad task, with segments lasting from stimulus onset until 2.5s after stimulus onset. Then, we applied Gratton and Coles ocular correction to correct for eye movements as implemented in BVA, and a standard artifact detection and rejection procedure in which segments were rejected that contained jumps larger than 30μV/ms, amplitude differences exceeding 150μV/ 200ms, and amplitude differences below 0.5 μV/ 100 ms. Note that only the channels that contained artifacts were deleted within the given segment, and not the entire segment. For all segments, we decomposed the signal into components of frequencies ranging from 1 to 128 Hz using a Fast Fourier Transform (FFT) (BVA, using a 10% hamming window). The resulting spectral data was exported to Matlab (Mathworks).

Appendix 3.E Details statistical analyses (see also *Figure 3.1* in main chapter)

Step A on IAPS data: Data reduction and construction of frequency by region of interest (F-ROI). After preprocessing the EEG data that was obtained while viewing the IAPS pictures in BVA, the data is specified at two dimensions: at frequencies and electrodes. The typical approach for conducting analyses on EEG data is via multiple ANOVA's on mean activity that may (or may not) be defined in a priori windows (Groppe, Urbach, & Kutas, 2011). Instead, we apply a non-parametric method in which we can properly use a large number of univariate tests via a mass-univariate analysis. More specifically, we use a cluster-based permutation test (from the Mass Univariate ERP Toolbox by David Groppe, 2011) to search for how and where in the brain, the difference between low and high arousal is consistently represented in the EEG data across participants. The reported p-values are corrected for multiple comparisons via this permutation test. The result of this analysis is a cluster of frequencies and electrodes where the activity represents a difference between low and high arousal states in a similar fashion across participants. We will therefore refer to this this cluster as the arousal Frequency by Region of Interest (F-ROI).

With a one-sample cluster based permutation test, we can search for a cluster where a given statistic consistently deviates from zero across participants. On the first level or participant level, we regressed the EEG data onto the three independent variables (IV's): valence, arousal, and interaction of valence and arousal (at each electrode, each frequency and each participant). The resulting regression coefficients were used to test three effects (i.e., valence, arousal, and the interaction effect) at the second level (i.e., the group level), and served as input to the one-sample cluster-based permutation test to search for clusters of electrodes and frequencies where the regression coefficients consistently deviate from zero across participants (hence, 3 tests). This permutation test is part of the aforementioned toolbox, and needs as additional input the number of permutations it has to execute, the test-wise p-value threshold for cluster inclusion of the data point, and the desired family-wise error rate (FWER). We set the number of permutations to 5000, the test-wise p-value threshold for cluster inclusion to .005, and the FWER to .05. The test is based on a 'cluster mass' statistic, and the output exists of a cluster of data points (frequencies, electrodes) where an effect is present across participants, with a p-value specific for that cluster and t-values for the specific data points. For an evaluation of the different phases of the cluster-based permutation test itself, we refer to the paper by Groppe et al. (2011) and the toolbox.

In addition to the main effect of arousal on the EEG data, the cluster based permutation test revealed no main effect of valence and no interaction effect. Since the literature on valence often reports asymmetry-effects (e.g., Davidson, 2004), we additionally transformed the data to specifically search for asymmetrical effects where we subtracted for each electrode, data at the electrode at the exact opposite side of the other hemisphere (except for the midline electrodes). We did not find additional significant clusters based on this transformed data. This means that we focus the next steps of the analyses solely on arousal.

Step B on IAPS data: Individual calibration. Here, we search for the optimal relation between the standardized and validated levels of arousal of the IAPS pictures on the one hand, and the participants' specific EEG response on the other hand, as explained in the main chapter.

In order to validate the use of the participant specific coefficients in the individual scaling of the EEG activity in response to the print ads, we checked whether the classification accuracies of the logistic regressions on the IAPS EEG data deviated from chance level across participants. This was indeed the case (M = 55.10%, $t = 5.05$, $p < .00005$). In addition to comparing the classification accuracy of the IAPS pictures to chance level, we applied permutations: we simulated the distribution of accuracies under the null hypothesis that the observations and class labels are independent, for each of the participants. For each of the participants, we sampled 50 observations uniformly at random from the EEG data with replacement, and assigned them the label low arousal. We repeated this step for high arousal and then estimated the corresponding classification accuracies. This overall procedure was repeated 1000 times and resulted in a distribution of accuracies under the null hypothesis for each participant. We attributed a percentile score to the participants' true classification accuracy by computing the percentage of observations in the simulated distribution with classification accuracies below the true classification accuracy. The median of the 31 participants' percentile scores was 95.3, indicating that the class labels and observations are not independent as the null hypothesis would predict and based on a 5% significance threshold (we report the median and not the mean because of the bimodal distribution of percentile scores created by the participants with a classification accuracy below chance level).

When inspecting the individual classification accuracies (see Supplementary Table 3.1), we noticed that classification accuracy is below chance level for four participants. For those participants, it does not make sense to use activity in the F-

ROI to distinguish between low and high levels of arousal. A possible explanation is that these participants did not follow the instructions, and were not genuinely trying to empathize with the IAPS pictures shown on the screen (as could be deducted from the log book in two cases). Because we did not want to exclude participants ad hoc based on results, they are still part of the data in the main chapter. The results of the subsequent analyses on ad level were robust to removing the four cases with a classification accuracy below 50, and three cases with a classification accuracy equal to 50 (chance level).

Supplementary Table 3.1. *Descriptives of Classification Accuracies across Participants*

Descriptives	Classification accuracy in %
M	55.1
SD	5.6
Range	45.0 - 65.0
95% CI	53.1 - 57.1

Appendix 3.F Analysis without individual calibration

As mentioned in the final section on the statistical analyses on the EEG data in the main chapter (section on EEG sample data), not conducting the individual calibration would result in less clean (/more crude) relationships between the EEG activity that is measured in response to the print ads, and the arousal process that it should represent across individuals. A less refined estimation of arousal would thereby exist of merely selecting EEG data from the arousal F-ROI as determined in the first step of the analysis (i.e., activity at 7 – 12 Hz, at central electrode sites), but measured while viewing the print ads, for every participant. By averaging this (Z-transformed) data across the selected electrodes and frequencies, and also across participants per advertisements, we similarly arrive at a neural arousal score for each print ad, thus making the advertisement our focal unit of analysis. This is a similar approach to the approach in study two, and the neural arousal score in study two represents (inverted) alpha activity instead of the probability that the ad evokes arousal. Finally, to test whether the neural arousal scores of the print advertisements are related to ad effectiveness in the population at large, we computed the correlation between the 'uncalibrated' neural arousal score of the ads, and averaged ratings of the ads by the population sample on Notability and Attitude. As one might expect, these resulting correlation coefficients are somewhat lower than when we use the calibrated neural arousal score, namely rho = .13 (p = .12) compared to rho = .20 for Notability when we used individual calibration, and rho = -.19 (p = .02) compared to rho = -.26 for Attitude when we used individual calibration (both p-values uncorrected for multiple comparisons). Please note, that neural arousal scores are inverted in order to avoid confusion about the inverse relation between alpha activity and arousal. These results illustrate the value of calibration to account for heterogeneity

Appendix 3.G EEG recording and analysis of Study 2

The EEG data was acquired in a similar manner as in study 1. All preprocessing was again done in Brain Vision Analyzer software. The data was first down-sampled to 256 Hz. Then the data was re-referenced to the averaged mastoids, and filtered with a low cutoff filter of 1 Hz and a notch filter of 50 Hz with a slope of 48 dB/octave. Thereafter, the data was segmented into 44 segments (one for each commercial), with segments lasting from the beginning to the end of the commercial. We then split the segments further into 50% overlapping segments of 256 data points. We applied Gratton and Coles ocular correction as implemented in BVA, and standard artifact detection and rejection criteria where segments are rejected that contain jumps larger than 30μV/ms, amplitude differences exceeding 150μV/ 200ms, and amplitude differences below 0.5 μV/ 100 ms. Note that only the channels that contained artifacts were deleted within the given segment, and not the entire segment. Then, data was decomposed into components of different frequencies using a FFT (using a 100% hanning window). Finally, we averaged the spectral data across all segments for each presentation of the commercial and for each participant individually. The resulting spectral data was exported to Matlab (Mathworks).

Chapter 4

EEG metrics relating to population-wide commercial success of movies: A meta-analysis

Introduction

Neuromarketing, which is the application of neuroscientific tools and methods to marketing theory and practice, has become more and more popular since its inception in 2002 (for reviews see e.g., Ariely & Berns, 2010; Levallois, Smidts, Wouters, 2019; Plassmann, Venkatraman, Huettel, & Yoon, 2015; Smidts et al., 2014). The main proposed explanation for this popularity is the hope that it reveals insight that cannot be obtained through traditional methods applied in marketing. Traditional methods such as surveys or focus groups, typically rely on the respondent indicating a preference, or reflecting on the experience of a product. Although such methods certainly have their merits, the disadvantage is that they depend on the respondent's ability to reflect on an experience and use of introspection, as well as on proneness to social desirability. Neuroscientific methods are free of these kinds of biases, providing more direct insight into the processes underlying a decision or experience.

The most commonly employed method in neuromarketing practice, and particularly in ad testing, is electroencephalography (EEG) (Smidts et al., 2014). The advantages of this method are that the costs are relatively low, and the temporal resolution high, compared to other methods such as functional magnetic resonance imaging (fMRI) (Ariely & Berns, 2010), making it especially suited for analysing dynamic stimuli such as TV commercials or movies. EEG measures the potential for electrical current to pass between two sites expressed in voltage (Luck, 2005), and these electrical properties reflect the summed activity in millions of neurons at the surface of the brain (Stern, Ray, Quigley, 2001). Since the introduction of EEG by Hans Berger in 1929, it has been noted that changes in an individuals' engagement of an activity, co-occur with changes in frequency (i.e., the number of oscillations per second) and amplitude of the EEG signal. This resulted in a convention to define oscillations in the EEG signal in terms of different frequency bands that are typically associated with psychological constructs. Particularly these psychological constructs are of interest to practitioners in neuromarketing, because they enable practitioners to assess the processing of marketing stimuli such as products or advertisements. Here, we will focus on the following frequency bands, or metrics, that have been studied extensively: theta, alpha, (frontal) alpha asymmetry, beta and gamma.

Ample research supports the association between activity in the *theta* frequency band (oscillations between 4 and 8 Hz) and memory processes (e.g., Buzsaki, 2005; Jensen, 2005; Jensen & Lisman, 2005; Kirk & Mackay, 2003; Vertes, 2005; for a review regarding the involvement of the different frequency bands in motivation, emotion, and their inhibitory control, see Knyazev, 2007). Memory and emotion are strongly related because of their tight link to relevance for survival, since emotional events are more likely to be relevant for survival, and thus less likely to be forgotten (Phelps, 2004). Indeed, theta activity has also been linked to processing of emotional information (Aftanas, Varlamov, Pavlov, Makhnev, Reva, 2001; Nishitani, 2003), as well as more specifically to losses or other negative outcomes (Cohen, Elger, & Ranganath, 2007; Marco-Pallares et al., 2008; Van de Vijver, Ridderinkhof, & Cohen, 2011). Moreover, in a marketing context, activity in the theta band has been shown to be positively related to memory for TV commercials (Vecchiato et al., 2010; Vecchiato et al., 2012).

Activity in the *alpha* frequency band (8 to 12 Hz) reflects the most dominantly present oscillations in the brain, and it has been shown to be inversely related to attention (e.g., Dockree et al., 2004; Klimesch, 1998) or more general information processing (Pfurtscheller & Lopes da Silva, 1999). In line with this, a strong negative correlation was found between activity in the alpha band and activity in lateral frontal and parietal areas that are known to support attentional processes, in a study in which EEG was simultaneously acquired with fMRI (Laufs et al., 2003). Attention has furthermore been examined in the context of emotion, sometimes also referred to as motivated or affective attention (Uusberg, Uibo, Kreegipuu, & Allik, 2013). In these kind of studies, pictures were presented that elicited different levels of emotional arousal, and thereby varied in degree of relevance to be processed. Indeed, results revealed a negative association between alpha band activity and - in this case- processing of arousing stimuli (DeCesarei & Codispoti, 2011; Simons, Detenber, Cuthbert, Schwartz, & Reiss, 2003, but see Aftanas, Varlamov, Pavlov, Makhnev, & Reva, 2002; Uusberg et al., 2013).

A third metric concerns *frontal asymmetry*. Particularly in the alpha band, the asymmetry in activity between hemispheres has been a subject of interest for more than 50 years. The results of early studies suggested that greater left than right frontal activation (i.e., increased alpha band activity in the right compared to the left) is associated with the experience of positive affect, and greater right than left frontal activation with the experience of negative affect (e.g., Davidson & Fox, 1982). In those studies, however, emotional valence (positive versus negative affect) was confounded with motivational direction (approach versus withdrawal states). When emotional states that differed in valence were evoked, for example by the presentation of positive and negative pictures, those states could have also been interpreted as approach and withdrawal motivational states (for a review see Harmon-Jones, Gable, & Peterson, 2010).

More recent research therefore aimed at distinguishing emotional valence from motivational direction when studying the asymmetry in frontal cortical activity. Results from those studies suggest that greater left than right frontal activity is associated with approach behavior, not necessarily positive affect, and greater right than left activity with withdrawal behavior and not necessarily negative affect. Evidence for this motivational direction theory independent of valence of the emotion is primarily originating from research on anger, an emotion negative in valence but associated with approach behavior. The findings are not only of correlational nature; both experimentally manipulating anger (e.g. Harmon-Jones, Sigelman, Bohlig, & Harmon-Jones, 2003) as well as manipulating the asymmetry in frontal activity using rTMS (e.g., Van Honk & Schutter, 2006) provided support for this theory. Thus, although frontal EEG asymmetry seems to reflect approach motivation, emotional valence may not necessarily be extracted from it. Still, approach-related behavior is often of appetitive, and even of pleasant nature, and indeed a decrease of activity in the alpha band across left frontal areas was found to be associated with the judged pleasantness of TV commercials (Vecchiato et al., 2011).

Introducing a fourth metric, activity in the *beta* band (oscillations between 12 to 30 Hz) has been related to reward processing in various contexts. Increased activity in the beta band has been found in monetary incentive tasks when participants received rewards, and when they anticipated rewards (Cohen et al., 2007; Kawasaki & Yamaguchi 2013; Marco-Pallares et al., 2008). Furthermore, Van de Vijver at al. (2011) showed that beyond the reception or anticipation of money, beta band activity was enhanced when participants received positive feedback (versus negative feedback), and that it predicted learning. It is thereby suggested to play a role in reinforcement learning, in which outcomes are "evaluated" in order to guide future reward-seeking behavior (Cohen et al., 2007). In a study with a more practical setup, the authors hypothesized an association between reward processing and brand evaluation (Lucchiari & Pravettoni, 2012).

Supporting their hypothesis, the results revealed that the experience of pleasure that was associated with a favourite brand, seemed to modulate activity in the beta band.

Finally, activity in the *gamma* band (oscillations > 30 Hz) has traditionally been associated with perceptual processing and attention. An increase in activity in the gamma band has been found when participants focused their attention on auditory stimuli (Tiitinen et al. 1993), on visual, moving stimuli (Gruber, Muller, Keil, Elbert, 1999), and also when they shifted their attention within static pictures in a visual search task (Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1997). In this visual search task, participants viewed a picture containing black blobs. The picture seemed meaningless at first, but when participants learned how to perceive a dog in this exact same picture, activity in the gamma band increased. It is thereby suggested that increased (synchrony of) gamma band activity enables large-scale communication and integration of parallel processes across distributed brain areas. Indeed, gamma band activity has not only been found to accompany focused attention, but also tasks involving object recognition (Keil, Muller, Ray, Gruber, & Elbert, 1999; Rodriguez, George, Lachaux, Martinerie, Renault, & Varela, 1999), and emotional evaluation (Müller, Gruber, & Keil, 2000), implying that gamma band activity is not indicative of these specific processes as such. In line with this, Herrmann et al. (2004) propose a model in which they explain activity in the gamma band in terms of a match between the bottom-up (sensory encoding) and top-down (e.g., attention, memory) information.

Based on this knowledge about how neural activity extracted from the overall EEG signal relates to psychological constructs, practitioners in neuromarketing formulate and develop specific metrics in order to assess the effectiveness of marketing stimuli (e.g., Pradeep, 2010 [founder & CEO NeuroFocus, Inc.]). However, by extracting such metrics, the assumptions are that the metrics tap into the associated psychological constructs, and subsequently, that they relate to (preferably population-wide) commercial success. Therefore, over the past decade, researchers in neuromarketing have started to investigate the possibility of using neural data collected in response to marketing stimuli in a relatively small group of people ('neural focus group', Falk, Berkman, & Lieberman, 2012), to forecast (the consequence of the stimulus on future) behavior of a larger group of people. The behavior of a large group of people thereby represents real-world market level success, and thus the term 'neuroforecasting' has been adopted in this context (Genevsky, Yoon, & Knutson, 2017; for an overview, see Knutson & Genevsky, 2018). Multiple fMRI studies revealed that, in particular, activity in the nucleus accumbens (NAcc) and medial prefrontal cortex (mPFC) predicts the aggregate behavior of larger groups such as song downloads (Berns & Moore, 2012), funding success (Genevsky et al., 2017), and ad related click through (Falk et al., 2016).

The previously mentioned examples of studies all employed fMRI, and although EEG is the most frequently used imaging technique in neuromarketing, the spatial resolution is much lower compared to fMRI. EEG records voltage changes at the scalp, with volume conduction of the brain as one of the difficulties resulting in the inverse problem (i.e., inferring combinations of multiple sources of activity underlying the measured activity). Thus, whether EEG can be employed to measure neural activity in response to marketing stimuli in order to predict market level success, is largely an open question. In recent years, a few studies have addressed this question (for a review, see Hakim & Levy, 2018). Dmochowski et al., (2014) were one of the first to show that the neural reliability across multiple viewers of naturalistic stimuli, as expressed by the intersubject correlation (ISC) of the EEG signal, could predict behavior of a large population. The neural reliablity predicted audience preferences in terms of tweet frequency during an episode of a popular television series ("The Walking Dead"), and population ratings of advertisements (SuperBowl commercials). In another study on super bowl advertisements, it was demonstrated that activity in the theta band contributed to predicting the number of you tube views (Guixeres et al., 2017). Activity in the theta, beta, and gamma band was positively associated with in-sample ad liking and recall in their study. Venkatraman et al. (2015) examined predicting market level success in the form of ad elasticity, and the results showed a moderate relationship between ad elasticity and the two included EEG metrics: (occipital) alpha band activity and frontal asymmetry.

In three recent studies, the neural activity in response to watching movie trailers was measured in order to investigate whether it related to population-wide commercial success of the movies they promoted. In a seminal study, a whole-brain analysis revealed that activity in the gamma band predicted commercial success of the movies as expressed in terms of box office, beyond a traditional measure such as liking, while activity in the beta band was related to in-sample preference (Boksem & Smidts, 2015). In a second study, the cross-brain-correlation (i.e., neural similarity), measured in activity in the alpha band, predicted future free recall of the movie trailers and population-level sales of the corresponding movies (Barnett & Cerf, 2017). In addition in a third study, neural similarity (cognitive-congruency) in the gamma band predicted corresponding movies' sales performance of premiere weekends, and of sales up to 4 weeks (but not neural similarity in the beta band) (Christoforou, Papadopoulos, Constantinidou, & Theodorou, 2017).

It is no coincidence that over the past five years, three articles have been published on the relationship found between neural activity as measured by EEG in response to watching movie trailers, and population-wide success of movies. In the risky movie business, these movie trailers serve as advertisements in order to promote commercially released movies. Although the global box-office for all movies that were released across the world reached $41.1 billion in 2018 (Motion Picture Association of America [MPAA], 2019), most movies are not profitable (De Vany & Walls, 1999). With such high stakes, the quality of the movie trailers is of paramount importance in order to attract the audience to watch the upcoming movies. Indeed, movie trailers are an important source of information for consumers in determining which movie to watch in the future (Barnett, White, & Cerf, 2016; Gazley, Clark, and Sinha 2011). All in all, it is therefore worthwhile to investigate whether EEG metrics can be applied effectively in the production and optimization stage of the movie trailer before its release.

Although all three studies on movie trailers show that EEG can be applied effectively in predicting population-wide success, both the number of participants and the number of stimuli are relatively small, as is the case in most neuroimaging studies. In addition, there are notable differences between the analyses that were conducted: Not only in pre-processing of the EEG data (e.g., presence of correction for eye movements), but also in the specific neural activity that was extracted, and even differences in the measures of the movies' commercial success. Therefore, our aim is to re-analyze all data from these studies using the same analysis strategy and to relate this neural data to the same market level outcome/ measure of success. This data set preferably includes data from the three studies on movie trailers that were published: Boksem & Smidts, 2015, Christoforou et al, 2017, and Barnett & Cerf, 2017. In addition, new data has been collected in our lab in two additional studies that have not yet been published on. At the time of writing, we are still in the process of acquiring the data from Barnett & Cerf, 2017, thus, here we only include the data from two published studies (Boksem & Smidts, 2015; Christoforou et al., 2017) and the two non-published studies. Our aim is to re-analyze this large dataset consisting of four studies in a similar fashion, in order to investigate the presence and robustness of effects in a larger sample of participants and stimuli.

More specifically, we extract neural activity from the five previously introduced frequency bands (i.e., theta, alpha, alpha asymmetry, beta, and gamma), often used in neuromarketing practice and found to be related to psychological constructs. We aimed to examine whether neural activity, represented by these five EEG metrics, in response to watching movie trailers is related to population-wide success of the corresponding movies expressed in terms of U.S. box office. Not only the size of the dataset (data on a total of 130 participants and 149 movie trailers) is an important feature of this study, but also the fact that we extract multiple EEG metrics (Hakim & Levy, 2018). Finally, for findings to be of value for practitioners, it is important to know whether measuring neural activity contributes to the prediction of commercial success above and beyond already available traditional measures. Therefore, we investigated whether neural activity is predictive of success above and beyond in-sample liking and genre of the movies, both assumed to be related to the movies' commercial success.

Methods

Participants

All participants had normal or corrected-to-normal vision and were fluent in English. Across all four studies, some participants were excluded from the final analyses due to technical problems with the EEG equipment and/ or problems with the recordings causing excessive artifacts in the EEG data. For the same reason, we excluded (these) participants in the current analysis.

In Study 1 (Boksem & Smidts, 2015), 32 students were recruited from a university population (in the Netherlands). They received 25 euro in return for their participation in the study. The final sample included in the analysis consisted of 29 participants (13 women, M = 21.5 years of age, SD = 2.8, Min = 18, Max = 28).

In Study 2 (new data), 40 students were recruited from a university population in the Netherlands. They received 25 euro in return for their participation in the study. The final sample included in the analysis consisted of 39 participants (21 women, M = 20.3 years of age, SD = 1.6, Min = 17, Max = 24).

In Study 3 (new data), 42 students were recruited from a university population in the Netherlands. They received 25 euro in return for their participation in the study. The final sample included in the analysis consisted of 39 participants (24 women, M = 21.9 years of age, SD = 3.2, Min = 18.3, Max = 34.9).

In Study 4 (Christoforou et al, 2017), 27 participants (16 women) were recruited in Cyprus (Median = 22 years of age, Min = 19 Max = 24). They received compensation in return for their participation in the study. Data from three participants was excluded in the original analysis, and additionally from one other participant in the current analysis.

This resulted in an overall sample of 130 participants across the four studies (between 17 and 35 years of age).

Stimuli

Across all studies, the stimuli consisted of official movie trailers in English. In studies 1, 2, and 3, the selection procedure of movies promoted by the trailers was largely similar. The Internet Movie Database (IMDb.com) was searched for the selection of the trailers in these studies. To ensure reasonable quality of the movies, the movies needed to have received a rating of at least 5.5 stars (out of 10) based on at least 1000 votes. Animated movies, cartoons, 3D movies, horror movies, movies with excessive violence, and remakes or sequels were excluded.

In Study 1, the stimulus set consisted of 56 trailers from movies that were released between 2000 and 2010, from the four genres action, adventure, drama, thriller. The final selection was based on ranking of U.S. box office results, and the best 150 movies were excluded to avoid familiarity with the movies. To ensure considerable variation in commercial success, movies were selected with rank 150, 200, 250, and so on, for each genre. Trailer duration had to be 2-2.5 minutes If the movie for this rank did not meet the criteria, the movie for the next rank was selected (e.g., 151 instead of 150). Commercial success of the movies that the movie trailers promoted as expressed in terms of U.S. box office varied between $4.4 million to $121 million ($M$ = $47 million, SD = $28 million).

For Study 2, the stimuli consisted of 18 trailers that were released between 2000 and 2010, from the genres drama, thriller, action, adventure, and fantasy. We used the same stimuli that were selected for the study of Chan et al. (2019). The selection was based on ranking of U.S. box office results, and the best 49 movies were excluded to avoid familiarity with the movies. To ensure considerable variation in commercial success, movies were selected with rank 50, 60, 70, and so on, until a box office of $100 million was reached. Minimal box office had to be $1 million. If the movie for this rank did not meet the criteria, the movie for the next rank was selected, and this resulted in a preselection of 168 movies. From these 168 movies, eighteen movies were selected that were seen the least by 60 participants from a pilot study. Commercial success of the final 18 movies that the movie trailers promoted as expressed in terms of U.S. box office varied between $1.4 million to $92.2 million ($M$ = $20.4 million, SD = $23.9 million).

For Study 3, the stimuli consisted of 60 trailers that were released between 2010 and 2016, from the genres adventure, drama, and thriller. The final selection was based on a balanced distribution of U.S. box office results, and the best 50 movies were excluded to avoid familiarity with the movies. Additional exclusion criteria were movies based on books, based on famous people, and based on computer games, because the popularity of the movie could depend less on the quality of the trailer/ movie, and an U.S. box office of less than \$5 million. Commercial success of the movies that the movie trailers promoted as expressed in terms of U.S. box office varied between \$6 million to \$126.5 million (*M* = \$54.9 million, *SD* = \$32.2 million).

In Study 4, the boxofficemojo.com website was searched for selection of the trailers. The stimuli in this study consisted of 15 trailers that were released in the second or third quarter of 2014, from the three genres action, adventure, thriller. To ensure reasonable quality and market reach of the movies, the movies had to be released in at least 1000 theaters with a total revenue of at least \$10 million. The final selection was based on ranking of box office performance, and focused only on the top 100 movies of each quarter. To ensure considerable variation in commercial success, movies were selected with both the highest and lowest rankings. Commercial success of the movies that the movie trailers promoted as expressed in terms of U.S. box office varied between \$14 million to \$333.2 million (*M* = \$115.6 million, *SD* = \$101.4 million).

This resulted in a total number of 149 movie trailers that were presented across four studies. Out of these trailers, four identical trailers were presented across two different studies. The commercial success as expressed in terms of U.S. box office of the movies that all movie trailers promoted varied between \$1.4 million and \$333.2 million (median = \$43.3 million, *M* = \$53.9 million, *SD* = \$47.8 million). The evaluations of the movies that the trailers promoted as expressed by IMDB ratings varied between 3.1 to 8.3, and the average evaluation was very similar across studies (*M* = 6.6 on a scale from 1-10). The variation in evaluation was larger in study 4 compared to in other studies, as was the case for commercial success.

In three out of the four studies (excluding Study 2 where all participants viewed the same trailers), the participants viewed only a subset of movie trailers out of the entire stimulus set. In order to construct a subset of trailers for a specific participant, in Study 1, trailers from the movie genre that the participants preferred the least (action, drama, adventure, or thriller) were removed first. Trailers from a movie that the participant had already seen were removed additionally. From the trailers that remained, six were selected randomly from each of the three remaining genres. The final stimulus subset that was presented to the participants thereby consisted of 18 trailers. In Study 3, the construction of a stimulus subset specific for each participant was very similar. In addition to excluding trailers from movies that participants had seen already, participants had to indicate their least preferred genre (adventure, drama, thriller). They viewed 20 trailers, preferably not from their least liked genre (only when too few trailers were left from other genres that had not been seen before). In Study 4, the stimulus set of 15 trailers was split into two subsets, one of eight trailers and one of seven trailers. Participants were randomly assigned to viewing one of the subsets. Participants viewed the stimulus subset a second time in this study, but we only included the first presentation of the trailers in the analysis to keep situations constant across studies.

This resulted in a total number of views by participants that was at least 2 per trailer, and maximally 39 per trailer (median = 12) (see *Figure 4.1*).

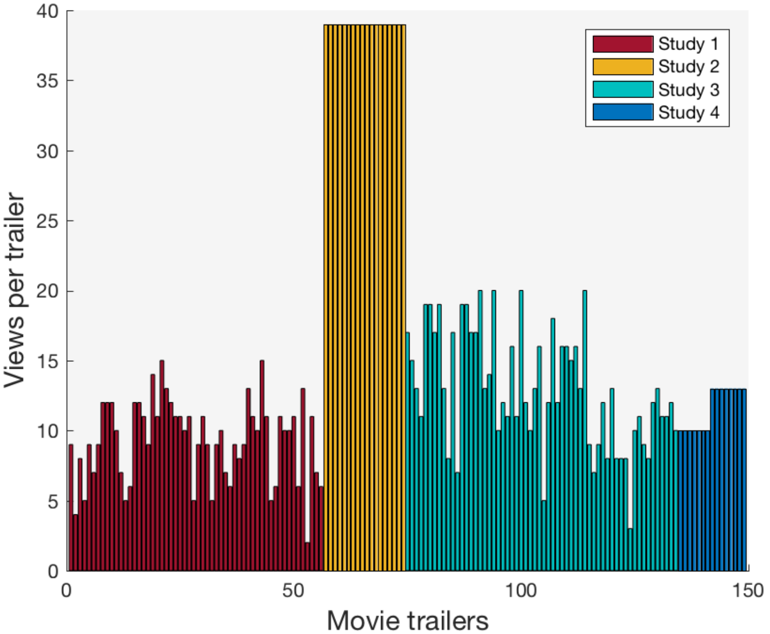


Figure 4.1. Number of views by participants per trailer across all four studies

Data collection procedure

Prior to the start of all data collection sessions, participants first received information and instructions about the tasks, and consent forms from all participants were obtained. Then, the participants were seated in a comfortable chair, the electrodes were applied, and the data recording prepared. In three out of the four studies, the recording occurred in an electrically shielded, sound-attenuated room that was dimly lit (for study 4, this condition is unknown). For Study 1, 2, and 3, a 19-inch PC monitor was positioned at 1.80 meters from the participants, while for Study 4 the distance from the monitor to the participants was 60 cm because of the simultaneous collection of eye-gaze data. In all studies, the stimuli were presented randomly to the participants. After the presentation of each trailer, an on-screen question was presented asking the participants to indicate the degree to which they liked the movie(trailer) on an 11-point scale (0-10) for study 1, 2, 3, and on a scale from 1-10 for study 4.

EEG recording and pre-processing

All EEG data was recorded with the BioSemi ActiveTwo system, and active Ag-AgCl electrodes fixed in an elastic cap. For the Studies 1, 2, and 3, data was recorded with 64 electrodes, and for Study 4 with 32 electrodes fitted in the corresponding standard caps according to the international 10/20 system.

We executed all preprocessing of the EEG data in Brain Vision Analyzer software (BVA; Brain Products). We possessed the original raw data of three out of the four studies: Study 1, Study 2, and Study 3. First, we down-sampled this data to 256 Hz, re-referenced the activity to the average mastoid signal, and filtered it with a 1 Hz low cutoff filter (24 dB/octave) and 50 Hz notch filter. The data of Study 4 that we obtained was also sampled with a rate of 256 Hz, but re-referenced to the average activity of all electrodes, since no additional (flat-type) electrodes were used other than the cap electrodes. In addition, this data was filtered with a 1.5 Hz low cutoff filter, and a 50 Hz notch filter.

Thereafter, we corrected the EEG activity for artifacts originating from eye movements using independent component analysis as implemented in the BVA software. In Study 1, 2, and 3, we used flat-type electrodes positioned above and below the left eye to detect vertical eye movements (i.e., blinks), and electrodes positioned to the outer canthi of both eyes to detect horizontal eye movements. In Study 4, we used electrode Fp1, and electrode F7 referenced to F8 for those purposes, respectively.

From this point onwards, all pre-processing steps were performed in an identical fashion. We split the continuous data into segments lasting from the beginning of the stimuli to the end of the stimuli, with each segment consisting of data recorded during watching one movie trailer. We then divided these segments further into segments consisting of 256 data points, with 50% overlap. We applied standard artifact detection and rejection criteria to these segments, where we rejected segments that contained jumps larger than 30µV/ms, amplitude differences exceeding 150µV/ 200ms, and amplitude differences below 0.5 µV/ 200 ms. Note that only the channels that contained artifacts were deleted within the given segment, and not the entire segment. Thereafter, we decomposed the data into different frequencies (1-128 Hz) using a Fast Fourier Transform (FFT, using a 100% hanning window). Finally, we averaged the frequency data across all segments for each video, and for each participant separately. The resulting frequency data was exported to MATLAB (MathWorks). Note that this results for each movie trailer in averaged frequency data across the entire stimulus duration (but for each electrode, for each frequency).

Statistical analyses

EEG metrics. To create the five EEG metrics theta, alpha, alpha asymmetry, beta, and gamma, we sampled neural activity from specific groups of electrodes, per frequency band, based on the literature (as previously presented in the Introduction). Before averaging the activity across such a group of electrodes and frequencies, we first transformed the data, starting with taking the natural logarithm in order to reduce skewness, followed by standardization (i.e., z-transformation of the data per participant, per electrode, per frequency, across stimuli). For the 64-electrode cap and the theta metric, we sampled activity at 4-8 Hz from electrodes AFz, F1, Fz, F2, FC2, FCz, FC1 (see *Figure 4.2.A*). For the alpha metric, we sampled activity at 8-12 Hz from electrodes CP2, CPz, CP1, P1, Pz, P2, POz (see *Figure 4.2.B*). We computed the alpha-asymmetry measure by subtracting the natural logarithm of alpha power at the left hemisphere from the natural logarithm of alpha power at the right hemisphere, which is most commonly reported (ln[right alpha] − ln[left alpha], e.g., Coan & Allen, 2004). Higher scores on this scale indicate relatively more left frontal activation, and lower scores relatively more right frontal activation, because of the putative inverse relation between alpha power and activation. We sampled activity at 8-12 Hz from right hemispheric electrodes at AF4, F6, F4, FC6, FC4, and from left hemispheric electrodes at AF3, F5, F3, FC5, FC3 (see *Figure 4.2.C*). For this metric, the data was standardized before averaging the activity across the group of electrodes and frequencies, and after subtraction of left from right electrodes. For the beta metric, we sampled activity at 16-20 Hz from electrodes AFz, F1, Fz, F2, FC2, FCz, FC1 (see *Figure 4.2.D*). For the gamma metric, we sampled activity at 75-95 Hz from electrodes, F1, F2, FC4, FC2, FC1, FC3, C3, C1, C2, C4 (see *Figure 4.2.E*).

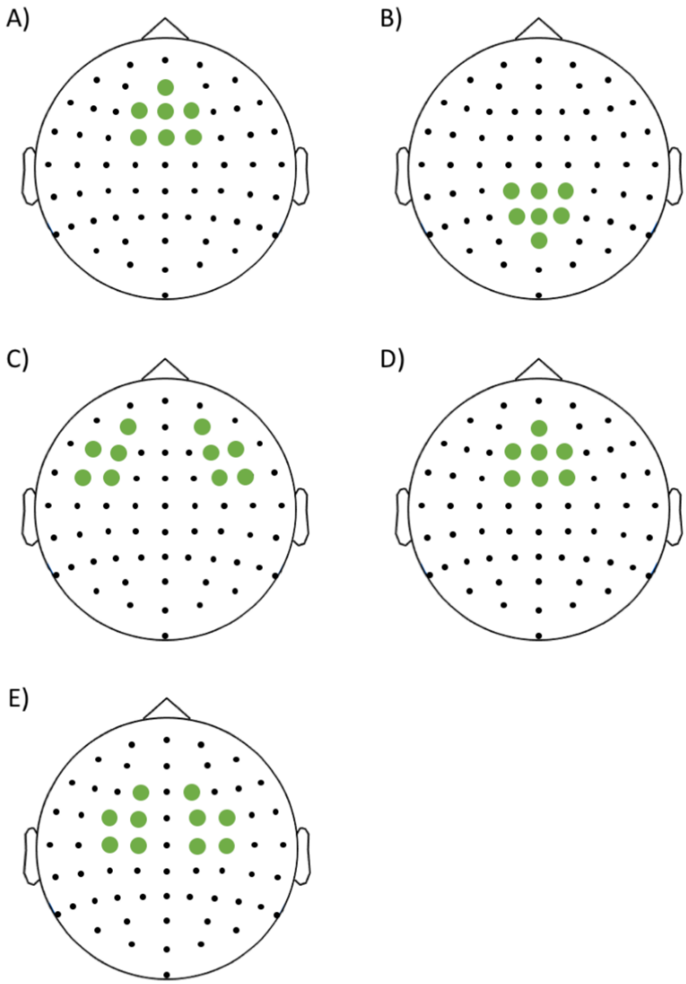


Figure 4.2. Electrode selection for calculation of each EEG metric. A) Theta 4-8 Hz. B) Alpha 8-12 Hz. C) Alpha asymmetry 8-12 Hz. D) Beta 16-20 Hz. E) Gamma 75-95 Hz.

In Study 4, a 32-electrode cap was used. Since this cap does not contain all electrode locations from the 64-electrode cap, we only sampled activity from those electrodes that were present in the before-mentioned metrics. For the 32-electrode cap and theta metric, we sampled activity at 4-8 Hz from electrodes FC1, FC2, Fz. For the alpha metric, we sampled activity at 8-12 Hz from electrodes CP1, CP2, Pz. For the beta metric, we sampled activity at 16-20 Hz from electrodes FC1, FC2, Fz. For the gamma metric, we sampled activity at 75-95 Hz from electrodes FC2, FC1, C3, C4. For alpha asymmetry, we sampled activity at 8-12 Hz from right hemispheric electrodes at AF4, F4, FC6, and from left hemispheric electrodes at AF3, F3, FC5.

Finally, we averaged the z-scored activity across these groups of electrodes and frequencies resulting in the five EEG metrics.

Genre classification. We took genre of the movie that the trailer promoted into account in the analysis in order to control for the effect of movie genre on box office. Since usually more than one genre is ascribed to a movie, and genres were not mutually exclusive, we re-classified the movies and their genres. IMDb.com was scraped for all genre tags belonging to each movie, which resulted in a total of 19 genres with at least one movie per genre. We narrowed down the number of genres and minimalized overlap by taking into account that the genres should differ in box office, potentially attract different audiences, and should contain enough movies within the genre to represent the (variety of movies per) genre. Ultimately, all trailers that the movies promoted should be assigned to at least one genre (but could be assigned to more genres). This resulted in assigning the movies/ trailers to a set of genres consisting of: action, biography, drama, mystery, thriller, and family/ comedy taken together (see Appendix 4.A for descriptives of box office per genre).

Statistical models/ tests. First, we re-scaled (z-transformed) the in-sample liking measure per participant. Second, in order to determine whether the EEG metrics provided unique information about commercial success beyond genre and liking, we fitted two linear mixed effects models (fitlme function from MATLAB) with box office as our DV. In the first, restricted model, we regressed box office onto the dummies for the six genres, and liking, and added random intercepts for study and participant. In the second model, we added the five EEG metrics as predictors to the first model. We compared the goodness of fit between the two models with and without the EEG metrics by conducting a likelihood-ratio test, and we report the fixed effects of predictors from the second, full model.

Results

We first checked whether multicollinearity imposed a problem on entering all EEG metrics in the same model by inspecting the correlations and VIF values among the metrics. This was not the case since the maximum correlation among metrics was .47 and the average VIF 1.31 (all < 1.57). We removed 39 observations from the movie trailer promoting the movie The Town in study 2, because it represented an outlier with respect to box office within that study (Z-score = 3.0; for robustness checks regarding sensitivity to outliers, including results using all data, see Appendix 4.C-4.D). The likelihood ratio test showed that entering the EEG metrics in addition to genre and liking to predict box office improved the fit of the model significantly ($\chi^2(5) = 13.83$, $p < .05$). The results of the full model including all EEG metrics showed that in addition to the significant effects of genre and liking on box office, only the effect of gamma was significant ($b = 2.73$, $t(2126) = 2.87$, $p < .005$, see Table 4.1). The effect of alpha asymmetry was marginally significant (but see robustness checks in Appendix 4.C-4.D: only the effect of gamma is insensitive to the presence of outliers). The results were very similar/ almost identical when liking was removed as predictor from the model.

Table 4.1. *Results for fixed effects in full model predicting box office*

Predictor	Estimate	*t* -value	*p* -value
Action	10.86	5.30	.000
Biography	-8.57	-3.28	.001
Drama	-6.79	-3.36	.001
Mystery	10.04	3.76	.000
Thriller	4.25	2.15	.032
Family_Comedy	29.09	9.28	.000
Liking	1.70	2.13	.034
Theta	-1.41	-1.12	.264
Alpha	0.81	0.64	.524
Alpha asymmetry	-2.50	-1.83	.068
Beta	-0.46	-0.38	.707
Gamma	2.73	2.87	.004

In order to rule out the possibility that the effect of gamma was caused or influenced by inclusion of the other EEG metrics in the model, we conducted backward elimination: We removed the least significant EEG metric as predictor, and estimated the model again, and repeated this process until only significant ($p < .05$) metrics were left in the final model. The results of the final model (as well as the models in between, see Appendix 4.B) showed that including the other metrics, hardly influenced the effect of gamma on box office (see Table 4.2).

Table 4.2. *Results for fixed effects in final model predicting box office*

Predictor	Estimate	*t* -value	*p* -value
Action	11.06	5.40	.000
Biography	-8.73	-3.34	.001
Drama	-6.86	-3.42	.001
Mystery	10.02	3.76	.000
Thriller	4.34	2.19	.028
Family_Comedy	29.21	9.32	.000
Liking	1.62	2.04	.042
Gamma	2.62	2.97	.003

For illustrative purposes only, we averaged the data across participants and presented the relationship between the gamma metric and box office per study (Figure 4.3, see Appendix 4.E for figures including the other metrics).

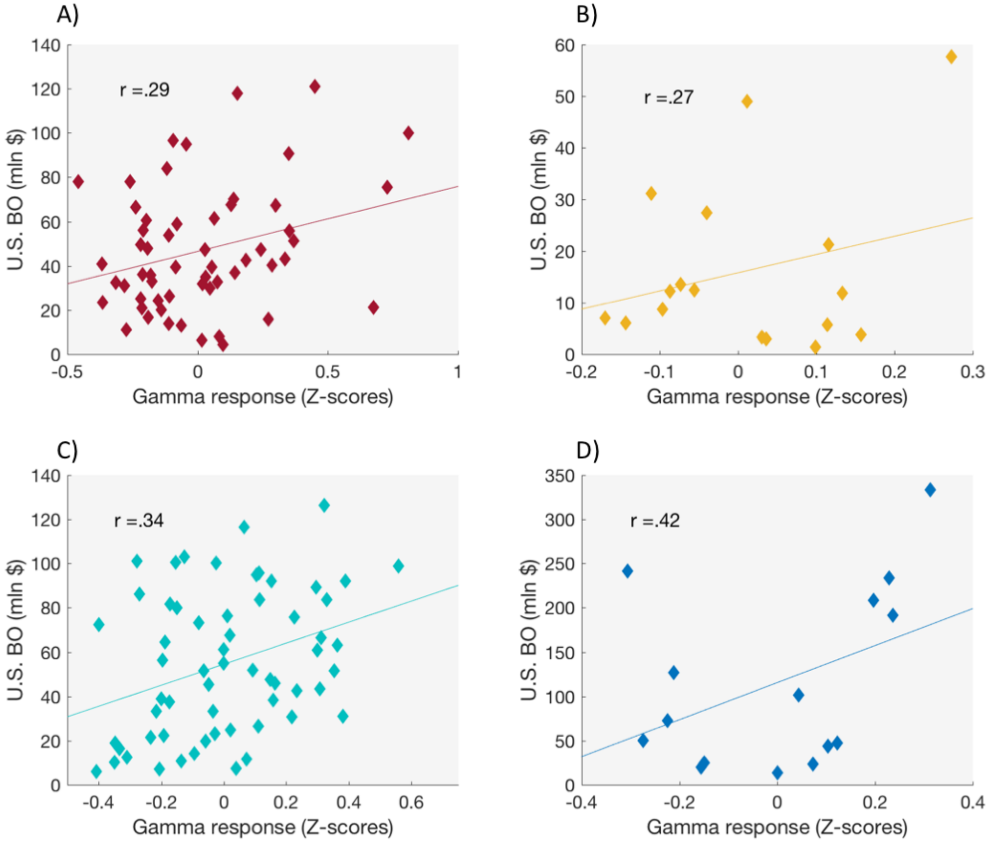


Figure 4.3. Relationship between gamma metric and box office per study on stimulus level. A) Study 1. B) Study 2. C) Study 3. D) Study 4. (r = Pearson correlation)

Discussion

In the current study, we combined the data from four studies in which neural activity in response to movie trailers was investigated using EEG, for a systematic re-analysis of all data. We examined whether five metrics (theta, alpha, alpha asymmetry, beta, and gamma) extracted from the EEG signal, are predictive of population-wide success of the corresponding movies expressed in terms of U.S. box office, above and beyond traditionally available information. The results show that the fit of the model improved significantly when we included the five EEG metrics in addition to the restricted model containing only genre and liking in predicting box office.

We tested the contribution of neural activity to the model conservatively by adding all five EEG metrics at once to the model containing genre and liking. Only the effect of gamma band activity appeared to be significant, and the effect was consistent when the other non-significant EEG metrics were removed from the model via backward elimination. The results of Boksem & Smidts (2015), were thus replicated in a much larger sample of participants and stimuli consisting of four studies (Christoforou et al. 2017 also found an effect of gamma but looked at similarity/ coherence in activity).

In considering an explanation for their finding, Boksem & Smidts (2015) already noted that gamma band activity has been associated with activity in the mPFC in a study with combined EEG and fMRI analysis (Mantini, Perrucci, Del Gratta, Romani, & Corbetta, 2007). Importantly, this is in line with the results of the fMRI studies referred to in the introduction, which revealed that also activity in this specific cortical area related to population-wide commercial success in various contexts (Falk et al., 2012; Falk et al., 2016; Kühn, Strelow, & Gallinat, 2016; Scholz, Baek, O'Donnell, Kim, Cappella, & Falk, 2017). Boksem & Smidts (2015) further suggested that enhanced activity in the gamma band may not only reflect that the viewed material captures attention, but also that it is more extensively processed and memorized. Indeed, the deeper lying "structure" of the hippocampus (critical for the encoding and recollection of events; Phelps, 2004) has extensive connections with the mPFC (Pochon et al., 2002), and intracranial EEG recordings revealed that successful memory formation was associated with the synchronization of gamma band activity between the hippocampus and other connected brain areas (Fell et al., 2001). All in all, these findings may suggest that increased gamma band activity during viewing a movie trailer could be associated with enhanced processing and memorization, which may have resulted in an increased likelihood of people going to watch the promoted movie, hence increasing box office.

Movie trailers not only act as an advertisement for the movie, but also let one already consume part of the product. As a consequence, movie trailers are very informative on the content of the movie and thus for determining whether to watch a movie in the future (Barnett et al., 2016; Gazley et al., 2011). In comparison to the other metrics, perhaps gamma activity is most closely related to integration of information across different brain areas in different stages of (perceptual) processing during viewing a trailer. We speculate that increased gamma band activity during a trailer hereby possibly reflects deeper processing of its content (of the trailer and thus part of the movie). Effective processing of the movie trailer narrative is almost a necessary condition for being able to decide whether one would like to watch the movie.

None of the other EEG metrics (i.e., theta, alpha, alpha-asymmetry, and beta) recorded during viewing the trailers, was significantly and consistently predictive of the movies' box office. The effect of the alpha asymmetry metric was marginally significant however, although sensitive to the presence of outliers, and negative. This would suggest that relatively more right versus left frontal activation (i.e., more alpha activity at left versus right hemisphere, because of putative reverse relation between alpha band activity and activation), associated with withdrawal motivation, is related to higher box office results. Reports about the asymmetry measure are not consistent/unambiguous in the literature though, both concerning ways to compute it and concerning the results (e.g., Hagemann, Naumann, Thayer, 2001; Murphy, Nimmo-Smith, & Lawrence, 2003), and our results suggest that this metric should be applied with caution in practice.

It should also be noted here that one has to be careful with interpreting the relationships between EEG metrics and psychological constructs because of the reverse inference problem (Poldrack, 2006). Merely based on the presence of specific neural activity, we can not infer that particular psychological functions were necessarily engaged in complex contexts such as advertising, in which no control exists over the presence and absence of all kinds of stimulus features. Nevertheless, it is relevant to know that such a consistent relationship exists between specific neural activity in a small group of people and real world commercial success (versus only a relationship with in-sample behavior), regardless of the conceptual interpretation.

A distinguishing feature of this study is that we examined the effects of *multiple* EEG metrics, and in a much larger sample of stimuli and participants than usual. By combining stimuli and participants from multiple studies, the data also becomes less homogeneous, and thus effects are more likely to be generalizable than effects from only one study. Conducting such a meta-analysis or re-analysis is important, because it provides insight into the robustness of findings from single, perhaps underpowered studies. Moreover, new results could even appear only in a larger sample with sufficient power. Importantly, Boksem & Smidts' (2015) finding that gamma band activity in response to trailers is related to total box office results, was replicated in a much larger sample here which increases confidence in the result.

An outstanding question for further research, however, is whether the results are specific for movie trailers as stimulus, and box office as dependent variable; it may be that activity in other frequency bands perhaps predicts success for other types of dynamic stimuli or other measures of commercial success (e.g., IMDB-ratings, more closely related to average attitude). In addition, not all types of neural activity contributing to an individual's decision or experience, are potentially equally suited to predict aggregate behavior, as proposed by Knutson & Genevsky (2018). This holds even more for activity recorded with EEG, because not knowing the exact source of the activity is problematic for understanding the mechanism contributing to an individual's choice (let alone aggregate behavior). Meta-analyses and/ or re-analyses with large data sets therefore do provide additional insight on the question of whether certain EEG metrics are better suited in predicting aggregate behavior or market level success.

In a very recent study, the EEG metrics formulated to denote specific psychological constructs, consisted of neural activity combined across frequency bands (Shestyuk, Kasinathan, Karapoondinott, Knight, & Gurumoorthy, 2019). By combining fronto-central alpha with beta asymmetry (suggested to index approach motivation), fronto-central alpha with theta activity (suggested to index attention), and fronto-central theta with gamma activity (suggested to index memory processing), it was found that each of these combined metrics recorded during viewing of prime-time TV shows was associated with TV viewership and Twitter volume. Further research could study whether these combined metrics are also predictive of commercial success of movie trailers.

Taking the combination of activity across frequency bands one step further, Hakim et al. (2018) sampled activity from multiple frequency bands and used machine learning algorithms to predict participants' individual future preferences and commercials' population success. Their results revealed that utilizing the EEG measurements added value beyond using questionnaires in predicting success. Future research should demonstrate the robustness of such machine learning findings if the goal of prediction becomes more important than the interpretation of the type of neural activity that is predictive. Further research could also explore the inter-subject correlation (ISC), an additional promising candidate metric not addressed in the current study, reflecting the reliability of neural activity across multiple people (e.g., Dmochowski et al., 2014).

In the current study, we showed that EEG metrics, and more specifically activity in the gamma band, contributed significantly beyond the traditionally available information such as genre and in-sample liking of the movie (trailer) in predicting box office. Investigating the contribution of neural measures above and beyond traditional measures is of particular relevance for managerial implications, and is not necessarily standard practice in research (Hakim & Levy, 2018). Here we assessed the added value of neural measures beyond in-sample liking. This in-sample liking is measured on a relatively small and not representative sample of the population. Expanding our work, future research could examine the contribution of neural activity beyond self-report measures collected in large samples (e.g., in online panels such as MTurk). Moreover, in order to estimate the relationships between EEG metrics, traditionally available information and box office, we fitted a prediction model to the entire data set in the present analysis. A next step would be to investigate how these relationships generalize to new data, for example, by leaving out a subset of the data, in order to truly test the prediction results of the model.

The present results demonstrated that recording EEG added value to predicting success in addition to asking participants how much they liked the movie (trailer), and knowing the movie genre. Neural activity in the gamma band in response to the trailers explained approximately 10 percent of the variance in box office (as inferred from the correlation coefficients at stimulus level, see *Figure 4.3*), and the coefficient of 2.73 would imply that an increase of one standard deviation in gamma band activity would relate to an increase of $2.73 million in box office (keeping other variables constant). Given the average box office of $53.9 million, this increase in gamma band activity would thereby be associated with an increase of roughly five percent of the total returns of the movie.

A potential direction for further research would be to explore the influence of the dynamics in activity in the gamma band over time on this effect. Perhaps the peaks in activity more strongly contribute to the predictive effect of gamma than the currently measured average gamma activity (see, for example, the peak-end rule posing that people tend to rely heavily on the intensity of peak and final moments, e.g., Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993; Do, Rupert, & Wolford, 2008). It might be very informative to see whether such a peak is related to responses to specific types of scenes or particular events in the narrative. Yet, maximizing average gamma band activity in response to movie trailers, for example, by cutting out scenes with low gamma responses, would potentially already be an effective strategy for neuromarketing practice, thus supporting decisions on release and production of the trailer as an advertisement for a movie, and thus potentially increase the revenue.

In sum, we demonstrated that EEG activity in response to viewing of movie trailers, is related to the corresponding movies' box office above and beyond traditional measures. More specifically, the metrics we extracted from the EEG signal were theta, alpha, alpha-asymmetry, beta, and gamma, but only gamma appeared to be a significant predictor of the movies' success. Furthermore, we conducted a comprehensive re-analysis in which we combined the samples of participants and stimuli from four studies. The present study and results thereby provides additional support for the growing body of research suggesting that measuring EEG in response to marketing stimuli may be useful for marketing purposes.

Appendix

Appendix 4.A Box office genre descriptives

Supplementary Table 4.1. *Descriptives of U.S. box office per genre across trailers*

Genre	Mean	SD	N trailers
Action	65.34	56.24	80
Biography	24.46	19.98	13
Drama	43.97	34.89	91
Mystery	50.56	35.73	17
Thriller	56.96	31.95	47
Family_Comedy	97.51	88.62	15

Appendix 4.B Backward elimination

In order to rule out the possibility that including other EEG metrics in the model influenced the effect of gamma on box office, we conducted backward elimination: We removed the least significant EEG metric as predictor and estimated the model again, and repeated this process until one EEG metric was left. The results in Supplementary Table 4.2 showed that removing the other metrics, hardly changed the effect of gamma on box office.

Supplementary Table 4.2. *Results for all models in backward elimination in predicting box office*

Full model				Model 1st predictor removed				Model 2nd predictor removed				Model 3rd predictor removed				Final model			
Predictor	*b*	*t*	*p*	Predictor	*b*	*t*	*p*	Predictor	*b*	*t*	*p*	Predictor	*b*	*t*	*p*	Predictor	*b*	*t*	*p*
Action	10.86	5.30	.000	Action	10.86	5.30	.000	Action	10.84	5.29	.000	Action	10.89	5.32	.000	Action	11.06	5.40	.000
Biography	-8.57	-3.28	.001	Biography	-8.58	-3.28	.001	Biography	-8.56	-3.28	.001	Biography	-8.59	-3.29	.001	Biography	-8.73	-3.34	.001
Drama	-6.79	-3.36	.001	Drama	-6.82	-3.38	.001	Drama	-6.74	-3.35	.001	Drama	-6.94	-3.46	.001	Drama	-6.86	-3.42	.001
Mystery	10.04	3.76	.000	Mystery	10.06	3.77	.000	Mystery	9.96	3.74	.000	Mystery	9.88	3.71	.000	Mystery	10.02	3.76	.000
Thriller	4.25	2.15	.032	Thriller	4.27	2.16	.031	Thriller	4.30	2.17	.030	Thriller	4.18	2.12	.034	Thriller	4.34	2.19	.028
Fam_Com	29.09	9.28	.000	Fam_Com	29.14	9.31	.000	Fam_Com	29.12	9.30	.000	Fam_Com	29.17	9.31	.000	Fam_Com	29.21	9.32	.000
Liking	1.70	2.13	.034	Liking	1.69	2.12	.034	Liking	1.65	2.08	.038	Liking	1.66	2.09	.037	Liking	1.62	2.04	.042
Theta	-1.41	-1.12	.264	Theta	-1.53	-1.25	.211	Theta	-1.22	-1.13	.259	Alpha asym	-2.49	-1.82	.069	Gamma	2.62	2.97	.003
Alpha	0.81	0.64	.524	Alpha	0.65	0.54	.587	Alpha asym	-2.50	-1.83	.068	Gamma	2.55	2.89	.004				
Alpha asym	-2.50	-1.83	.068	Alpha asyr	-2.49	-1.82	.068	Gamma	2.62	2.96	.003								
Beta	-0.46	-0.38	.707	Gamma	2.60	2.93	.003												
Gamma	2.73	2.87	.004																

Appendix 4.C Robustness check 1: Results including all data

In order to rule out the possibility that removing an outlier (The Town in the original analysis in the manuscript) changed the results, we conducted the same analysis again, including backward elimination, but with the complete dataset (see Supplementary Table 4.3). Except for the small decrease in the effect of gamma, the results were very similar (significant effect of gamma, marginally significant of alpha asymmetry).

Supplementary Table 4.3. *Results for all models in backward elimination in predicting box office, including all data*

Full model				Model 1st predictor removed				Model 2nd predictor removed				Model 3rd predictor removed				Final model			
Predictor	*b*	*t*	*p*	Predictor	*b*	*t*	*p*	Predictor	*b*	*t*	*p*	Predictor	*b*	*t*	*p*	Predictor	*b*	*t*	*p*
Action	9.92	4.73	.000	Action	9.93	4.73	.000	Action	9.89	4.71	.000	Action	9.95	4.74	.000	Action	10.12	4.82	.000
Biography	-9.82	-3.67	.000	Biography	-9.82	-3.67	.000	Biography	-9.79	-3.65	.000	Biography	-9.83	-3.67	.000	Biography	-9.97	-3.72	.000
Drama	-4.59	-2.22	.026	Drama	-4.60	-2.23	.026	Drama	-4.47	-2.17	.030	Drama	-4.72	-2.30	.021	Drama	-4.65	-2.27	.024
Mystery	8.51	3.11	.002	Mystery	8.52	3.12	.002	Mystery	8.36	3.06	.002	Mystery	8.24	3.02	.003	Mystery	8.37	3.07	.002
Thriller	10.01	5.07	.000	Thriller	10.02	5.08	.000	Thriller	10.09	5.12	.000	Thriller	9.96	5.05	.000	Thriller	10.11	5.13	.000
Fam_Com	31.83	9.93	.000	Fam_Com	31.86	9.95	.000	Fam_Com	31.84	9.94	.000	Fam_Com	31.91	9.95	.000	Fam_Com	31.95	9.96	.000
Liking	2.28	2.80	.005	Liking	2.27	2.80	.005	Liking	2.21	2.73	.006	Liking	2.22	2.74	.006	Liking	2.18	2.69	.007
Theta	-2.09	-1.63	.103	Theta	-2.17	-1.74	.081	Theta	-1.63	-1.48	.138	Alpha asym	-2.44	-1.75	.080	Gamma	2.29	2.55	.011
Alpha	1.23	0.94	.345	Alpha	1.12	0.92	.358	Alpha asym	-2.45	-1.76	.078	Gamma	2.22	2.47	.014				
Alpha asym	-2.45	-1.76	.079	Alpha asym	-2.45	-1.76	.079	Gamma	2.31	2.56	.010								
Beta	-0.30	-0.24	.813	Gamma	2.28	2.53	.011												
Gamma	2.36	2.45	.014																

Appendix 4.D Robustness check 2: Results without Guardians of the Galaxy as outlier

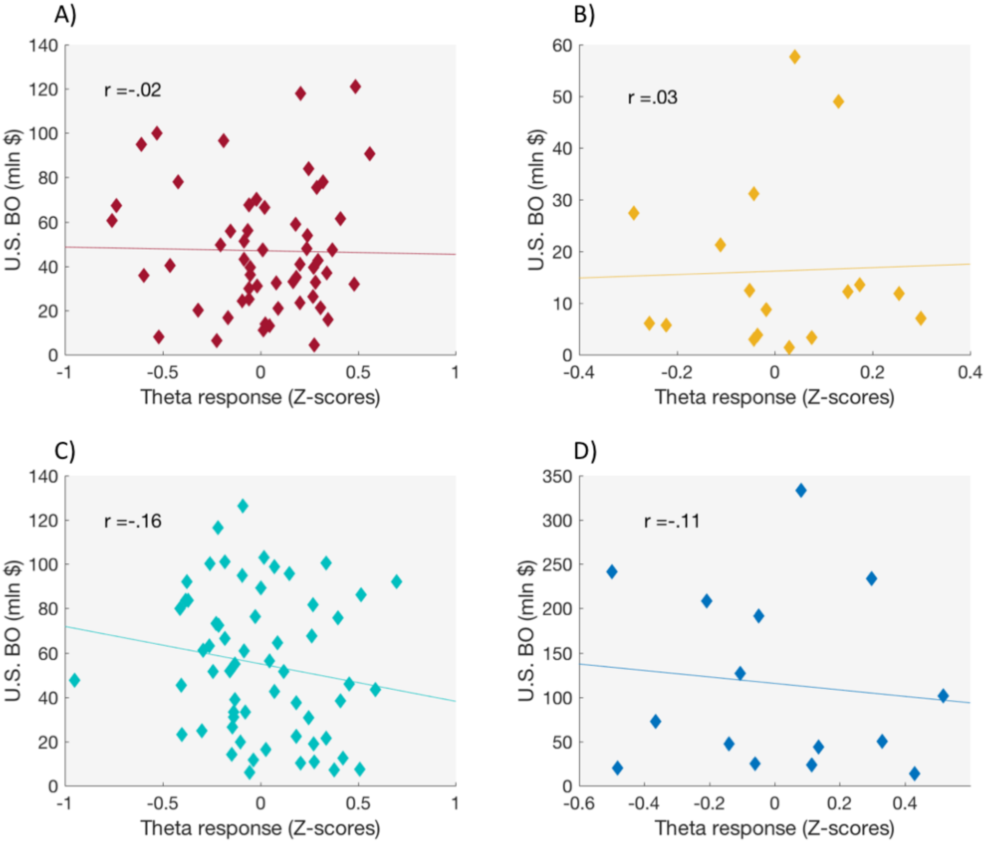
Fitting the linear mixed effects model resulted in a couple of large residuals which were caused by a trailer with the largest box office (Guardians of the Galaxy from Study 4). In order to rule out the possibility that these observations influenced the results, we conducted the same analysis again, including backward elimination, but without these observations (see Supplementary Table 4.4). Compared to the original analysis in the manuscript, the effect of gamma decreased only slightly, but the marginal significant effect of alpha asymmetry disappeared.

Supplementary Table 4.4. *Results for all models in backward elimination in predicting box office, without additional 'outlier'*

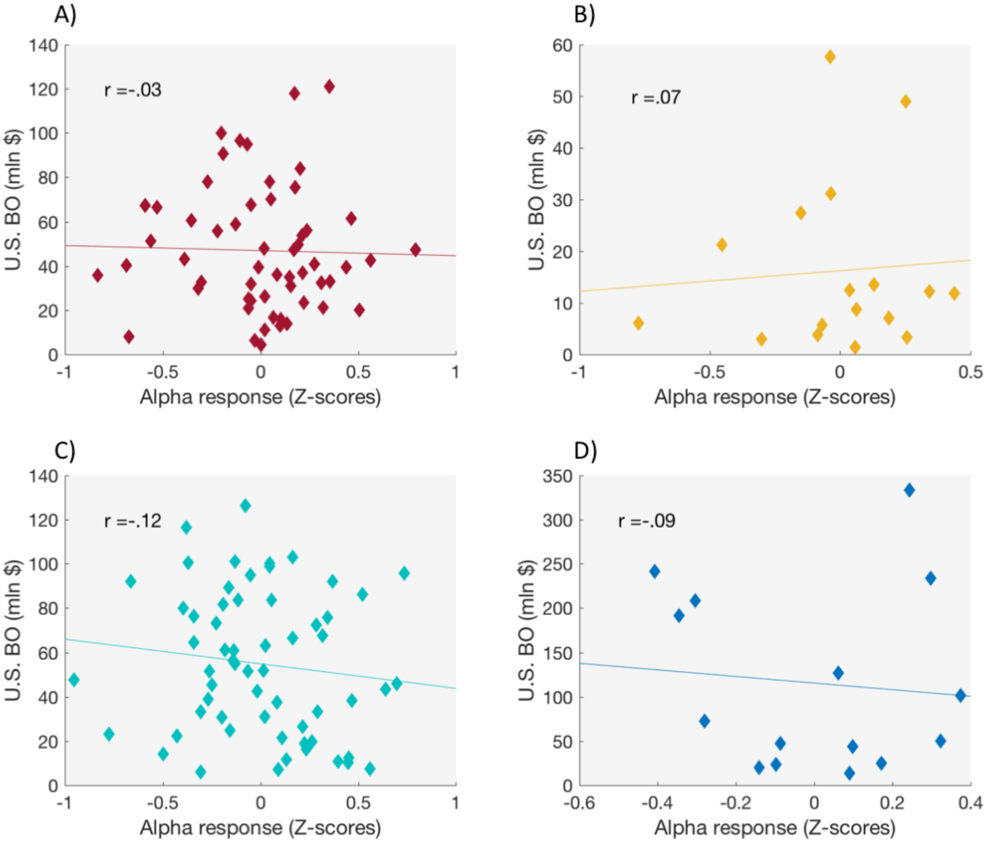
Full model				Model 1st predictor removed				Model 2nd predictor removed				Model 3rd predictor removed				Final model			
Predictor	*b*	*t*	*p*	Predictor	*b*	*t*	*p*	Predictor	*b*	*t*	*p*	Predictor	*b*	*t*	*p*	Predictor	*b*	*t*	*p*
Action	9.74	5.18	.000	Action	9.74	5.18	.000	Action	9.74	5.18	.000	Action	9.85	5.24	.000	Action	9.91	5.27	.000
Biography	-9.76	-4.07	.000	Biography	-9.76	-4.07	.000	Biography	-9.76	-4.07	.000	Biography	-9.86	-4.11	.000	Biography	-9.89	-4.12	.000
Drama	-7.20	-3.88	.000	Drama	-7.19	-3.89	.000	Drama	-7.21	-3.90	.000	Drama	-7.16	-3.87	.000	Drama	-7.42	-4.03	.000
Mystery	9.40	3.84	.000	Mystery	9.39	3.84	.000	Mystery	9.40	3.85	.000	Mystery	9.49	3.88	.000	Mystery	9.37	3.84	.000
Thriller	4.63	2.55	.011	Thriller	4.64	2.55	.011	Thriller	4.64	2.56	.011	Thriller	4.74	2.61	.009	Thriller	4.59	2.53	.011
Fam_Com	17.45	5.95	.000	Fam_Com	17.45	5.95	.000	Fam_Com	17.46	5.96	.000	Fam_Com	17.47	5.96	.000	Fam_Com	17.54	5.98	.000
Liking	1.76	2.39	.017	Liking	1.75	2.40	.017	Liking	1.76	2.40	.016	Liking	1.73	2.36	.018	Liking	1.74	2.38	.018
Theta	-1.54	-1.33	.184	Theta	-1.52	-1.39	.164	Theta	-1.57	-1.59	.113	Theta	-1.57	-1.58	.114	Gamma	2.27	2.79	.005
Alpha	0.08	0.07	.948	Alpha asyr	-1.62	-1.28	.199	Alpha asym	-1.62	-1.28	.199	Gamma	2.35	2.90	.004				
Alpha asym	-1.62	-1.28	.199	Beta	-0.13	-0.12	.906	Gamma	2.31	2.84	.005								
Beta	-0.15	-0.13	.894	Gamma	2.34	2.70	.007												
Gamma	2.35	2.69	.007																

Appendix 4.E Illustration of relationship between EEG metrics and box office per study averaged across participants

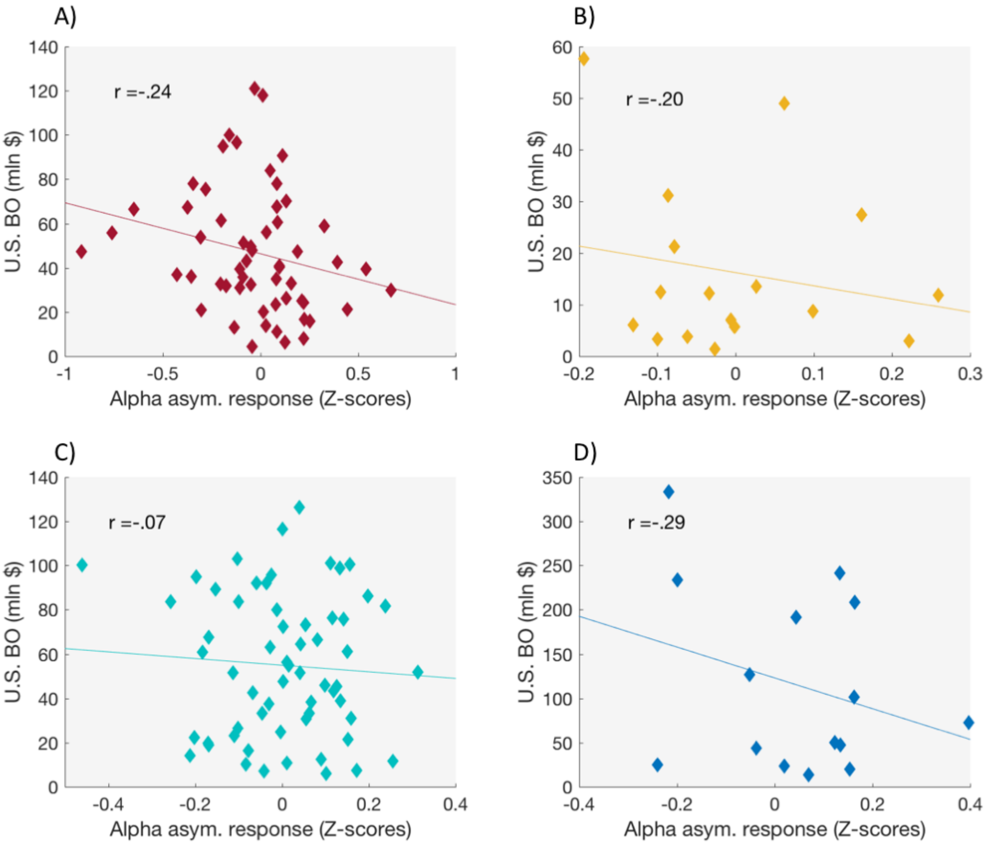
Figures are based on the same data as presented in the main chapter.



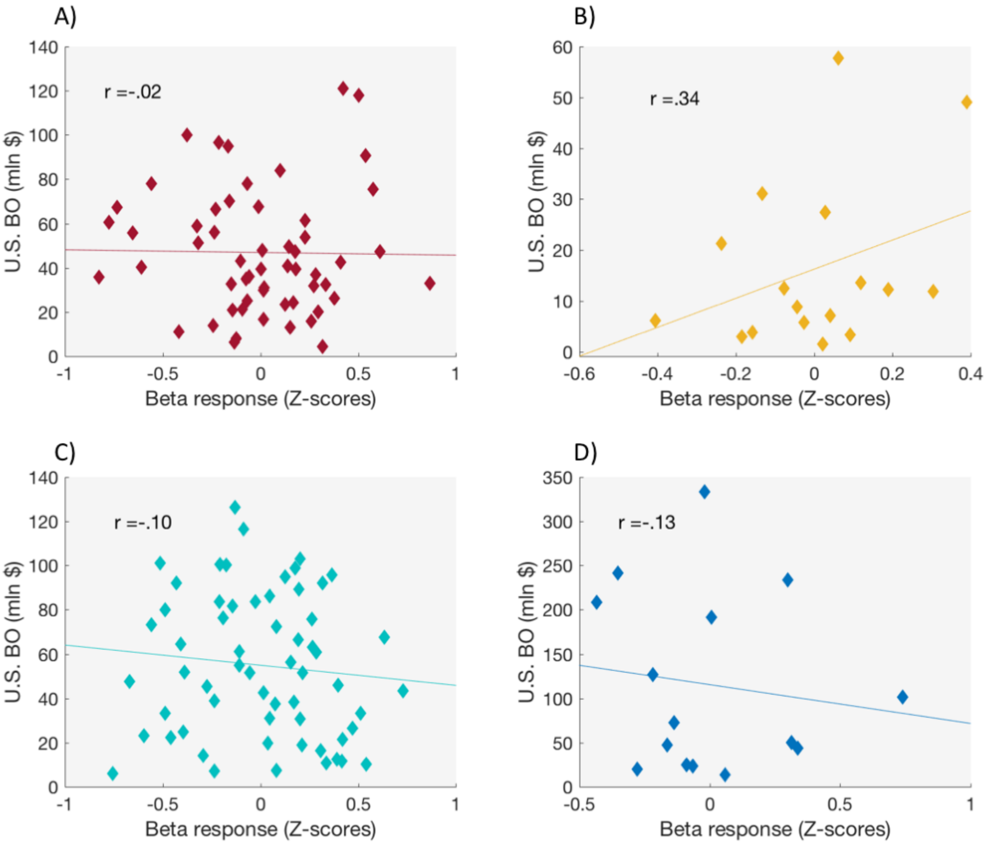
Supplementary Figure 4.1 Relationship between theta metric and box office per study on stimulus level. A) Study 1. B) Study 2. C) Study 3. D) Study 4. (r= Pearson correlation)



Supplementary Figure 4.2. Relationship between alpha metric and box office per study on stimulus level. A) Study 1. B) Study 2. C) Study 3. D) Study 4. (r = Pearson correlation)



Supplementary Figure 4.3. Relationship between alpha asymmetry metric and box office per study on stimulus level. A) Study 1. B) Study 2. C) Study 3. D) Study 4. (r = Pearson correlation)



Supplementary Figure 4.4. Relationship between beta metric and box office per study on stimulus level. A) Study 1. B) Study 2. C) Study 3. D) Study 4. (r = Pearson correlation)

Chapter 5

General discussion

In this dissertation, I investigate the (emotional) response elicited by marketing related stimuli using EEG, and to what extent such a response is associated with advertising effectiveness. It thereby extends existing knowledge by elucidating the two proposed aims of neuromarketing: offering additional insight into implicit emotional processes and contributing to predicting behavioral, market level, responses. Chapter 2 focusses on processing of the marketing stimulus itself, whereas the focus in Chapter 3 and 4 is additionally shifted towards the effectiveness of the marketing stimulus at the population level. In each chapter I discussed the detailed findings. Here I provide a short summary of the core findings.

In Chapter 2, we aimed to distinguish different emotional experiences elicited by audiovisual stimuli designed to evoke particularly happy, sad, fear and disgust responses, using EEG and a multivariate pattern approach. Although many studies revealed that emotions and their dynamics have a profound impact on cognition and behavior, it has proven difficult to unobtrusively measure emotions. We show that we were able to classify these four emotional experiences based on respondent's EEG activity well above chance level. Importantly, we retained all the information (frequency and topography) present in the EEG data. This allowed us to interpret the differences between emotional experiences in terms of component psychological processes such as attention and arousal that are suggested to be associated with the observed activation patterns. In addition, we illustrate how this method of classifying emotional experiences can be applied on a moment-by-moment basis in order to track dynamic changes in the emotional response over time.

In Chapter 3, we showed a refined way of first estimating how arousal is represented in the brain via a separate task (using EEG), and thereafter using this representation to measure arousal in response to advertisements. As such, we estimate the relationship between the a priori identified process (arousal) and external measures of ad effectiveness (as measured by notability, attitude toward the ad, and choice) in the population at large. Across two studies, the results show that the neural measure of arousal in response to advertisements is positively associated with notability of ads in the population at large, but negatively associated with attitude and behavioral measures toward these ads.

In Chapter 4, we combined the data from four studies in which neural activity in response to movie trailers was investigated using EEG, for a systematic re-analysis of all data (n = 130 participants, k = 145 stimuli). We examined whether five metrics (theta, alpha, alpha asymmetry, beta, and gamma) extracted from the EEG signal, are predictive of population-wide success of the corresponding movies expressed in terms of U.S. box office, above and beyond traditionally available information. The results show that the fit of the multilevel model improved significantly when we included the five EEG metrics in addition to the model containing only genre and liking in predicting box office. This suggests that the brain signal contains unique information that contributes to explaining the effectiveness of the stimulus. Importantly, only the effect of gamma band activity appeared to be significant, and the effect was consistent when the other non-significant EEG metrics were removed from the model via backward elimination. The results of Boksem & Smidts (2015) were hereby replicated in a much larger sample of participants and stimuli consisting of four studies.

In addition to these substantive findings, this dissertation also contributes methodologically to the neuromarketing field by i) applying novel multivariate methods to decode emotional experiences, ii) using a localizer task in an EEG study to reduce the reverse inference problem that commonly plagues neuroimaging research, and iii) conducting a major meta-analysis to address the issue of small samples sizes regarding both participants and stimuli in neuromarketing research.

Managerial Implications

Although I discussed study specific managerial implications in the respective chapters, I will focus here on the overall implications of the research presented in this dissertation. The core managerial contribution of the current research consists of showing a novel way of measuring customer experience, in addition to demonstrating associations between EEG metrics measured in response to marketing stimuli and their effects at the population level.

In Chapter 2 I show how the toolbox of practitioners in the neuromarketing industry can be expanded in order to increase understanding of consumers' processing of marketing stimuli, beyond mere attention or arousal. Using machine learning tools, we developed an algorithm that enables classifying the emotional response as happy, sad, fear or disgust based on the EEG signal recorded during an audiovisual stimulus. This method provides insight on the moment-by-moment specific emotional effect that a marketing stimulus such as a TV-commercial, website or customer service has, on consumers. By being able to measure and track these specific emotional experiences in "real time", the diagnostic value strongly improves and it thus may provide advertisers and marketers with more concrete and actionable insights on how to design and optimize the stimulus. The programming code is now publicly available in order for its application to be extended.

In Chapter 3 I found that a refined way of measuring arousal does not seem to be very discrepant from how attention or arousal is commonly measured in practice (that is, although most neuromarketing practitioners are not disclosing their precise measures, arousal is usually measured by extracting alpha band activity at the posterior side of the head). Our findings thus provide a stronger underpinning of current practice. Importantly, the findings of this chapter also imply that different levels of neural arousal or alpha activity evoked in response to advertisements, cannot automatically assumed to be associated with ad effectiveness in general. Ads that are neurally arousing may be effective in the sense that they are highly notable and thus enhance awareness, but may not necessarily be effective in terms of being positively evaluated or indeed acted upon. Neuromarketing practitioners should thus take such caution when advising clients on their arousal measures.

Chapter 4 demonstrates that out of the five EEG metrics, only gamma band activity (measured in response to movie trailers) appeared to be related to real-world market level success. This implies that practitioners should focus especially on the average and dynamic pattern of gamma band activity when optimizing movie trailers towards box office effects. The costs of measuring gamma of a trailer for a relatively small sample of customers are negligible in comparison to the production costs of the movie and may yield important insights for creating a more effective trailer. Future research should further develop and examine the robustness of these findings by including other stimuli than movie trailers and different outcome measures from box office.

Directions for Future Research

Although the results of the studies in this dissertation offer important implications for neuromarketing (practice), further research is needed to address additional important questions.

In Chapter 3 and Chapter 4 I average the neural activity measured in response to the stimuli across the entire stimulus duration, and examine how it relates to the response at market level. I did not assess, however, the dynamics of neural activity. Perhaps peaks in neural activity or other dynamic patterns possibly contribute more strongly to the predictive effects than the average activity. For example, one might consider the peak-end rule, posing that people tend to rely heavily on the intensity of peak and final moments (e.g., Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993), or primacy and recency effects (Ebbinghaus, 1902). It might therefore be very informative to investigate whether such influential peaks relate to responses to specific types of scenes or particular events in the narrative, and to study the importance of the storyline. Liu, Shi, Teixeira, & Wedel (2018) provide an interesting modeling procedure on how to select scenes and optimize clips, based on facial-expression tracking, but more research is needed to investigate how the EEG metrics and/ or processes measured over time are related to effectiveness of the stimuli.

Although I find an association between neural activity measured in response to the stimuli and market level effectiveness in Chapter 3 and Chapter 4 (and I do not rely on reverse inference, and extract multiple EEG features, respectively), additional statistical tests would be needed to answer the question of how well neural data contributes to the *prediction* of market level success. Ideally, methods used in data science should be adopted in order to test the true *prediction* results of the models by leaving out a subset of the data, and thus to test how such relationships generalize to new data (Hakim & Levy, 2018). In addition, other metrics computed from the EEG signal such as the intersubject or cross-brain correlation (Dmochowski et al., 2014; Barnett & Cerf, 2017), or emotional responses classified based on the method developed in Chapter 2 are potential valuable candidates to include in future prediction analyses. This should be a next step in this line of research.

Furthermore, it is important (perhaps especially for neuromarketing practice) to gain more insight into the unique contribution of recording EEG, beyond the application of other methods in (neuro)marketing. In Chapter 4 we took in-sample self-reported liking of the stimuli into account, representing a traditional survey measure, when examining the additional value of EEG in predicting market level success. This approach could be extended by not merely taking liking into account, but also other self-report variables (e.g., self-reported predicted success), or even larger out-of-sample self-reports collected from online panels such as Amazon Mechanical Turk for example. In addition, it would be recommended to study the contribution of EEG more extensively in direct comparison to other neurophysiological methods (e.g., building on the seminal study by Venkatraman et al., 2015). Aligning EEG with other scalable methods such as eye tracking (e.g., Teixeira, Wedel, & Pieters, 2010) and facial expression detection (e.g., Lewinski, Fransen, & Tan, 2014; Teixeira, Wedel, & Pieters, 2012) that measure a consumers' response directly, could prove to be very fruitful and beneficial in assessing each methods' unique contribution and added value.

Finally, a weakness of EEG is its low spatial resolution. Aligning EEG with fMRI, by applying both methods to the same stimuli, would result in more knowledge about the source of activity of EEG metrics. It would thereby provide stronger underpinnings for neuromarketing in practice and stimulate development of the field by identifying underlying mechanisms.

Final Note

All in all, the current dissertation provides additional support for the growing body of research suggesting that measuring EEG in response to marketing stimuli is useful for marketing purposes, albeit within limits/ specific contexts. The interdisciplinary field of neuromarketing is still relatively new and this dissertation is an attempt at strengthening the foundation upon which academia and practitioners from neuromarketing industry can build the field together.

References

Aftanas, L. I., Varlamov, A. A., Pavlov, S. V., Makhnev, V. P., & Reva, N.V. (2002). Time-dependent cortical asymmetries induced by emotional arousal: EEG analysis of event-related synchronization and desynchronization in individually defined frequency bands. *International Journal of Psychophysiology, 44*, 67-82.

Aftanas, L. I., Varlamov, A. A., Pavlov, S.V., Makhnev, V. P., & Reva, N. V. (2001). Affective picture processing: Event-related synchronization within individually defined human theta band is modulated by valence dimension. *Neuroscience Letters, 303*, 115-118.

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84,* 888-918.

Ariely, D., & Berns, G. S. (2010). Neuromarketing: the hope and hype of neuroimaging in business. *Nature Reviews Neuroscience, 11*, 284-292.

Armony, J. L., & Dolan, R. J. (2002). Modulation of spatial attention by fear-conditioned stimuli: an event-related fMRI study. *Neuropsychologia, 40*, 817-826.

Bagozzi, R. P., Gopinath, M., & Nyer, P. U. (1999). The Role of Emotions in Marketing. *Journal of the Academy of Marketing Science, 27*, 184–206.

Barnett, S. B. & Cerf, M. (2017). A ticket for your thoughts: Method for predicting content recall and sales using neural similarity of moviegoers. *Journal of Consumer Research, 44*, 160–181.

Barnett, S. B., White, H. M., & Cerf, M. (2016). Keep it simple stimuli: Brain-vetted elements of movie trailers predict opening weekend ticket sales. *Advances in Consumer Research, 44*, 285-289

Barrett, L. F., & Wager, T. D. (2006). The structure of emotion: Evidence from neuroimaging studies. *Current Directions in Psychological Science, 15*, 79-83.

Bartra, O., McGuire J. T., Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage, 76*, 412-427.

Basar, E., Basar-Eroglu, C., Karakas, S., & Schurmann, M. (1999) Are cognitive processes manifested in event-related gamma, alpha, theta and delta oscillations in the EEG? *Neuroscience Letters, 259*, 165–168.

Basar, E., Basar-Eroglu, C., Karakas, S., & Schurmann, M. (1999). Oscillatory brain theory: A new trend in neuroscience. *IEEE Engineering in Medicine and Biology, 18*, 56-66.

Baumgartner, H., Sujan, M., & Padgett, D. (1997). Patterns of Affective Reactions to Advertisements: The Integration of Moment-to-Moment Responses into Overall Judgments. *Journal of Marketing Research, 34*, 219–232.

Baumgartner, T., Esslen, M., & Jäncke, L. (2006). From emotion perception to emotion experience: Emotions evoked by pictures and classical music. *International Journal of Psychophysiology, 60*, 34-43.

Berger, H. (1929). Über das elektrenkephalogramm des menschen. *Archiv für Psychiatrie und Nervenkrankheiten, 87*, 527-570.

Berns, G. S., & Moore, S. E. (2012). A neural predictor of cultural popularity. *Journal of Consumer Psychology, 22*, 154–160.

Boksem, M. A. S., & Smidts, A. (2014). Brain responses to movie trailers predict individual preferences for movies and their population-wide commercial success. *Journal of Marketing Research, 52,* 482–492.

Brown, S. P., & Stayman, D. M. (1992). Antecedents and consequences of attitude toward the ad: A meta-analysis. *Journal of Consumer Research, 19,* 34-51.

Burke, M. C., & Edell, J. A. (1989). The impact of feelings on ad-based affect and cognition. *Journal of Marketing Research, 26*, 69-83.

Buzsaki, G., (2005). Theta rhythm of navigation: Link between path integration and landmark navigation, episodic and semantic memory. *Hippocampus, 15*, 827-840.

Chan, H. Y., Smidts, A., Schoots, V. C., Dietvorst, R. C., Boksem, M. A. S. (2019). Neural similarity at temporal lobe and cerebellum predicts out-of-sample preference and recall for video stimuli. *Neuroimage, 197*, 391-401.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin, 25*, 215-224.

Christoforou, C., Papadopoulos, T. C., Constantinidou, F., & Theodorou, M. (2017). Your brain on the movies: A computational approach for predicting box-office performance from viewer's brain responses to movie trailers. *Frontiers in Neuroinformatics, 11*, 1-13.

Coan, J. A., & Allen, J. J. B. (2004). Frontal EEG asymmetry as a moderator and mediator of emotion. *Biological Psychology, 67*, 7-49.

Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage, 35*, 968–978.

Coles, M. G. H, & Rugg, M. D. (1995). Event-related brain potentials: an introduction. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 1-26). Oxford: Oxford University Press.

Couwenberg, L. E., Boksem, M. A. S., Dietvorst, R. C., Worm, L., Verbeke, W. J. M. I., & Smidts, A. (2017). Neural responses to functional and experiential ad appeals: Explaining ad effectiveness. *International Journal of Research in Marketing, 34*, 355-366.

Davidson, R. J., & Fox, N. A. (1982). Asymmetrical brain activity discriminates between positive and negative affective stimuli in human infants. *Science, 218*, 1235-1237.

Davidson, R.J. (2004). What does the prefrontal cortex "do" in affect: perspectives on frontal EEG asymmetry research. *Biological Psychology, 67*, 219–233.

Davidson, R.J., & Fox, N., 1982. Asymmetric brain activity discriminates between positive and negative affective stimuli in 10 month old infants. *Science, 218*, 1235–1237.

DeVany, A., & Walls, W. D. (1999). Uncertainty in the movie industry: Does star power reduce the terror of the box office? *Journal of Cultural Economics, 23*, 285–318.

DeCesarei, A., & Codispoti, M. (2011). Affective modulation of the LPP and α-ERD during picture viewing. *Psychophysiology, 48,* 1397-1404.

Dholakia, U. M., & Morwitz, V. G. (2002). The scope and persistence of mere measurement effects: Evidence from a field study of customer satisfaction measurement. *Journal of Consumer Research, 29*, 159-167.

Dietterich, T. G., & Bakiri, G. (1995). Solving multiclass learning problems via error-correcting output codes. *Journal of Artificial Intelligence Research, 2*, 263-286.

Dmochowski, J. P., Bezdek, M. A., Abelson, B. P., Johnson, J. S., Schumacher, E. H., & Parra, L. C. (2014). Audience preferences are predicted by temporal reliability of neural processing. *Nature Communications, 5,* 1-9.

Do, A. M., Rupert, A. V., & Wolford, G. (2008). Evaluations of pleasurable experiences: The peak-and-end rule. *Psychonomic Bulletin & Review, 15*, 96-98.

Dockree, P. M., Kelly, S. P., Roche, R. A. P., Hogan, M. J., Reilly, R. B., & Robertson, I. H. (2004). Behavioral and physiological impairments of sustained attention after traumatic brain injury. *Brain Research and Cognitive Brain Research, 2*0, 403 414.

Dolan, R.J. (2002). Emotion, cognition, and behavior. *Science, 298*, 1191-1194.

Ebbinghaus, H. (1902). *Grundzuge der Psychologie*. Leipzig, Germany: Veit & Comp.

Edell, J. A., & Burke, M. C. (1987). The power of feelings in understanding advertising effects. *Journal of Consumer Research, 14,* 421-433.

Elster, J. (1998). Emotions and economic theory. *Journal of Economic Literature, 36*, 47-74.

Falk, E. B., Berkman, E. T., & Lieberman, M. D. (2012). From neural responses to population behavior: Neural focus group predicts population-level media effects. *Psychological Science, 23*, 439–445.

Falk, E. B., O'Donnell, M. B., Tompson, S., Gonzalez, R., Dal Cin, S., Strecher, V., Cummings, K. M., & An, L. (2016). Functional brain imaging predicts public health campaign success. *Social Cognitive and Affective Neuroscience, 11*, 204 -214.

Feldman, J. M., & Lynch, J. G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intentions, and behavior. *Journal of Applied Psychology, 73*, 421-435.

Fell, J., Klaver, P., Lehnertx, K., Grunwald, T., Schaller, C., Elger, C. E., & Fernándex, G. (2001). Human memory formation is accompanied by rhinal-hippocampal coupling and decoupling. *Nature Neuroscience, 4*, 1259-1264.

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research, 20*, 303-315.

Folkes, V., Koletsky, S., & Graham, J. (1987). A Field Study of Causal Inferences and Consumer Reaction: The View from the Airport. *Journal of Consumer Research, 13*, 534-539.

Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition and Emotion, 19*, 313-332.

Gazley, A., Clark, G., & Sinha, A. (2011). Understanding preferences for motion pictures. *Journal of Business Research, 64*, 854–61.

Genevsky, A., Yoon, C., & Knutson, B. (2017). When brain beats behavior: Neuroforecasting crowdfunding outcomes. *The Journal of Neuroscience, 37*, 8625-8634.

Goncharova, I. I., McFarland, D. J., Vaughan, T. M., & Wolpaw, J. R. (2003). EMG contamination of EEG: Spectral and topographical characteristics. *Clinical Neurophysiology, 114*, 1580-1593.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*, 1711–1725.

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition and Emotion, 9*, 87-108.

Gruber, T., Müller, M. M., Keil, A., & Elbert, T. (1999). Selective visual-spatial attention alters induced gamma band responses in the human EEG. *Clinical neurophysiology, 110*, 2074–2085.

Guixeres, J., Bigné, E., Azofra, J. M. A., Raya, M. A., Granero, A. C., Hurtado, F. F., & Ornedo, V. N. (2017). Consumer neuroscience-based metrics predict recall, liking and viewing rates in online advertising. *Frontiers in Psychology, 8*, 1–14.

Hakim, A., & Levy, D. J. (2018). A gateway to consumers' minds: Achievements, caveats, and prospects of electroencephalography-based prediction in neuromarketing. *WIREs Cognitive Science, 10*: e1485.

Hakim, A., Klorfeld, S., Sela, T., Friedman, D., Shabat-Simon, M., & Levy, D. J. (2018). Pathways to consumers minds: Using machine learning and multiple EEG metrics to increase preference prediction above and beyond traditional measurements. *bioRxiv,* 317073.

Harmon-Jones, E., Gable, P.A., & Peterson, C.K. (2010). The role of asymmetric frontal cortical activity in emotion-related phenomena: A review and update. *Biological Psychology, 84*, 451-462.

Harmon-Jones, E., Sigelman, J. D., Bohlig, A., & Harmon-Jones, C. (2003). Anger, coping, and frontal cortical activity: The effect of coping potential on anger-induced left frontal activity. *Cognition and Emotion, 17*, 1-24.

Herrmann, C. S., Munk, M. H. J., & Engel, A. K. (2004). Cognitive functions of gamma-band activity: Memory match and utilization. *Trends in Cognitive Sciences, 8*, 347–355.

Holbrook, M. B., & Batra, R. (1987). Assessing the role of emotions as mediators of consumer responses to advertising. *Journal of Consumer Research, 14,* 404-420.

Jensen, O. (2005). Reading the hippocampal code by theta phase-locking. *Trends in Cognitive Sciences, 9*, 551-553.

Jensen, O., & Lisman, J. E. (2005). Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. *Trends in Neurosciences, 28*, 67-72.

Kahneman, D., Fredrickson, D. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science, 4*, 401-405.

Kassam, K. S., Markey, A. R., Cherkassky, V. L., Loewenstein, G., & Just, M. A. (2013). Identifying emotions on the basis of neural activation. *PLoS ONE, 8*, e66032.

Kawasaki, M., & Yamaguchi, Y. (2013). Frontal theta and beta synchronizations for monetary reward increase visual working memory capacity. *Social Cognitive and Affective Neuroscience, 8*, 523–30.

Keil, A., Müller, M. M., Gruber, T., Wienbruch, C., Stolarova, M., & Elbert, T. (2001). Effects of emotional arousal in the cerebral hemispheres: a study of oscillatory brain activity and event-related potentials. *Clinical Neurophysiology, 112,* 2057–2068.

Keil, A., Müller, M. M., Ray, W. J., Gruber, T., & Elbert, T. (1999). Human gamma band activity and perception of a gestalt. *The Journal of Neuroscience, 19*, 7152-7161.

Kirk, I. J. & Mackay, J. C. (2003). The role of theta-range oscillations in synchronising and integrating activity in distributed mnemonic networks. *Cortex, 39*, 993-1008.

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences, 16,* 606–617.

Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., & Schwaiger, J. (1998). Induced alpha band power changes in the human EEG and attention. *Neuroscience Letters, 244*, 73-76.

Knutson, B, & Genevsky, A. (2018). Neuroforecasting aggregate choice. *Current Directions in Psychological Science, 27*, 110–115.

Knyazev, G. G. (2007). Motivation, emotion, and their inhibitory control mirrored in brain oscillations. *Neuroscience and Biobehavioral Reviews, 31*, 377-395.

Kotler, P., & Keller, K. L. (2007). *A framework for marketing management*. Upper Saddle River, New Jersey: Pearson Education, Inc.

Kover, A. J. (1995). Copywriters' implicit theories of communication: An exploration. *Journal of Consumer Research, 21*, 596-611.

Kragel, P. A., & LaBar, K. S. (2016). Decoding the nature of emotion in the brain. *Trends in Cognitive Sciences, 20*, 444-455.

Kühn, S., Strelow, E., & Gallinat, J. (2016). Multiple "buy buttons" in the brain: Forecasting chocolate sales at point-of-sale based on functional brain activation using fMRI. *NeuroImage, 136*, 122–128.

LaBarbera, P. A., & Tucciarone, J. D. (1995). GRS reconsidered: a behavior-based approach to evaluating and improving the sales potency of advertising. *Journal of Advertising Research, 35*, 33-53.

Labroo, A. A., Dhar, R., & Schwarz, N. (2008). Of frog wines and frowning watches: Semantic priming, perceptual fluency, and brand evaluation. *Journal of Consumer Research, 34*, 819–831.

Lang, P. J., & Bradley, M. M. (2010). Emotion and the motivational brain. *Biological psychology, 84*, 437–450.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. University of Florida, Gainesville, FL.

Laufs, H., Krakow, K., Sterzer, P., Eger, E., Beyerle, A., Salek-Haddadi, A., & Kleinschmidt, A. (2003). Electroencephalographic signatures of attentional and cognitive default modes in spontaneous brain activity fluctuations at rest. *Proceedings of the National Academy of Sciences, USA, 100*, 11053–11058.

Lee, A. Y., & Labroo, A. A. (2004). The Effect of conceptual and perceptual fluency on brand evaluation. *Journal of Marketing Research, 41*, 151-165.

Levallois, C., Smidts, A., & Wouters, P. (2019). The emergence of neuromarketing investigated through online public communications (2002–2008). *Business History*, DOI: 10.1080/00076791.2019.1579194

Lewinski, P., Fransen, M. L., & Tan, E. S. H. (2014). Predicting advertising effectiveness by facial expressions in response to amusing persuasive stimuli. *Journal of Neuroscience, Psychology, and Economics, 7*, 1-14.

Lin, Y. P., Wang, C. H., Jung, T. P., Wu, T. L., Jeng, S. K., Duann, J. R., & Chen, J. H. (2010). EEG-based emotion recognition in music listening. *IEEE Transactions on Biomedical Engineering, 57*, 1798-1806.

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L.F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences, 35*, 121-202.

Liu, X., Shi, S. W., Teixeira, T., & Wedel, M. (2018). Video content marketing: The making of clips. *Journal of Marketing, 82*, 86–101.

Loewenstein, G. (2000). Emotions in economic theory and economic behavior. *American Economic Review, 90*, 426-432.

Lucas, D. B. (1960). The ABCs of ARF's Parm. *Journal of Marketing, 25*, 9-20.

Lucchiari, C., & Pravettoni, G. (2012). The effect of brand on EEG modulation: A study on mineral water. *Swiss Journal of Psychology, 71*, 199-204.

Luck, S. J. (2005). An introduction to the event-related potential technique. Cambridge, MA: MIT press.

MacKenzie, S. B., Lutz, R. J., & Belch, G. E. (1986). The role of attitude toward the ad as a mediator of advertising effectiveness: A test of competing explanations. Journal of Marketing Research, 23, 130-143.

Mantini, D., Perrucci, M. G., Del Gratta, C., Romani, G. L., & Corbetta, M. (2007). Electrophysiological signatures of resting state networks in the human brain. Proceedings of the National Academy of Sciences, USA, 104, 13170-13175.

Marco-Pallares, J., Cucurell, D., Cunillera, T., García, R., Andrés-Pueyo, A., Münte, T. F., & Rodríguez-Fornells, A. (2008). Human oscillatory activity associated to reward processing in a gambling task. Neuropsychologia, 46, 241–8.

Morris, J. D., Woo, C., Geason, J. A., & Kim, J. (2002). The power of affect: Predicting intention. Journal of Advertising Research, 42, 7-17.

Müller, M. M., Gruber, T., & Keil, A. (2000). Modulation of induced gamma band activity in the human EEG by attention and visual information processing. International Journal of Psychophysiology, 38, 283-299.

Müller, M. M., Keil, A., Gruber, T., & Elbert, T. (1999). Processing of affective pictures modulates right-hemispheric gamma band EEG activity. Clinical Neurophysiology, 110, 1913-1920.

Murugappan, M., Nagarajan, R., & Yaacob, S. (2011). Combining spatial filtering and wavelet transform for classifying human emotions using EEG Signals. Journal of Medical and Biological Engineering, 31, 45-51.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.

Nishitani, N. (2003). Dynamics of cognitive processing in the human hippocampus by neuromagnetic and neurochemical assessments. Neuroimage, 20, 561–571.

Nyilasy, G., Canniford, R., & Kreshel, P. J. (2013). Ad agency professionals' mental models of advertising creativity. European Journal of Marketing, 47, 1691-1710.

Ohman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. Journal of Experimental Psychology, 130, 466-478.

Olney, T. J., Holbrook, M. B., & Batra, R. (1991). Consumer responses to advertising: The effects of ad content, emotions, and attitude toward the ad on viewing time. Journal of Consumer Research, 17, 440-453.

Percy, L. & Rossiter, J. R. (1992). A model of brand awareness and brand attitude advertising strategies. Psychology and Marketing, 9, 263-274.

Perria, L., Rosadini, G., & Rossi, G. F. (1961). Determination of side of cerebral dominance with amobarbital. *Archives of Neurology, 4*, 175-181.

Peters, E., Vastfjall, D., Garling, T., & Slovic, P. (2006). Affect and decision making: A "hot" topic. *Journal of Behavioral Decision Making, 19*, 79-85.

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology, 110*, 1842–1857.

Pham, M. T., Geuens, M., & De Pelsmacker, P. (2013). The influence of ad-evoked feelings on brand evaluations: Empirical generalizations from consumer responses to more than 1000 TV commercials. *International Journal of Research in Marketing, 30,* 383–394.

Phelps, E. A. (2004). Human emotion and memory: Interactions of the amygdala and hippocampal complex. *Current Opinion in Neurobiology, 14*, 198-202.

Plassmann, H., Venkatraman, V., Huettel, S., & Yoon, C. (2015). Consumer neuroscience: Applications, challenges, and possible solutions. *Journal of Marketing Research, 52*, 427-435.

Pochon, J. B., Levy, R., Fossati, P., Lehericy, S., Poline, J. B., Pillon, B., . . . Dubois, B. (2002). The neural system that bridges reward and cognition in humans: An fMRI study. *Proceedings of the National Academy of Sciences, USA, 99*, 5669–5674.

Poels, K., & DeWitte, S. (2006). How to capture the heart? Reviewing 20 years of emotion measurement in advertising. *Journal of Advertising Research, 46*, 18–37.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10,* 59–63.

Pradeep, A. K. (2010). *The Buying Brain: Secrets for Selling to the Subconscious Mind*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review, 8*, 364-382.

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science, 9*, 45–48.

Rodriguez, E., George, N., Lachaux, J.-P., Martinerie, J., Renault, B., & Varela, F. J. (1999). Perception's shadow: Long-distance synchronization of human brain activity. *Nature, 397*, 430-433.

Rugg, M. D., & Coles, M. G. H. (Eds.) (1995). Electrophysiology of mind: Event-related brain potentials and cognition. Oxford, England: Oxford University Press, (Chapter 1).

Schmitt, B. (1999). Experiential Marketing. *Journal of Marketing Management, 15*, 53-67

Scholz, C., Baek, E. C., O'Donnell, M. B., Kim, H. S., Cappella, J. N., & Falk, E. B. (2017). A neural model of valuation and information virality. *Proceedings of the National Academy of Sciences, USA, 114*, 2881–2886.

Shackman, A. J., & Wager, T. D. (2019) The emotional brain: Fundamental questions and strategies for future research. *Neuroscience Letters, 693*, 68-74.

Shestyuk, A. Y., Kasinathan, K., Karapoondinott, V., Knight, R. T., Gurumoorthy, R. (2019). Individual EEG measures of attention, memory, and motivation predict population level TV viewership and Twitter engagement. *PLoS ONE, 14*, e0214507.

Simons, R. F., Detenber, B. H., Cuthbert, B. N., Schwartz, D. D., & Reiss, J. E. (2003). Attention to television: Alpha power and its relationship to image motion and emotional content. *Media Psychology, 5,* 283-301.

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research, 177*, 1333-1352.

Smidts, A. (2002). Kijken in het brein: Over de mogelijkheden van neuromarketing. ERIM Inaugural Address Series Research in Management. Retrieved from http://hdl.handle.net/1765/308

Smidts, A., Hsu, M., Sanfey, A.G., Boksem, M.A.S., Ebstein, R.B., Huettel, S.A., Kable, J.W., Karmarkar, U.R., et al. (2014). Advancing consumer neuroscience. *Marketing Letters, 25*, 257–267.

Spotts, H. E., Weinberger, M. G., & Parsons, A. L. (1997) Assessing the use and impact of humor on advertising effectiveness: A contingency approach. *Journal of Advertising, 26*, 17-32,

Stahl, S. M. (2008). Stahl's essential psychopharmacology: Neuroscientific basis and practical applications. (3th ed.). New York: Cambridge University Press, (Chapter 1).

Stayman, D. M., & Aaker, D. A. (1988). Are all the effects of ad-induced feelings mediated by AAd? *Journal of Consumer Research, 15,* 368-373.

Stern, R. M., Ray, W. J., & Quigley, K. S. (2001). *Psychophysiological Recording*. New York: Oxford University Press.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1997). Oscillatory gamma band (30-70 Hz) activity induced by a visual search task in humans. *The Journal of Neuroscience, 17*, 722-734.

Teixeira, T., Wedel, M., & Pieters, R. (2010). Moment-to-moment optimal branding in TV-commercials: Preventing avoidance by pulsing. *Marketing Science, 2*9, 783-804.

Teixeira, T., Wedel, M., & Pieters, R. (2012). Emotion-Induced Engagement in Internet Video Advertisements. *Journal of Marketing Research, 49*, 144–159.

Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement Science Review, 2*, 1-11.

Tiitinen, H., Sinkkonen, J., Reinikainen, K., Alho, K., Lavikainen, J., & Naatanen, R. (1993). Selective attention enhances the auditory 40-Hz transient response in humans. *Nature, 364*, 59–60.

Uusberg, A., Uibo, H., Kreegipuu, K., & Allik, J. (2013). EEG alpha and cortical inhibition in affective attention. *International Journal of Psychophysiology, 89,* 26-36.

Van 't Wout, M., Chang, L. J., & Sanfey, A. G. (2010). The influence of emotion regulation on social interactive decision-making. *Emotion, 10*, 815-821.

Van de Vijver, I., Ridderinkhof, K. R., & Cohen, M. X. (2011). Frontal oscillatory dynamics predict feedback learning and action adjustment. *Journal of Cognitive Neuroscience, 23*, 4106-4121.

Van Honk, J., & Schutter, D. J. L. G. (2006). From affective valence to motivational direction: the frontal asymmetry of emotion revised. *Psychological Science, 17*, 963–965.

Vecchiato, G., Astolfi, L., De Vico Fallani, F., Cincotti, F., Mattia, D., Salinari, S, Soranzo, R., & Babiloni, F. (2010). Changes in brain activity during the observation of TV commercials by using EEG, GSR and HR measurements. *Brain Topography, 23*, 165-179.

Vecchiato, G., Toppi, J., Astolfi, L., Cincotti, F., De Vico Fallani, F., Maglione, A. G., Borghini, G., Cherubino, P., et al. (2012). The added value of the electrical neuroimaging for the evaluation of marketing stimuli. *Bulletin of the Polish Academy of Sciences: Technical Sciences, 60*, 419-426.

Vecchiato, G., Toppi, J., Astolfi, L., De Vico Fallani, F., Cincotti, F., Mattia, D., ... Babiloni, F. (2011). Spectral EEG frontal asymmetries correlate with the experienced pleasantness of TV commercial advertisements. *Medical and Biological Engineering and Computing, 49*, 579–583.

Venkatraman, V., Dimoka, A., Pavlou, P. A., Vo, K., Hampton, W., Bollinger, B., . . . Winer, R. S. (2015). Predicting advertising success beyond traditional measures: New insights from neurophysiological methods and market response modeling. *Journal of Marketing Research, 52*, 436–452.

Vertes, R. P. (2005). Hippocampal theta rhythm: A tag for short-term memory. *Hippocampus, 15*, 923–935.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: A meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience, 3*, 255-274.

Wang, Y. J., & Minor, M. S. (2008). Validity, reliability, and applicability of psychophysiological techniques in marketing research. *Psychology & Marketing, 25*, 197–232.

Wiles, J. A., & Cornwell, T. B. (1990). A review of methods utilized in measuring affect, feelings, and emotion in advertising. *Current Issues & Research in Advertising, 13*, 241–275.

Winkielman, P., & Berridge, K. C. (2004). Unconscious emotion. *Current Directions in Psychological Science, 13*, 120-123.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*, 151-175.

English summary

The application of neuroscience methods and insights to the field of marketing theory and practice, has increased in popularity over the past two decades. This dissertation extends existing knowledge by elucidating two proposed aims of neuromarketing, using EEG: offering additional insight into implicit processes (here, emotions) and contributing to predicting behavioral, market level, responses or ‘advertising effectiveness’.

Emotions are fundamental in guiding our behavior and they have been studied extensively in marketing. However, it has proved difficult to measure emotional experiences unobtrusively, particularly for dynamic stimuli. The first chapter therefore demonstrates a method that could provide insight on the moment-by-moment specific emotional effect that a marketing stimulus, such as a TV-commercial, has, on consumers. In the second chapter, the relationship between an a priori identified process (arousal) and external measures of ad effectiveness in the population at large (as measured by notability, attitude toward the ad, and choice), is investigated in one and the same study. The third chapter shows a systematic re-analysis of data from four studies in which neural activity in response to a similar stimulus (here, movie trailers) was investigated using EEG to examine the association with population-wide commercial success of the movies.

In addition to the substantive findings, this dissertation also contributes methodologically to the neuromarketing field by i) applying novel multivariate methods to decode emotional experiences, ii) using a localizer task in an EEG study to reduce the reverse inference problem that commonly plagues neuroimaging research, and iii) conducting a major meta-analysis to address the issue of small samples sizes regarding both participants and stimuli in neuromarketing research.

Dutch summary/ Nederlandse samenvatting

Marketing is gericht op het identificeren van en tegemoetkomen aan behoeften die mensen hebben (Kotler & Keller, 2007). Hierbij worden nieuwe producten en diensten ontwikkeld die klanten waarderen, en passende communicatie in de vorm van reclame ontworpen om consumenten te informeren en te overtuigen. Hoe ziet een goed ontworpen film trailer eruit die zodanig de interesse wekt bij mensen dat ze sterk overwegen de desbetreffende film te gaan zien? Om deze ontwerpen aan te kunnen laten sluiten op de behoeften van mensen, is het belangrijk de consument goed te begrijpen. Met onder andere dit doel voor ogen, is de afgelopen jaren het toepassen van methoden en inzichten uit de neurowetenschappen in marketing, in populariteit toegenomen. Naar deze toepassingen wordt gerefereerd met de term *neuro*marketing, die dateert uit het jaar 2002 (Levallois, Smidts, & Wouters, 2019).

Elektro-encefalografie (EEG) is één van de methoden uit de neurowetenschappen die wordt toegepast in marketing. Hierbij wordt hersenactiviteit gemeten door middel van electroden die gemonteerd zijn in een elastische cap op het hoofd. De hersenen bestaan uit een netwerk van miljarden hersencellen (neuronen) die met elkaar communiceren en met EEG kun je iets van deze hersenactiviteit meten aan de schedeloppervlakte in de vorm van elektrische voltageverschillen. Deze methode wordt vooral toegepast in neuromarketing in de praktijk, omdat het relatief goedkoop is en een goede temporele resolutie heeft, wat goed van pas komt bij het analyseren van de hersenreactie op dynamische marketing stimuli zoals tv-reclames of films.

In dit proefschrift onderzoeken we hoe de hersenen reageren op marketing gerelateerde stimuli met behulp van EEG en in hoeverre die reactie gerelateerd is aan de effectiviteit van deze stimuli in de markt. We vergroten daarmee bestaande kennis door twee doelen die neuromarketing heeft, te verhelderen. Het eerste doel betreft het verschaffen van extra inzicht in psychologische processen die ten grondslag liggen aan beslissingen van de consument. De focus in dit proefschrift is hierbij gericht op het meten van emoties. Het tweede doel is bijdragen aan het voorspellen van de effectiviteit van marketing gerelateerde stimuli, zoals reclames.

Hoewel veel onderzoek heeft aangetoond dat emoties een grote invloed hebben op cognitie en gedrag, blijkt het moeilijk deze emotionele ervaringen te meten, vooral over de tijd. In Hoofdstuk 2 focussen we ons op het meten van verschillende emotionele ervaringen met behulp van EEG, door het toepassen van een algoritme dat voor dit doel gebruik maakt van patronen in de data. Voor het oproepen van de emotionele ervaringen hebben we proefpersonen audiovisuele stimuli (video's) getoond, die speciaal waren samengesteld om respectievelijk, blijdschap, verdriet, angst of afkeer op te wekken. Onze resultaten laten zien dat we met het algoritme, op basis van het EEG-signaal, goed kunnen onderscheiden welke emoties worden ervaren tijdens het bekijken van de video's. Op basis van het hersensignaal kunnen we met redelijke zekerheid vaststellen of mensen blij zijn of juist verdrietig. Met onze methode benutten we de frequenties en topografie van het EEG-signaal die de emoties onderscheidden, zodat we de verschillen tussen emoties konden interpreteren in termen van psychologische processen die hier mogelijk aan ten grondslag liggen, zoals aandacht of arousal (de intensiteit van een emotionele respons). Tenslotte illustreren we hoe deze methode toegepast kan worden om van seconde tot seconde veranderingen in de emotionele ervaring te meten, door een analyse te maken van de emotionele reacties op de video *Up*. In de neuromarketingpraktijk kan het toepassen van dit algoritme erg nuttig zijn om consumenten ervaringen preciezer te meten door bijvoorbeeld te kunnen bepalen welke scenes van een filmtrailer of advertentie veel blijdschap of juist afkeer oproepen, en welke scenes weinig reactie veroorzaken en dus verwijderd zouden kunnen worden.

In Hoofdstuk 3 kijken we niet alleen naar de kwaliteit van de marketing stimulus zelf, maar ook naar hoe het EEG-signaal gerelateerd is aan het effect van de stimulus. We onderzoeken de mate waarin de emotionele reactie op advertenties is geassocieerd met de evaluatie van en het gedrag in reactie op de advertenties, op marktniveau. De nadruk ligt hierbij op het meten van 'arousal', een belangrijke component van gevoelens die opgeroepen kunnen worden door reclame. Met arousal bedoelen we de intensiteit van een (emotionele) respons. We houden bij dit onderzoek bovendien rekening met het probleem van *reverse inference* ('omgekeerde afleiding') bij het analyseren van de EEG-data. Het *reverse inference* probleem houdt in dat hoewel het opwekken van een hoog arousal niveau (bijvoorbeeld door mensen plaatjes van spinnen of slangen te laten zien) kan resulteren in het meten van verminderde activiteit in de alpha band frequentie (8-12 Hz) (i.e., *forward inference* of 'voorwaarts afleiden'), daarom niet per definitie ook het omgekeerde hoeft te gelden, namelijk dat het meten van verminderde activiteit in de alpha band frequentie simpelweg betekent dat arousal hoog was. Dit komt doordat ook andere psychologische processen (zoals aandacht en algemene alertheid) kunnen resulteren in verminderde alpha activiteit en tevens dat een verhoogd niveau van arousal ook kan leiden tot andere veranderingen in hersenactiviteit (bijvoorbeeld verhoogde activiteit in een ander frequentiegebied dan alpha).

We verkleinen het probleem van *reverse inference* door eerst te bekijken hoe arousal wordt gerepresenteerd in het brein met behulp van EEG in een *aparte taak*, en door vervolgens deze representatie te gebruiken om de arousal respons in reactie op advertenties te meten. Hiermee onderzoeken we in één dezelfde studie wat de relatie is tussen een vooraf bepaald proces (hier, arousal) gemeten in een kleine groep mensen, en externe maten van effectiviteit van de advertenties bij andere mensen. Effectiviteit van de advertenties is op marktniveau gemeten met opvallendheid van de advertentie, de attitude t.o.v. de advertentie, en gedrag (doorklikken naar de product website). De resultaten laten zien dat arousal gemeten met EEG in reactie op de advertenties inderdaad positief samenhangt met de opvallendheid van de advertenties, maar negatief met de houding ten opzichte van, en gedrag in reactie op de advertenties. De bevindingen bieden hierbij ondersteuning voor hoe met EEG arousal het beste gemeten kan worden in de neuromarketingpraktijk, maar laten ook zien dat opgepast moet worden met advies over de effectiviteit van een advertentie op basis van deze maat. Meer arousing en opvallende reclames worden niet noodzakelijkerwijs ook positiever beoordeeld.

In het vierde hoofdstuk richten we ons ook op de bijdrage van EEG-metingen aan het voorspellen van de effectiviteit van marketing gerelateerde stimuli, op marktniveau. We kijken hierbij specifiek naar de hersenactiviteit gemeten met EEG in reactie op filmtrailers en naar in hoeverre deze activiteit gerelateerd is aan het marktsucces van de desbetreffende films. Hiervoor bekijken we of de voorspellende effecten die zijn gevonden in één enkele reeds gepubliceerde studie (Boksem en Smidts, 2015), ook zijn te vinden in een veel grotere steekproef van proefpersonen en stimuli. We combineren data uit vier studies, zodat we een totale dataset analyseren van $n = 130$ proefpersonen en $k = 145$ stimuli (films). Met een systematische her-analyse hebben we getest of vijf EEG-maten van frequentiegebieden die vaak zijn onderzocht (theta, alpha, alpha asymmetrie, beta en gamma hersenactiviteit), bijdragen aan het voorspellen van het succes van de films. Belangrijk daarbij is dat we bekijken of de EEG-maten van toegevoegde waarde zijn náást film genre en de door de proefpersonen zelf gerapporteerde waardering van de films. De EEG-maten bleken inderdaad van toegevoegde waarde te zijn, maar opmerkelijk was dat alleen het effect van *gamma* activiteit significant was. Deze resultaten suggereren dat activiteit in de gamma frequentieband, gemeten in reactie op de trailers, unieke informatie bevat die bijdraagt aan het verklaren van het succes van de films. We moeten voorzichtig zijn met het interpreteren van deze bevindingen (zie het eerdergenoemde *reverse inference* probleem), maar we zouden kunnen speculeren dat dit effect van gamma mogelijk te maken heeft met het vasthouden van aandacht en verdere integratie van informatie. Ongeacht de verklaring, zou dit betekenen dat er in de praktijk gericht naar activiteit in de gamma band gekeken kan worden om trailers te optimaliseren m.b.t. de opbrengst van de film en om dus effectievere trailers te creëren.

In aanvulling op bovenstaande inhoudelijke bevindingen, levert deze dissertatie ook een aantal methodologische bijdragen aan het gebied van neuromarketing door: i) een nieuwe multivariate methode toe te passen om emotionele ervaringen te decoderen, ii) een lokalisatie-taak te gebruiken om het probleem van omgekeerde afleiding te verkleinen, en iii) een meta-analyse uit te voeren om het probleem van kleine steekproeven in neuromarketing onderzoek aan te pakken. Al met al draagt deze dissertatie bij aan de groeiende hoeveelheid onderzoek die laat zien dat het gebruik van EEG in reactie op marketing stimuli nuttig kan zijn voor marketing doeleinden. Het interdisciplinaire gebied van neuromarketing is echter nog relatief nieuw en volop in ontwikkeling. Deze dissertatie is daarmee een poging om het fundament te versterken waarop de praktijk en wetenschap samen verder kunnen bouwen.

About the author

Esther Eijlers was born in Vlissingen (The Netherlands) on November 10th, 1987. She obtained her Bachelor's and Master's degree in Psychology (Brain & Cognition, cum laude) from the Erasmus University Rotterdam. Prior to starting her PhD, she worked as a research assistant under the supervision of dr. Maarten Boksem and prof. Ale Smidts at the department of Marketing Management, at the Erasmus Research Institute of Management. With her research lying at the intersection of marketing and psychology, she uses electroencephalography (EEG) to investigate the brain's response elicited by marketing related stimuli, and explores the role of underlying emotional processes associated with evaluating advertisements. She has presented her work at several international conferences, and published one of the dissertation chapters in Public Library of Science (PLoS) One.

Portfolio

Education

- M.Sc. in Psychology - Brain & Cognition 2009-2010
  Cum Laude/ Award for best master thesis
  Erasmus University Rotterdam
- B.Sc. in Psychology 2006-2009
  Erasmus University Rotterdam

Publications

- Eijlers, E., Smidts, A., & Boksem, M.A.S. (2019). Implicit measurement of emotional experience and its dynamics. *PLoS One*, *14*: e0211496.

Work in progress

- Arousal and advertising success: Neural measures suggest that arousing ads stand out more but are liked less (*submitted for review*)
- EEG metrics relating to population-wide commercial success of movies: A meta-analysis

Conference presentations

- Identifying emotions with EEG (2018)
  European Society for Cognitive and Affective Neuroscience Conference, Leiden, The Netherlands
- Neural measures suggest that arousing ads stand out more but are liked less (2017)
  NeuroPsychoEconomics Conference, Antwerp, Belgium
- Neural measures of evoked emotions in predicting advertising effectiveness (2016)
  Society for NeuroEconomics Conference, Berlin, Germany
- The value of neural measures in evaluating the quality of advertisements (2014)
  Society for NeuroEconomics Conference, Miami, Florida
- The value of neural measures in evaluating the quality of advertisements (2014)
  Tilburg Institute for Behavioral Economics Research Symposium, Tilburg, The Netherlands

Doctoral teaching experience

Instructor Neuromarketing elective - EEG recording and analysis practicals 2014-2018

Bachelor internship coordinator (Dept. Marketing Management) 2017-2018

Instructor Research Training & Bachelor Thesis course 2016-2017

Doctoral coursework

Skill Courses *(Erasmus Research Institute of Management)*

- English (CPE)
- Scientific Integrity
- Publishing Strategy

Core Courses on Research Methodology *(Erasmus Research Institute of Management)*

- Statistical Methods
- Programming

Advanced Specialization or Methodology Courses

- Introduction to Data Analysis with R
- Multilevel Analysis in SPSS
- Generating and Implementing Interesting Research Ideas
- Experimental Methods in Business Research
- Topics in Consumer Behavior: Advances in Consumer Neuroscience
  Erasmus Research Institute of Management, Rotterdam, The Netherlands

- EDEN Doctoral Seminar on Consumer Research
  European Institute for Advanced Studies in Management, Brussels, Belgium
- Programming in Presentation
  Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands
- Neuroanatomy
  Donders Graduate School for Cognitive Neuroscience, Radboud University Nijmegen, The Netherlands
- The tool-kit of cognitive neuroscience 2014: advanced data analysis and source modelling of EEG and MEG data
  Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands

ERIM PhD Series

The ERIM PhD Series contains PhD dissertations in the field of Research in Management defended at Erasmus University Rotterdam and supervised by senior researchers affiliated to the Erasmus Research Institute of Management (ERIM). All dissertations in the ERIM PhD Series are available in full text through the ERIM Electronic Series Portal: http://repub.eur.nl/pub. ERIM is the joint research institute of the Rotterdam School of Management (RSM) and the Erasmus School of Economics (ESE) at the Erasmus University Rotterdam (EUR).

Dissertations in the last four years

Ahmadi, S., A motivational perspective to decision-making and behavior in organizations, Promotors: Prof. J.J.P. Jansen & Prof. T.J.M. Mom, EPS-2019-477-S&E, https://repub.eur.nl/pub/116727

Akemu, O., Corporate Responses to Social Issues: Essays in Social Entrepreneurship and Corporate Social Responsibility, Promotors: Prof. G.M. Whiteman & Dr. S.P. Kennedy, EPS-2017-392-ORG, https://repub.eur.nl/pub/95768

Albuquerque de Sousa, J.A., International stock markets: Essays on the determinants and consequences of financial market development, Promotors: Prof. M.A. van Dijk & Prof. P.A.G. van Bergeijk, EPS-2019-465-F&A, https://repub.eur.nl/pub/115988

Alexiou, A., Management of Emerging Technologies and the Learning Organization: Lessons from the Cloud and Serious Games Technology, Promotors: Prof. S.J. Magala, Prof. M.C. Schippers & Dr. I. Oshri, EPS-2016-404-ORG, https://repub.eur.nl/pub/93818

Alserda, G.A.G., Choices in Pension Management, Promotors: Prof. S.G. van der Lecq & Dr. O.W. Steenbeek, EPS-2017-432-F&A, https://repub.eur.nl/pub/103496

Arampatzi, E., Subjective Well-Being in Times of Crises: Evidence on the Wider Impact of Economic Crises and Turmoil on Subjective Well-Being, Promotors: Prof. H.R. Commandeur, Prof. F. van Oort & Dr. M.J. Burger, EPS-2018-459-S&E, https://repub.eur.nl/pub/111830

Avci, E., Surveillance of Complex Auction Markets: a Market Policy Analytics Approach, Promotors: Prof. W. Ketter, Prof. H.W.G.M. van Heck & Prof. D.W. Bunn, EPS-2018-426-LIS, https://repub.eur.nl/pub/106286

Balen, T.H. van, Challenges of Early Stage Entrepreneurs : the Roles of Vision Communication and Team Membership Change, Promotors: Prof. J.C.M van den Ende & Dr. M. Tarakci, EPS-2018-468-LIS, https://repub.eur.nl/pub/115654

Benschop, N, Biases in Project Escalation: Names, frames & construal levels, Promotors: Prof. K.I.M. Rhode, Prof. H.R. Commandeur, Prof. M. Keil & Dr. A.L.P. Nuijten, EPS-2015-375-S&E, https://repub.eur.nl/pub/79408

Bernoster, I., Essays at the Intersection of Psychology, Biology, and Entrepreneurship, Promotors: Prof. A.R. Thurik, Prof. I.H.A. Franken & Prof. P.J.F Groenen, EPS-2018-463-S&E, https://repub.eur.nl/pub/113907

Beusichem, H.C. van, Firms and Financial Markets: Empirical Studies on the Informational Value of Dividends, Governance and Financial Reporting, Promotors: Prof. A. de Jong & Dr. G. Westerhuis, EPS-2016-378-F&A, https://repub.eur.nl/pub/93079

Bouman, P., Passengers, Crowding and Complexity: Models for Passenger Oriented Public Transport, Promotors: Prof. L.G. Kroon, Prof. A. Schöbel & Prof. P.H.M. Vervest, EPS-2017-420-LIS, https://repub.eur.nl/pub/100767

Bunderen, L. van, Tug-of-War: Why and when teams get embroiled in power struggles, Promotors: Prof. D.L. van Knippenberg & Dr. L. Greer, EPS-2018-446-ORG, https://repub.eur.nl/pub/105346

Burg, G.J.J. van den, Algorithms for Multiclass Classification and Regularized Regression, Promotors: Prof. P.J.F. Groenen & Dr. A. Alfons, EPS-2018-442-MKT, https://repub.eur.nl/pub/103929

Chammas, G., Portfolio concentration, Promotor: Prof. J. Spronk, EPS-2017-410-F&E, https://repub.eur.nl/pub/94975

Consiglio, I., Others: Essays on Interpersonal and Consumer Behavior, Promotor: Prof. S.M.J. van Osselaer, EPS-2016-366-MKT, https://repub.eur.nl/pub/79820

Cranenburgh, K.C. van, Money or Ethics: Multinational corporations and religious organisations operating in an era of corporate responsibility, Promotors: Prof. L.C.P.M. Meijs, Prof. R.J.M. van Tulder & Dr. D. Arenas, EPS-2016-385-ORG, https://repub.eur.nl/pub/93104

Darnihamedani, P., Individual Characteristics, Contextual Factors and Entrepreneurial Behavior, Promotors: Prof. A.R. Thurik & S.J.A. Hessels, EPS-2016-360-S&E, https://repub.eur.nl/pub/93280

Dennerlein, T., Empowering Leadership and Employees' Achievement Motivations: the Role of Self-Efficacy and Goal Orientations in the Empowering Leadership Process, Promotors: Prof. D.L. van Knippenberg & Dr. J. Dietz, EPS-2017-414-ORG, https://repub.eur.nl/pub/98438

Depecik, B.E., Revitalizing brands and brand: Essays on Brand and Brand Portfolio Management Strategies, Promotors: Prof. G.H. van Bruggen, Dr. Y.M. van Everdingen and Dr. M.B. Ataman, EPS-2016-406-MKT, https://repub.eur.nl/pub/93507

Duijzer, L.E., Mathematical Optimization in Vaccine Allocation, Promotors: Prof. R. Dekker & Dr. W.L. van Jaarsveld, EPS-2017-430-LIS, https://repub.eur.nl/pub/101487

Duyvesteyn, J.G., Empirical Studies on Sovereign Fixed Income Markets, Promotors: Prof. P. Verwijmeren & Prof. M.P.E. Martens, EPS-2015-361-F&A, https://repub.eur.nl/pub/79033

El Nayal, O.S.A.N., Firms and the State: An Examination of Corporate Political Activity and the Business-Government Interface, Promotor: Prof. J. van Oosterhout & Dr. M. van Essen, EPS-2018-469-S&E, https://repub.eur.nl/pub/114683

Elemes, A., Studies on Determinants and Consequences of Financial Reporting Quality, Promotor: Prof. E. Peek, EPS-2015-354-F&A, https://repub.eur.nl/pub/79037

Erlemann, C., Gender and Leadership Aspiration: The Impact of the Organizational Environment, Promotor: Prof. D.L. van Knippenberg, EPS-2016-376-ORG, https://repub.eur.nl/pub/79409

Faber, N., Structuring Warehouse Management, Promotors: Prof. M.B.M. de Koster & Prof. A. Smidts, EPS-2015-336-LIS, https://repub.eur.nl/pub/78603

Feng, Y., The Effectiveness of Corporate Governance Mechanisms and Leadership Structure: Impacts on strategic change and firm performance, Promotors: Prof. F.A.J. van den Bosch, Prof. H.W. Volberda & Dr. J.S. Sidhu, EPS-2017-389-S&E, https://repub.eur.nl/pub/98470

Fernald, K., The Waves of Biotechnological Innovation in Medicine: Interfirm Cooperation Effects and a Venture Capital Perspective, Promotors: Prof. E. Claassen, Prof. H.P.G. Pennings & Prof. H.R. Commandeur, EPS-2015-371-S&E, https://repub.eur.nl/pub/79120

Fisch, C.O., Patents and trademarks: Motivations, antecedents, and value in industrialized and emerging markets, Promotors: Prof. J.H. Block, Prof. H.P.G. Pennings & Prof. A.R. Thurik, EPS-2016-397-S&E, https://repub.eur.nl/pub/94036

Fliers, P.T., Essays on Financing and Performance: The role of firms, banks and board, Promotors: Prof. A. de Jong & Prof. P.G.J. Roosenboom, EPS-2016-388-F&A, https://repub.eur.nl/pub/93019

Frick, T.W., The Implications of Advertising Personalization for Firms, Consumer, and Ad Platfroms, Promotors: Prof. T. Li & Prof. H.W.G.M. van Heck, EPS-2018-452-LIS, https://repub.eur.nl/pub/110314

Fytraki, A.T., Behavioral Effects in Consumer Evaluations of Recommendation Systems, Promotors: Prof. B.G.C. Dellaert & Prof. T. Li, EPS-2018-427-MKT, https://repub.eur.nl/pub/110457

Gaast, J.P. van der, Stochastic Models for Order Picking Systems, Promotors: Prof. M.B.M de Koster & Prof. I.J.B.F. Adan, EPS-2016-398-LIS, https://repub.eur.nl/pub/93222

Ghazizadeh, P., Empirical Studies on the Role of Financial Information in Asset and Capital Markets, Promotors: Prof. A. de Jong & Prof. E. Peek. EPS-2019-470-F&A https://repub.eur.nl/pub/114023

Giurge, L., A Test of Time; A temporal and dynamic approach to power and ethics, Promotors: Prof. M.H. van Dijke & Prof. D. De Cremer, EPS-2017-412-ORG, https://repub.eur.nl/pub/98451

Gobena, L., Towards Integrating Antecedents of Voluntary Tax Compliance, Promotors: Prof. M.H. van Dijke & Dr. P. Verboon, EPS-2017-436-ORG, https://repub.eur.nl/pub/103276

Groot, W.A., Assessing Asset Pricing Anomalies, Promotors: Prof. M.J.C.M. Verbeek & Prof. J.H. van Binsbergen, EPS-2017-437-F&A, https://repub.eur.nl/pub/103490

Hanselaar, R.M., Raising Capital: On pricing, liquidity and incentives, Promotors: Prof. M.A. van Dijk & Prof. P.G.J. Roosenboom, EPS-2018-429-F&A-9789058925404, https://repub.eur.nl/pub/113274

Harms, J. A., Essays on the Behavioral Economics of Social Preferences and Bounded Rationality, Promotors: Prof. H.R. Commandeur & Dr. K.E.H. Maas, EPS-2018-457-S&E, https://repub.eur.nl/pub/108831

Hekimoglu, M., Spare Parts Management of Aging Capital Products, Promotor: Prof. R. Dekker, EPS-2015-368-LIS, https://repub.eur.nl/pub/79092

Hendriks, G., Multinational Enterprises and Limits to International Growth: Links between Domestic and Foreign Activities in a Firm's Portfolio, Promotors: Prof. P.P.M.A.R. Heugens & Dr. A.H.L Slangen, EPS-2019-464-S&E, https://repub.eur.nl/pub/114981

Hengelaar, G.A., The Proactive Incumbent: Holy grail or hidden gem? Investigating whether the Dutch electricity sector can overcome the incumbent's curse and lead the sustainability transition, Promotors: Prof. R.J. M. van Tulder & Dr. K. Dittrich, EPS-2018-438-ORG, https://repub.eur.nl/pub/102953

Hogenboom, A.C., Sentiment Analysis of Text Guided by Semantics and Structure, Promotors: Prof. U. Kaymak & Prof. F.M.G. de Jong, EPS-2015-369-LIS, https://repub.eur.nl/pub/79034

Hollen, R.M.A., Exploratory Studies into Strategies to Enhance Innovation-Driven International Competitiveness in a Port Context: Toward Ambidextrous Ports, Promotors: Prof. F.A.J. Van Den Bosch & Prof. H.W. Volberda, EPS-2015-372-S&E, https://repub.eur.nl/pub/78881

Jacobs, B.J.D., Marketing Analytics for High-Dimensional Assortments, Promotors: Prof. A.C.D. Donkers & Prof. D. Fok, EPS-2017-445-MKT, https://repub.eur.nl/pub/103497

Jia, F., The Value of Happiness in Entrepreneurship, Promotors: Prof. D.L. van Knippenberg & Dr. Y. Zhang, EPS-2019-479-ORG, https://repub.eur.nl/pub/115990

Kahlen, M. T., Virtual Power Plants of Electric Vehicles in Sustainable Smart Electricity Markets, Promotors: Prof. W. Ketter & Prof. A. Gupta, EPS-2017-431-LIS, https://repub.eur.nl/pub/100844

Kampen, S. van, The Cross-sectional and Time-series Dynamics of Corporate Finance: Empirical evidence from financially constrained firms, Promotors: Prof. L. Norden & Prof. P.G.J. Roosenboom, EPS-2018-440-F&A, https://repub.eur.nl/pub/105245

Karali, E., Investigating Routines and Dynamic Capabilities for Change and Innovation, Promotors: Prof. H.W. Volberda, Prof. H.R. Commandeur & Dr. J.S. Sidhu, EPS-2018-454-S&E, https://repub.eur.nl/pub/106274

Keko, E., Essays on Innovation Generation in Incumbent Firms, Promotors: Prof. S. Stremersch & Dr. N.M.A. Camacho, EPS-2017-419-MKT, https://repub.eur.nl/pub/100841

Kerkkamp, R.B.O., Optimisation Models for Supply Chain Coordination under Information Asymmetry, Promotors: Prof. A.P.M. Wagelmans & Dr. W. van den Heuvel, EPS-2018-462-LIS, https://repub.eur.nl/pub/109770

Khattab, J., Make Minorities Great Again: a contribution to workplace equity by identifying and addressing constraints and privileges, Promotors: Prof. D.L. van Knippenberg & Dr. A. Nederveen Pieterse, EPS-2017-421-ORG, https://repub.eur.nl/pub/99311

Kim, T. Y., Data-driven Warehouse Management in Global Supply Chains, Promotors: Prof. R. Dekker & Dr. C. Heij, EPS-2018-449-LIS, https://repub.eur.nl/pub/109103

Klitsie, E.J., Strategic Renewal in Institutional Contexts: The paradox of embedded agency, Promotors: Prof. H.W. Volberda & Dr. S. Ansari, EPS-2018-444-S&E, https://repub.eur.nl/pub/106275

Kong, L., Essays on Financial Coordination, Promotors: Prof. M.J.C.M. Verbeek, Dr. D.G.J. Bongaerts & Dr. M.A. van Achter. EPS-2019-433-F&A, https://repub.eur.nl/pub/114516

Koolen, D., Market Risks and Strategies in Power Systems Integrating Renewable Energy, Promotors: Prof. W. Ketter & Dr. R. Huisman, EPS-2019-467-LIS, https://repub.eur.nl/pub/115655

Krämer, R., A license to mine? Community organizing against multinational corporations, Promotors: Prof. R.J.M. van Tulder & Prof. G.M. Whiteman, EPS-2016-383-ORG, https://repub.eur.nl/pub/94072

Kyosev, G.S., Essays on Factor Investing, Promotors: Prof. M.J.C.M. Verbeek & Dr. J.J. Huij, EPS-2019-474-F&A, https://repub.eur.nl/pub/116463

Lamballais, T., Optimizing the Performance of Robotic Mobile Fulfillment Systems, Promotors: Prof. M.B.M de Koster & Prof. R. Dekker & Dr. D. Roy, EPS-2019-411-LIS, https://repub.eur.nl/pub/116477

Lee, C.I.S.G., Big Data in Management Research: Exploring New Avenues, Promotors: Prof. S.J. Magala & Dr. W.A. Felps, EPS-2016-365-ORG, https://repub.eur.nl/pub/79818

Legault-Tremblay, P.O., Corporate Governance During Market Transition: Heterogeneous responses to Institution Tensions in China, Promotor: Prof. B. Krug, EPS-2015-362-ORG, https://repub.eur.nl/pub/78649

Lenoir, A.S., Are You Talking to Me? Addressing Consumers in a Globalised World, Promotors: Prof. S. Puntoni & Prof. S.M.J. van Osselaer, EPS-2015-363-MKT, https://repub.eur.nl/pub/79036

Leung, W.L., How Technology Shapes Consumption: Implications for Identity and Judgement, Promotors: Prof. S. Puntoni & Dr. G. Paolacci, EPS-2019-485-MKT, https://repub.eur.nl/pub/117432

Li, D., Supply Chain Contracting for After-sales Service and Product Support, Promotor: Prof. M.B.M. de Koster, EPS-2015-347-LIS, https://repub.eur.nl/pub/78526

Li, X., Dynamic Decision Making under Supply Chain Competition, Promotors: Prof. M.B.M de Koster, Prof. R. Dekker & Prof. R. Zuidwijk. EPS-2018-466-LIS, https://repub.eur.nl/pub/114028

Liu, N., Behavioral Biases in Interpersonal Contexts, Promotors: Prof. A. Baillon & Prof. H. Bleichrodt, EPS-2017-408-MKT, https://repub.eur.nl/pub/95487

Ma, Y., The Use of Advanced Transportation Monitoring Data for Official Statistics, Promotors: Prof. L.G. Kroon & Dr. J. van Dalen, EPS-2016-391-LIS, https://repub.eur.nl/pub/80174

Maas, A.J.J., Organizations and their external context: Impressions across time and space, Promotors: Prof. P.P.M.A.R. Heugens & Prof. T.H. Reus, EPS-2019-478-S&E, https://repub.eur.nl/pub/116480

Maira, E., Consumers and Producers, Promotors: Prof. S. Puntoni & Prof. C. Fuchs, EPS-2018-439-MKT, https://repub.eur.nl/pub/104387

Mell, J.N., Connecting Minds: On The Role of Metaknowledge in Knowledge Coordination, Promotor: Prof. D.L. van Knippenberg, EPS-2015-359-ORG, https://repub.eur.nl/pub/78951

Meulen, D. van der, The Distance Dilemma: the effect of flexible working practices on performance in the digital workplace, Promotors: Prof. H.W.G.M. van Heck & Prof. P.J. van Baalen, EPS-2016-403-LIS, https://repub.eur.nl/pub/94033

Moniz, A, Textual Analysis of Intangible Information, Promotors: Prof. C.B.M. van Riel, Prof. F.M.G de Jong & Dr. G.A.J.M. Berens, EPS-2016-393-ORG, https://repub.eur.nl/pub/93001

Mulder, J., Network design and robust scheduling in liner shipping, Promotors: Prof. R. Dekker & Dr. W.L. van Jaarsveld, EPS-2016-384-LIS, https://repub.eur.nl/pub/80258

Neerijnen, P., The Adaptive Organization: the socio-cognitive antecedents of ambidexterity and individual exploration, Promotors: Prof. J.J.P. Jansen, P.P.M.A.R. Heugens & Dr. T.J.M. Mom, EPS-2016-358-S&E, https://repub.eur.nl/pub/93274

Okbay, A., Essays on Genetics and the Social Sciences, Promotors: Prof. A.R. Thurik, Prof. Ph.D. Koellinger & Prof. P.J.F. Groenen, EPS-2017-413-S&E, https://repub.eur.nl/pub/95489

Oord, J.A. van, Essays on Momentum Strategies in Finance, Promotor: Prof. H.K. van Dijk, EPS-2016-380-F&A, https://repub.eur.nl/pub/80036

Peng, X., Innovation, Member Sorting, and Evaluation of Agricultural Cooperatives, Promotor: Prof. G.W.J. Hendriks, EPS-2017-409-ORG, https://repub.eur.nl/pub/94976

Pennings, C.L.P., Advancements in Demand Forecasting: Methods and Behavior, Promotors: Prof. L.G. Kroon, Prof. H.W.G.M. van Heck & Dr. J. van Dalen, EPS-2016-400-LIS, https://repub.eur.nl/pub/94039

Petruchenya, A., Essays on Cooperatives: Emergence, Retained Earnings, and Market Shares, Promotors: Prof. G.W.J. Hendriks & Dr. Y. Zhang, EPS-2018-447-ORG, https://repub.eur.nl/pub/105243

Plessis, C. du, Influencers: The Role of Social Influence in Marketing, Promotors: Prof. S. Puntoni & Prof. S.T.L.R. Sweldens, EPS-2017-425-MKT, https://repub.eur.nl/pub/103265

Pocock, M., Status Inequalities in Business Exchange Relations in Luxury Markets, Promotors: Prof. C.B.M. van Riel & Dr. G.A.J.M. Berens, EPS-2017-346-ORG, https://repub.eur.nl/pub/98647

Pozharliev, R., Social Neuromarketing: The role of social context in measuring advertising effectiveness, Promotors: Prof. W.J.M.I. Verbeke & Prof. J.W. van Strien, EPS-2017-402-MKT, https://repub.eur.nl/pub/95528

Protzner, S., Mind the gap between demand and supply: A behavioral perspective on demand forecasting, Promotors: Prof. S.L. van de Velde & Dr. L. Rook, EPS-2015-364-LIS, https://repub.eur.nl/pub/79355

Reh, S.G., A Temporal Perspective on Social Comparisons in Organizations, Promotors: Prof. S.R. Giessner, Prof. N. van Quaquebeke & Dr. C. Troster, EPS-2018-471-ORG, https://repub.eur.nl/pub/114522

Riessen, B. van, Optimal Transportation Plans and Portfolios for Synchromodal Container Networks, Promotors: Prof. R. Dekker & Prof. R.R. Negenborn, EPS-2018-448-LIS, https://repub.eur.nl/pub/105248

Rietdijk, W.J.R., The Use of Cognitive Factors for Explaining Entrepreneurship, Promotors: Prof. A.R. Thurik & Prof. I.H.A. Franken, EPS-2015-356-S&E, https://repub.eur.nl/pub/79817

Roza, L., Employee Engagement in Corporate Social Responsibility: A collection of essays, Promotor: Prof. L.C.P.M. Meijs, EPS-2016-396-ORG, https://repub.eur.nl/pub/93254

Rösch, D., Market Efficiency and Liquidity, Promotor: Prof. M.A. van Dijk, EPS-2015-353-F&A, https://repub.eur.nl/pub/79121

Schie, R. J. G. van, Planning for Retirement: Save More or Retire Later?, Promotors: Prof. B. G. C. Dellaert & Prof. A.C.D. Donkers, EOS-2017-415-MKT, https://repub.eur.nl/pub/100846

Schoonees, P., Methods for Modelling Response Styles, Promotor: Prof. P.J.F. Groenen, EPS-2015-348-MKT, https://repub.eur.nl/pub/79327

Schouten, K.I.M., Semantics-driven Aspect-based Sentiment Analysis, Promotors: Prof. F.M.G. de Jong, Prof. R. Dekker & Dr. F. Frasincar, EPS-2018-453-LIS, https://repub.eur.nl/pub/112161

Schouten, M.E., The Ups and Downs of Hierarchy: the causes and consequences of hierarchy struggles and positional loss, Promotors; Prof. D.L. van Knippenberg & Dr. L.L. Greer, EPS-2016-386-ORG, https://repub.eur.nl/pub/80059

Sihag, V., The Effectiveness of Organizational Controls: A meta-analytic review and an investigation in NPD outsourcing, Promotors: Prof. J.C.M van den Ende & Dr. S.A. Rijsdijk, EPS-2019-476-LIS, https://repub.eur.nl/pub/115931

Smit, J., Unlocking Business Model Innovation: A look through the keyhole at the inner workings of Business Model Innovation, Promotor: Prof. H.G. Barkema, EPS-2016-399-S&E, https://repub.eur.nl/pub/93211

Straeter, L.M., Interpersonal Consumer Decision Making, Promotors: Prof. S.M.J. van Osselaer & Dr. I.E. de Hooge, EPS-2017-423-MKT, https://repub.eur.nl/pub/100819

Stuppy, A., Essays on Product Quality, Promotors: Prof. S.M.J. van Osselaer & Dr. N.L. Mead. EPS-2018-461-MKT, https://repub.eur.nl/pub/111375

Subaşi, B., Demographic Dissimilarity, Information Access and Individual Performance, Promotors: Prof. D.L. van Knippenberg & Dr. W.P. van Ginkel, EPS-2017-422-ORG, https://repub.eur.nl/pub/103495

Suurmond, R., In Pursuit of Supplier Knowledge: Leveraging capabilities and dividing responsibilities in product and service contexts, Promotors: Prof. J.Y.F Wynstra & Prof. J. Dul. EPS-2018-475-LIS, https://repub.eur.nl/pub/115138

Szatmari, B., We are (all) the champions: The effect of status in the implementation of innovations, Promotors: Prof. J.C.M van den Ende & Dr. D. Deichmann, EPS-2016-401-LIS, https://repub.eur.nl/pub/94633

Toxopeus, H.S., Financing sustainable innovation: From a principal-agent to a collective action perspective, Promotors: Prof. H.R. Commandeur & Dr. K.E.H. Maas. EPS-2019-458-S&E, https://repub.eur.nl/pub/114018

Turturea, R., Overcoming Resource Constraints: The Role of Creative Resourcing and Equity Crowdfunding in Financing Entrepreneurial Ventures, Promotors: Prof. P.P.M.A.R Heugens, Prof. J.J.P. Jansen & Dr. I. Verheuil, EPS-2019-472-S&E, https://repub.eur.nl/pub/112859

Valogianni, K., Sustainable Electric Vehicle Management using Coordinated Machine Learning, Promotors: Prof. H.W.G.M. van Heck & Prof. W. Ketter, EPS-2016-387-LIS, https://repub.eur.nl/pub/93018

Vandic, D., Intelligent Information Systems for Web Product Search, Promotors: Prof. U. Kaymak & Dr. Frasincar, EPS-2017-405-LIS, https://repub.eur.nl/pub/95490

Verbeek, R.W.M., Essays on Empirical Asset Pricing, Promotors: Prof. M.A. van Dijk & Dr. M. Szymanowska, EPS-2017-441-F&A, https://repub.eur.nl/pub/102977

Vermeer, W., Propagation in Networks: The impact of information processing at the actor level on system-wide propagation dynamics, Promotor: Prof. P.H.M.Vervest, EPS-2015-373-LIS, https://repub.eur.nl/pub/79325

Versluis, I., Prevention of the Portion Size Effect, Promotors: Prof. Ph.H.B.F. Franses & Dr. E.K. Papies, EPS-2016-382-MKT, https://repub.eur.nl/pub/79880

Vishwanathan, P., Governing for Stakeholders: How Organizations May Create or Destroy Value for their Stakeholders, Promotors: Prof. J. van Oosterhout & Prof. L.C.P.M. Meijs, EPS-2016-377-ORG, https://repub.eur.nl/pub/93016

Vlaming, R. de, Linear Mixed Models in Statistical Genetics, Promotors: Prof. A.R. Thurik, Prof. P.J.F. Groenen & Prof. Ph.D. Koellinger, EPS-2017-416-S&E, https://repub.eur.nl/pub/100428

Vries, H. de, Evidence-Based Optimization in Humanitarian Logistics, Promotors: Prof. A.P.M. Wagelmans & Prof. J.J. van de Klundert, EPS-2017-435-LIS, https://repub.eur.nl/pub/102771

Vries, J. de, Behavioral Operations in Logistics, Promotors: Prof. M.B.M de Koster & Prof. D.A. Stam, EPS-2015-374-LIS, https://repub.eur.nl/pub/79705

Wagenaar, J.C., Practice Oriented Algorithmic Disruption Management in Passenger Railways, Promotors: Prof. L.G. Kroon & Prof. A.P.M. Wagelmans, EPS-2016-390-LIS, https://repub.eur.nl/pub/93177

Wang, P., Innovations, status, and networks, Promotors: Prof. J.J.P. Jansen & Dr. V.J.A. van de Vrande, EPS-2016-381-S&E, https://repub.eur.nl/pub/93176

Wang, R., Corporate Environmentalism in China, Promotors: Prof. P.P.M.A.R Heugens & Dr. F. Wijen, EPS-2017-417-S&E, https://repub.eur.nl/pub/99987

Wasesa, M., Agent-based inter-organizational systems in advanced logistics operations, Promotors: Prof. H.W.G.M van Heck, Prof. R.A. Zuidwijk & Dr. A. W. Stam, EPS-2017-LIS-424, https://repub.eur.nl/pub/100527

Wessels, C., Flexible Working Practices: How Employees Can Reap the Benefits for Engagement and Performance, Promotors: Prof. H.W.G.M. van Heck, Prof. P.J. van Baalen & Prof. M.C. Schippers, EPS-2017-418-LIS, https://repub.eur.nl/pub/99312

Wiegmann, P.M., Setting the Stage for Innovation: Balancing Diverse Interests through Standardisation, Promotors: Prof. H.J. de Vries & Dr. K. Blind, EPS-2019-473-LIS, https://repub.eur.nl/pub/114519

Wijaya, H.R., Praise the Lord! : Infusing Values and Emotions into Neo-Institutional Theory, Promotors: Prof. P.P.M.A.R. Heugens & Prof. J.P. Cornelissen, EPS-2019-450-S&E, https://repub.eur.nl/pub/115973

Williams, A.N., Make Our Planet Great Again: A Systems Perspective of Corporate Sustainability, Promotors: Prof. G.M. Whiteman & Dr. S. Kennedy, EPS-2018-456-ORG, https://repub.eur.nl/pub/111032

Witte, C.T., Bloody Business: Multinational investment in an increasingly conflict-afflicted world, Promotors: Prof. H.P.G. Pennings, Prof. H.R. Commandeur & Dr. M.J. Burger, EPS-2018-443-S&E, https://repub.eur.nl/pub/104027

Ye, Q.C., Multi-objective Optimization Methods for Allocation and Prediction, Promotors: Prof. R. Dekker & Dr. Y. Zhang, EPS-2019-460-LIS, https://repub.eur.nl/pub/116462

Ypsilantis, P., The Design, Planning and Execution of Sustainable Intermodal Port-hinterland Transport Networks, Promotors: Prof. R.A. Zuidwijk & Prof. L.G. Kroon, EPS-2016-395-LIS, https://repub.eur.nl/pub/94375

Yuan, Y., The Emergence of Team Creativity: a social network perspective, Promotors: Prof. D. L. van Knippenberg & Dr. D. A. Stam, EPS-2017-434-ORG, https://repub.eur.nl/pub/100847

Yuferova, D., Price Discovery, Liquidity Provision, and Low-Latency Trading, Promotors: Prof. M.A. van Dijk & Dr. D.G.J. Bongaerts, EPS-2016-379-F&A, https://repub.eur.nl/pub/93017

Zhang, Q., Financing and Regulatory Frictions in Mergers and Acquisitions, Promotors: Prof. P.G.J. Roosenboom & Prof. A. de Jong, EPS-2018-428-F&A, https://repub.eur.nl/pub/103871

Zuber, F.B., Looking at the Others: Studies on (un)ethical behavior and social relationships in organizations, Promotor: Prof. S.P. Kaptein, EPS-2016-394-ORG, https://repub.eur.nl/pub/94388