The use of event-related potentials in chronic back pain patients
Vossen, C.; Vossen, H.G.M.; van de Wetering, W.; Marcus, M.; van Os, J.; Lousberg, R.

Published in:
Low back pain

DOI:
10.5772/35563

Citation for published version (APA):
The Use of Event-Related Potentials in Chronic Back Pain Patients

Carine Vossen, Helen Vossen, Wiesje van de Wetering, Marco Marcus, Jim van Os and Richel Lousberg
Maastricht University Medical Centre, The Netherlands

1. Introduction

Chronic back pain is one of the most common pain syndromes, with a lifetime incidence of 60% to 90%. An important question in the field of chronic pain is how acute pain transits to a chronic pain state: why do some persons develop chronic pain while others do not? Approximately 10% to 20% of patients with chronic low back pain (CLBP) still have persisting complaints after 6 weeks (Bekkering et al., 2003). If more insight is gained into chronification mechanisms and, as a consequence, the ability to predict which individual with acute pain develops chronic pain, it may become possible to intervene in the process at an early stage.

In acute pain states, pain is often causally related to physical damage, whereas this relationship is less pronounced in chronic pain states. With increasing duration of pain complaints, other factors, such as psychological, cognitive, and environmental factors, are likely to become more involved (Gamsa, 1994). As a result, pain is conceptualized as a multidimensional phenomenon, making pain measurement complex. However, due to the subjective nature of the pain experience, it can not be measured directly. In fact, only derivatives of pain can be measured. The most frequently measured aspect of pain is its intensity. Two often used measures are the Visual Analog Scale (VAS) and Numeric Rating Scale (NRS). Despite some limitations, their psychometric properties have been demonstrated to be adequate (Jensen et al., 1986; Seymour, 1982). To evaluate several other components of pain, pain-related aspects, and risk factors for chronic back pain (such as fear avoidance, inadequate coping strategies, etc.), clinicians use questionnaires. Although many of these instruments provide reliable and valid results (for an overview, see the Handbook of Pain Assessment, edited by Turk & Melzack, 2011), all subjective measures have the potential for several forms of bias (Magnusson et al., 1995). In an attempt to measure relatively unbiased pain responses, a large number of studies in the 70s and 80s investigated the usefulness of psychophysiological recordings. The results of these studies showed that small but significant correlations could be demonstrated between the subjective pain experience on the one hand and skin conductance, heart rate, electromyography, and finger pulse volume on the other (Flor & Meyer, 2011). The most promising results, however, were obtained from experiments studying event (pain)-related potentials (ERPs), a measure that is derived from electroencephalography (EEG). This technique has been used to study
cerebral responses to (non-)noxious stimuli and to gain more insight into the cortical processing of pain. Pain-ERP studies are typically performed in a laboratory setting under strict experimental control. In contrast to the other aforementioned psychophysiological measures, specific ERP components correlate relatively highly with subjective pain estimates (Bromm et al., 1984; Chen et al., 1979; Miltner et al., 1989). In addition, ERPs have been shown to contain information not only on pain intensity but also on many other important (pain-related) factors, such as habituation, personality, and coping strategies. In other words, making use of ERPs, a large number of aspects of the pain construct, interrelations, and mechanisms can be quantified and studied (see Loesers “onion” model in Fig. 1).

![Fig. 1. Loesers “onion” model, modified with the addition of the position of questionnaires and event-related potentials.](image)

This chapter starts with a general description of event-related potentials and their components. Second, factors related to the pain-ERP, such as personality and genetics, are discussed. Special attention will be paid to methods of analyzing the ERP signal. After some discussion of methodological considerations, we will propose an alternative ERP analysis. In addition, we will present preliminary data to illustrate the usefulness of this alternative method. In the last section, some future directions of ERP will be discussed.

2. ERP structure

The ERP represents a cortical response to a specific stimulus—for instance, a sound, a light signal, or a pain stimulus. Event-related potentials are regarded as manifestations of specific (psycho)physiological, stimulus-related processes. An ERP is derived from EEG. Electrodes are attached to specific locations on the scalp (Jasper, 1958). Potential differences between the scalp electrodes and a reference electrode are sampled at a certain frequency, most commonly between 500 and 5000 Hz (cycles per second). The essence of an ERP is that the signal (the cortical reaction to the stimulus) has to be discriminated from background EEG noise. The procedure to achieve this goal is to compute an average of a number of time-locked EEG samples, called epochs or segments. As a general rule, it can be stated that the larger the number of epochs, the better the signal-to-noise ratio. Within each person or
condition, the averaging procedure results in a voltage-by-time graph. When averaging ERPs from different persons or conditions, the result is called a grand average.

![Grand average pain-ERP](image)

**Fig. 2. Example of a grand average pain-ERP on Cz.**

In Figure 2, a grand average pain-ERP (Cz location) is shown. This ERP was obtained from a paradigm in which subjects received a series of 150 painful and non-painful shocks of 10 ms in duration (Vossen, 2010c). As can be seen from Figure 2, three peaks can be clearly distinguished. The most common way to describe these peaks is by polarity and latency (time in ms after stimulus onset). For instance, the large positive peak between 250 and 300 ms is called P300. There is also a system that numbers the sequence of the peaks; e.g., the second positive peak is called P2. In this ERP, P300 is the second positive peak, and thus, P2 and P300 are abbreviations pointing to the same peak. In ERP experiments, researchers try to explain the meaning of the peaks. Which (stimulus) characteristics or processes are 'responsible' for the amplitude and latency of the peaks? Early peaks, also called 'exogenous' components, are believed to represent stimulus parameters, such as the intensity and other properties of the stimulus. Later components are thought to be representations of 'endogenous' processes, such as attention (Luck, 2005). Further, it is known that (slight) differences in the paradigm that is used (not only with respect to the stimulus but also to instructions, environmental characteristics, time of the day, etc.) result in changes in the ERP. These changes pertain to latency and amplitude effects and even profound morphological changes. It should be noted that an ERP has a high temporal but low spatial resolution. The latter means that it is difficult to draw valid conclusions on the underlying cerebral source that is generating the electrical activity (Makeig, 2004).

### 3. Influences on the pain-evoked potentials

In this section, we will discuss the influence of a variety of stimulus-related and person-related factors on the pain-ERP. Special attention will be paid to the predictive relationship of pain-ERPs to the clinical experience of low back pain.
3.1 Stimulus intensity and subjective pain experience

Intensity is an essential property of a stimulus. Is there a relationship between the intensity of a stimulus and peaks of the ERP? Accumulating evidence confirms the relationship between stimulus intensity and increased peak values of N200 and P300 (Becker et al., 1993; Bromm & Meier, 1984; Stowell, 1977). Does an increase in certain peak amplitudes relate to the amount of pain that a subject experiences? One of the first studies that investigated this association demonstrated that an increase in the N200 and P300 peak amplitudes was accompanied by an increase in VAS scores \((r = 0.67-0.77)\) (Harkins & Chapman, 1978). These results were replicated by Chen (1979) and Garcia-Larrea (1997). These authors also found a strong linear association \((r = 0.67 \text{ and } r = 0.41, \text{ respectively})\) between the N200/P300 peak-to-peak amplitude and subjective pain ratings. Miltner and colleagues (1987), however, could not find such a relationship when investigating habituation of noxious stimuli. They observed a significant decrease in peak-to-peak amplitudes of the N150-P360 across trials, but without a corresponding effect on VAS ratings. They suggested that the association between ERP amplitudes and subjective ratings might not be as strong as was claimed previously.

The subjective pain experience seems to be limited not only to these peak amplitude effects. Kanda and colleagues (2002) discovered a late positive component around 600 ms that was associated with pain report, which they called the ‘intensity assessment-related potential’ (IAP). The IAP was not influenced by intensity, suggesting that this component solely reflects the psychological processes of pain. A recent study, again investigating the relationship between stimulus intensity, peak amplitudes, and subjective pain experience, was performed by Vossen and colleagues (2011). Their methodological comments on common pain-ERP analyses relate to the problems inherent in the averaging technique (see also paragraph 4). The authors argued that averaging eliminates any unwanted ‘noise’ in the ERP, but in doing so, it assumes no difference between repeated trials. This assumption can not hold, since single trials likely differ from one another because of processes, such as habituation. Moreover, averaging across trials eliminates all information about possible within-subject correlations between ERP and subjective pain. As an alternative, they introduced multilevel random regression analysis, applied on pain-ERP, making it possible to model time (habituation), stimulus intensity, and their random within-person effects. The findings of this study show that the relationships between these three variables are confounded and moderated by several other variables, such as the intensity of the previous stimulus. This means that a certain pain rating after a stimulus also depends on the intensity of the previous stimulus, which makes clinical sense. A pain patient who is asked to evaluate his/her perceived pain is highly likely to base this evaluation on previous pain experiences. In sum, the relationship between stimulus intensity, ERP peaks, and pain experience is probably far more complex than previously thought.

3.2 Habituation and pain-evoked potentials

Habituation is the process that refers to a decrease in a behavioral response to a repeatedly presented stimulus (Thompson & Spencer, 1966). It could be hypothesized that altered habituation might be an explanation for the chronicification of pain. It is thought that chronic pain patients may have a deficit in habituation or even an inability to habituate to painful experiences, resulting in persistent pain. Older studies, using pain rating as an outcome
measure, reported mixed results. One study investigated the habituation difference between CLBP patients and controls, using eight successive trials of the cold pressure test (Brandt & Schmidt, 1987). Healthy controls could be divided into a subgroup that habituated over trials and a subgroup that sensitized. The CLBP group did not habituate or sensitiz over time. Additionally, they found a lower pain tolerance in CLBP patients while reporting higher pain ratings. It was hypothesized that CLBP patients had already undergone a learning process in which sensitization had taken place. Arntz and colleagues (1991) also studied habituation in CLBP patients and controls. They did not observe a difference in habituation between the groups, measured by pain intensity ratings, EMG, and heart rate. A third study (Peters et al., 1989) confirmed the results of Brandt & Schmidt (1987) but did not find differences in physiological measures, such as heart rate and skin conductance. More recently, Smith and colleagues (2008) reported differences in habituation of subjective pain ratings between women with fibromyalgia and pain-free controls. They found that women with fibromyalgia habituated at a lower rate to repeated heat stimuli.

In addition, there are some recent studies that have used ERP as a measure to study habituation. They are suggestive of a deficit in habituation in chronic pain patients, although different chronic pain populations were used. Valeriani and colleagues (2003) studied habituation in response to painful CO$_2$ laser stimulation in migraine patients. They found reduced habituation of ERP amplitudes in migraine sufferers compared to pain-free controls. In disconfirmation, another study found that patients with migraine did not show any habituation, whereas healthy controls did (De Tommaso, 2005). Vossen et al. (2010c) studied habituation in a group of chronic low back pain patients compared to pain-free controls, measuring ERP in response to 20 painful stimuli. They found a significant interaction between group and trial number on the P300 component at C4 and T4. This means that chronic low back pain patients appeared to habituate to a lesser degree than pain-free controls. They also examined the influence of state-depression on habituation, using the BDI score. The results revealed a significant three-way interaction between BDI, group, and trial, suggesting that the difference in habituation between groups depends on the level of depression. Only in the presence of depression did CLBP patients show a deficit in habituation. Interestingly, a recent study in fibromyalgia patients also found evidence for reduced habituation of the N200 vertex component, facilitated by the presence of symptoms of depression (De Tommaso, 2011). In conclusion, habituation seems to be different in chronic back pain patients compared to controls but is probably also influenced by factors, such as depression.

3.3 Influence of neuroticism on pain-evoked potentials

“Personality” can be defined as a dynamic and organized set of characteristics possessed by a person that uniquely influence his or her cognitions, motivations, and behaviors in various situations (Ryckman, 2008). Individuals with different personalities will differ in reaction to a specific situation. Likewise, it is conceivable that persons with diverse personality structures will react differently to a pain stimulus. This theoretical claim is frequently being confirmed in clinical practice. There is a large variety in ‘pain behaviors’ when patients are confronted with painful medical procedures (injections, stitches, etc.). One of the most important personality factors that are known to influence the experience of pain is neuroticism (Wade et al., 1992). Neuroticism is defined as a tendency to experience negative
emotions in stressful situations (Costa & McCrae, 1980). One of the most commonly used questionnaire measuring neuroticism, is the NEO-Big 5 (Costa & McCrae, 1985), which also gives information on six neuroticism facets—namely anxiety, impulsivity, depression, self-consciousness, irritability, and vulnerability.

There are several mechanisms that explain the hypothesized relationship between neuroticism and pain. In the first explanation, the relationship between neuroticism and pain is thought to arise from over-reporting of pain-related complaints, an exaggerated expression of disturbance, and a more focussed view on bodily states. Persons with high levels of neuroticism tend to be more aware of their bodily states than others and thus report more physical complaints (Groth-Marnat & Fletcher, 2000). Costa and McCrae (1980) suggest that in patients with high levels of neuroticism, physical complaints can be viewed as exaggerations of bodily concerns, linking neuroticism to hypochondria. In yet another explanation, neuroticism can be seen as a vulnerability factor. When a patient is confronted with a stressor, such as low back pain, patients with high neuroticism levels already might perceive pain as threatening at lower thresholds, which consequently may evoke catastrophic thoughts (Goubert et al., 2004). In a study on a large sample (n = 1441) of CLBP patients, Bendebba found a correlation between the severity of perceived pain and psychological distress. A correlation between psychological distress and the duration of the complaint, however, could not be demonstrated (Bendebba, 1997).

Note that the aforementioned studies are based on data derived from questionnaires. ERP can be used as an additional tool to get more insight into the mechanism(s) of pain and neuroticism. Vossen et al. performed a study in 75 healthy subjects in which they studied the influence of neuroticism and two NEO-big 5 facets (depression and anxiety) on the pain-ERP. They found that subjects with relatively high neuroticism scores showed more positive ERP amplitudes. These amplitude effects were observed frontally in a broad latency range, from 250 to 1500 ms, with significant effects between 340-400 ms, 730-860 ms, and 1240-1450 ms, suggesting stronger pain processing. Comparing the effects of the neuroticism facets anxiety and depression, opposite effects were found. Anxiety was associated with a negative effect (enlarging negative amplitudes) early in the ERP (100-200 ms), whereas depression exerted an opposite effect in the same latency range. Another study of 14 healthy participants also found anxiety to be related with a larger N140 component (Warbrick, 2006). In addition, the authors also observed no effect of anxiety in the P300 range. It is known that the N140 increases when attention to a stimulus is heightened (Kida et al., 2004). It is plausible that participants focus their attention more when they are more anxious, thus increasing the amplitude of N140. This would support the idea of a greater focus on bodily states in anxious patients. In conclusion, the number of pain-ERP studies investigating the relationship between personality factors (especially neuroticism) is relatively small. More experimental data are needed to unravel mechanisms involving pain report, personality, and cortical pain processing.

3.4 The influence of gene polymorphisms on pain-evoked potentials

There is a rising interest in genetic factors in pain research, as they likely explain a substantial portion of the interindividual differences in pain perception and response to pain treatment (Fillingim et al., 2008). Twin studies give the opportunity to determine the proportion of variability in pain response that is accounted for by genes (heritability).
Heritability for migraine has been estimated at 35% (Stam et al., 2010), 55% for menstrual pain (Treloar et al., 1998), and 33% to 50% for low back pain (Bengtsson & Thorson, 1991 Battie et al., 2007). The main aim in pain genetics, however, is to identify the actual genes and gene polymorphisms that influence the pain pathways. Linkage and association studies have attempted to identify specific genes that affect the peripheral nervous system through the voltage-gated sodium channels on the one hand and genes that affect the central nervous system and modulate sensory-discriminatory and affective-evaluative elements of pain perception that affects the central nervous system on the other (for an extensive overview, see Foulkes & Wood, 2008). Many candidate genes have been proposed, but the effects of these genes are small and even together, if true, explain only a fraction of the heritability involved. A possible explanation for this is the complexity of pain as a phenotype. Measurement of the pain experience plays an important part in this. In human studies, three single-nucleotide polymorphisms (SNPs) have been proposed to impact pain perception: COMT Val158Met (rs4680), BDNF Val66Met (rs6265), and OPRM1 A118G (rs1799797). COMT Val158Met is a gene polymorphism that alters the activity of the COMT enzyme, which degrades catecholamines, such as dopamine, epinephrine, and norepinephrine (Nackley et al., 2006). It has been demonstrated that Met/Met homozygotes have decreased mu-opioid system activation in response to pain (Zubieta et al., 2001, 2003); however, further replication is required. Brain-derived neurotrophic factor (BDNF) is a neurotrophin that supports the growth, differentiation, and survival of neurons in both the peripheral and the central nervous system. BDNF is released when nociceptors are activated and is involved in the activity-dependent pathogenesis of nociceptive pathways, which may lead to chronification of pain (Merighi et al., 2008; Sen et al., 2008). One piece of genetic variation within the BDNF gene is a valine-to-methionine substitution at codon 66 (Val66Met), resulting in reduced secretion of the BDNF protein and impaired BDNF signaling. The Met carriers are believed to be more sensitive to pain; however, here, replication in large and systematic studies is also required.

Experimental designs represent a particularly powerful approach to study genetic effects on psychological phenotypes, as they allow for controlled conditions and investigation of underlying mechanisms (van Os et al., 2008).

Lötsch and colleagues (2006) studied the influence of the G allele of the OPRM1 A118G polymorphism on ERP pain processing of experimental pain stimuli. This polymorphism replaces adenine with guanine, increasing the receptor affinity of b-endorphin 3-fold, resulting in decreased pain responses (Bond et al., 1998; Filligim et al., 2005). Lötsch and colleagues concluded that ERP amplitudes (N1 component) of carriers of the G allele were, on average, half as high as the amplitude of the non-carriers, suggesting lower pain processing for the G allele carriers.

In a more recent study, we investigated the influence of the COMT Val158Met, BDNF Val66Met, and BDNF Val66Met polymorphisms on pain using ERPs (Vossen et al., 2010a). The sample of this study consisted of chronic low back pain patients, as well as healthy controls. The results suggest that the COMT Val158Met and the BDNF Val66Met polymorphisms influence the cortical processing of experimental electrical pain stimuli. However, no main gene effects were observed. Rather, genetic effects appeared to be moderated by the concurrent presence of chronic pain complaints. In the presence of chronic pain, the COMT Met allele and the BDNF Met allele augmented cortical pain processing at
the N2 and P1 components, respectively, whilst reducing pain processing in pain-free controls. The findings of Lötsch and colleagues (2006) concerning the OPRM1 A118G polymorphism could not be replicated in our study. The influence of chronic pain complaints on gene effects may indicate a gene-environment interaction and may even implicate epigenetic modification. It is clear that genetic findings remain preliminary, and well-conducted systematic studies with larger samples sizes are required.

Up to now, limited attention has gone out to gene-environment interplay in pain research, especially with event-related potentials as pain measure. In future studies investigating gene-environment interplay, the complexity of the phenotype and the overall small direct effect of genes (Manolio et al., 2009) should be considered. Longitudinal designs using event-related potentials can contribute to the study of genetic influences in causal pathways of the chronicification of pain.

3.5 The predictive relationship of pain-ERPs to clinical experience of pain

Since chronic low back pain is a very common problem and accompanied by high costs in health care, it is important to be able to predict the likelihood of developing chronic disabling back pain. In 2010, Chou & Chelleke performed a systematic review of 20 studies to investigate the usefulness of individual risk factors for chronicification in low back pain. Identified risk factors, although individually relatively weak, were maladaptive pain coping behaviors, nonorganic signs, functional impairment, a poor general health status, and presence of psychiatric comorbidities. They also reviewed risk-predicting instruments, which are usually based on self-reported questionnaires. To date, no instrument has been used routinely and no recommendations exist, since evidence is insufficient (Chou & Chelleke, 2010). Could the pain-ERP serve as a predictor for chronic low back pain? The experimentally induced pain-ERP has been demonstrated to be a relatively objective measure of experimental pain compared to subjective pain ratings (Becker et al., 2000; Bromm, 1984; Stowell, 1977). An important issue, however, concerns the relevance and translation of the experimentally induced pain-ERP to pain in daily life. Stated in another way, can the pain-ERP serve as a predictor for clinical pain? There are two fundamental problems, which are related to the meaning of experimentally induced pain and its generalizability to pain in daily life. First, the characteristics of experimentally induced pain stimuli are typically not comparable with those of clinical pain (e.g., the intensity and duration). Second, in an experimental environment, the subject has at least partial control of the experimentally induced pain (escape is possible by stopping the participation), a controllability that cannot be exerted in a clinical setting. Consequently, a straightforward translation of experimentally induced pain to clinical pain is simply not possible. Nevertheless, event-related potentials have already been used to predict depression (Kemp et al., 2006), awakening from a coma (Daltrozzo et al., 2006), and in the discrimination of Alzheimer’s disease from controls (Benvenuto et al., 2002). The prediction of pain, using ERPs, has not been studied intensively. To our knowledge, only one study investigated the prediction of chronic low back pain complaints (Vossen et al., 2010b). Even-related potentials in response to experimental pain were measured in 75 CLBP patients. The ERP mean amplitudes of the peaks (N1, P1, N2, P3) served as predictors for the mean pain ratings, registered during a 2-week period after the experiment. The N2 component of Cz and C4 appeared to be significantly related to the daily pain ratings, collected over 2
consecutive weeks. Surprisingly, the ERP variables related more strongly to the clinical pain ratings than the accompanying subjective ratings of the experimental pain stimuli. Although care must be taken in the interpretation of these results, the findings suggest that it might be possible to make inferences on clinical pain, based on experimentally derived pain-ERPs. More studies are needed to confirm these results and to investigate the usefulness of predicting long-term complaints in patients with acute or chronic back pain.

4. Analyzing event-related potentials

In the last 2 decades, the methods of analyzing ERPs have not been changed essentially. In the first paragraph, we will describe the most commonly used method. In the second paragraph, we will discuss several issues concerning the methodology. Additionally, we will introduce an alternative method and present preliminary results.

4.1 Common methods in analyzing pain-ERPs

In an experimental ERP paradigm, stimuli are repeated to allow averaging of the epochs and to compare different experimentally induced conditions. Although many variants are possible, the most common procedure of ERP analysis is as follows (Luck, 2005; Mouraux & Iannetti, 2008): The first step is to filter the raw EEG data. The second step is the creation of segments or epochs, based on markers of the stimuli in the EEG. The duration of these epochs varies but is usually between 500 and 1500 ms. The third step concerns identification of invalid epochs. Epochs are qualified as valid or not, depending on whether an epoch is likely to be confounded by an artifact: electrical activity that does not arise from the brain—for example, an eyelid movement, tension of the muscles in the head and neck, or electrical activity from the heart. Basically, there are two procedures for dealing with invalid epochs. The first method is a rejection of invalid epochs in the computation of the averaged ERP. This simply reduces the maximum number of analyzable segments, and as a result, information is lost. The second option is a correction for confounding effects by commonly accepted statistical algorithms, such as the so-called Gratton and Coles ocular correction (Gratton, 1983). To date, it is not clear which of these two methods is preferable. After the averaging of all epochs per individual (step 4), a grand average of ERP segments across subjects (per experimental condition) is calculated (step 5). The next action is to carefully identify peaks with their corresponding latency windows: a time range surrounding a specific peak. The seventh step is to apply these ‘peak latency windows’ at the within-subject epoch level: within the defined latency window, the maximum (or minimum) amplitude is determined. These maximum amplitudes form the input for the computation of the peak average for each individual. The final action is to use these maximum or minimum amplitudes as a dependent variable in statistical analyses, such as ANOVA (Hoormann et al., 1998).

4.2 Methodological considerations

Although this procedure of ERP analysis is plausible, functional, and generally accepted, there are some critical issues that need to be considered, particularly given recent developments in statistics that may provide superior analytical approaches. First, each time-locked EEG segment consists of the aimed signal and a noise element (all background
ongoing EEG activity). Averaging of the trials will separate the signal from the noise, because the signal element is thought to be constant in every trial, while the noise element is considered to be random. However, one can dispute the fact that a signal is constant over trials in pain experiments, since processes, such as habituation, play an important role (Woestenburg et al., 1983; Vossen et al., 2006). In addition, within-subject variance (trial-to-trial variance) is lost by averaging, which may contain clinically important information on cortical processes. Second, it is known that in consecutive trials, the latency of maximum peak values is likely to differ. Although it is possible to take the variability of latency into account in the analysis (as a covariate), this solution is not ideal, since the trial-to-trial latency information is lost. A third unsolved problem is how to deal with peak values located on the borders of the latency window. A final critical point regards the fact that peaks contain information on many processes: it is generally known that P300 is sensitive for attention, evaluation, stimulus intensity, and many other stimulus-related and person-related factors (Zaslansky et al., 1996). Multilevel random regression analysis, as already discussed in paragraph 3.2, tackles the problems associated with averaging and habituation. Nonetheless, the methodological problems concerning peak definition and peak measurement can not be solved with multilevel analysis.

Without a doubt, averaged maximized peak values carry important information. However, theoretically spoken, each (!) latency point contains meaningful information. To be able to analyze amplitude information that is not related to peaks, area under the curve (AUC) can be computed for specific latency ranges. Usually, AUC is applied to quantify peaks as well as to calculate averaged group differences located on the flanks/limbs of a peak (Luck, 2005). AUC is not often applied in pain-ERPs, since it has been postulated that more noise is introduced when averaging trials (Picton, 2000). However, when using AUCs of single-trial data, the problem of introducing noise is substantially reduced.

5. Introducing an alternative method

In this section, we will discuss an alternative method for analyzing ERP data. We will present preliminary results using this method in a previously used dataset of CLBP patients and pain-free controls.

5.1 Fixed-interval AUC segment analysis

From a statistical point of view, the main goal of ERP analysis is to explain variance in the pain-ERP as much as possible, using a series of predictors. It seems reasonable to focus not only on the explanation of maximum peak amplitudes but also on effects in other latencies. This assertion is supported by the fact that several above-cited publications (e.g., Kanda et al., 2002; Vossen et al., 2006) report stimulus- and person-related effects in non-peak latencies. We felt that the concept of AUC is valuable but should be applied to small fixed intervals, independent of peaks. This implicates a partitioning of the whole epoch in small event-related fixed interval areas (ERFIs). To illustrate this line of reasoning, three averaged pain-ERPs are presented. In the first picture, three ERPs are depicted, each representing another level of stimulus intensity. As can be observed, there are intensity effects on P1 and N2 and a large effect on P2: the larger the intensity, the larger the peak amplitude. However, the intensity effects are not limited to these peaks. The effect on the P2
already starts at approximately 200 ms and lasts at least until 500 ms. Also, the intensity effect between P1 and N2 can not be ignored. The second graph illustrates the effects on habituation. The three ERPs represent three blocks of trials delivered in the experiment. Again, habituation is not restricted visually to the peaks. Also, habituation seems to reduce the amplitude. In the third graph, two grand averages are shown of ERPs on Cz: one from a pain-free control group (n = 76) and the other from a group of chronic low back patients (n = 75). There seem to be small amplitude group effects on the P1 and N2 but not on the P2. In addition, there seem to be non-peak-related group effects. Care must be taken in the interpretation of differences observed in these grand averages, since they represent a reduced, oversimplified representation of the pain experience. In the intensity ERPs, the information on habituation is averaged out and vice versa.

Based on these observations, we performed a number of (unpublished) pilot analyses on ERP datasets of previous studies to determine a pragmatic width of AUC segments (ERFIAs). This led to our choice of segments of 20 ms. In our view, this seems to be a reasonable compromise between specific AUC segments that are too large on the one hand and segments that are too small, resulting in multiple testing problems on the other.

We decided to reanalyze part of the data pertaining to the PhD thesis, defended by H. Vossen. We focused the preliminary analysis on three electrodes, namely C3, C4, and Cz, because these locations represent the sensomotoric cortex and are of anatomical importance in pain processing (Kupers & Kehlet, 2006). Also, we restricted the range to 0-500 ms post-stimulus. The reanalysis took place in an explorative, hypothesis-generating fashion. Basically, we were interested in to what degree the proposed ERFIA method would yield significant relationships between stimulus intensity and habituation and to what degree these findings would correspond to known results, based on peak analyses. In addition, there was one special point of interest: Do the ERPs of chronic pain patients differ from pain-free controls, analyzed with fixed-interval AUCs?

Fig. 3. Grand ERPs of stimulus intensity on Cz.
5.2 Study design

The dataset we are using is based on previously collected raw EEG data. For a detailed description of the protocol, we refer to Vossen et al., 2010a & 2011. Here, a summary of the design is given. Seventy-six pain-free subjects and 75 patients with chronic low back pain participated in the study. All CLBP subjects suffered from low back pain for at least 6 months and were recruited from the general population. All subjects underwent a rating paradigm of 150 semi-randomly presented electrical stimuli. The used stimuli, administered intracutaneously on the top of the left middle finger, consisted of electrical pulses, each with a duration of 10 ms, and an inter-stimulus interval (ISI) ranging from 9 to 11 seconds. Before starting the experiment, the sensory and pain threshold were determined. In the experiment, five different intensities, based on the participant’s pain threshold, were administered. The
The Use of Event-Related Potentials in Chronic Back Pain Patients

five used intensities were -50% and -25% below the pain threshold, the pain threshold itself (0%), and 25% and 50% above the pain threshold. After each stimulus, subjects were asked to rate the intensity on a numeric rating scale (NRS) ranging from 0 to 100. During the entire experiment, EEG was recorded with a 1000-Hz sampling rate. The ERP epochs were selected from the continuous EEG and segmented at 200 ms prior to the stimulus to 500 ms post-stimulus. For each stimulus, we calculated 20-ms ERFIas in the range of 0 to 500 ms. ERFIA segments with EOG activity exceeding +25 mA or -25 mA were excluded from the analysis. The calculated ERFIAs were used as dependent variables in a multilevel random regression model (see equations 1 and 2). This resulted in 25 separate multilevel regression analyses per electrode. All analyses were performed with SPSS 18.0.

\[ Y_{ti} = \beta_0 + \beta_1 \text{intensity}_{linear} + \beta_2 \text{trial}_{linear} + \beta_3 \text{trial}_{quadratic} + \beta_4 \text{trial}_{inverse} + \beta_5 \text{group} + \beta_6 \text{age} + \beta_7 \text{gender} + \beta_8 \text{sensation threshold} + \beta_9 \text{pain threshold} + \beta_{10} \text{intensity}_{linear} \text{of previous trial} + \epsilon_t + u_{1t} \text{intensity}_{linear} + u_{2t} \text{trial}_{linear} + u_{3t} \text{trial}_{inverse} + u_{4t} \text{trial}_{quadratic} \]  

(1)

\[ Y_{ti} = \beta_0 + \beta_1 \text{intensity}_{linear} + \beta_2 \text{trial}_{linear} + \beta_3 \text{trial}_{quadratic} + \beta_4 \text{trial}_{inverse} + \beta_5 \text{group} + \beta_6 \text{age} + \beta_7 \text{gender} + \beta_8 \text{sensation threshold} + \beta_9 \text{pain threshold} + \beta_{10} \text{intensity}_{linear} \text{of previous trial} + \beta_{11} \text{group*intensity}_{linear} + \beta_{12} \text{group*trial}_{linear} + \beta_{13} \text{group*trial}_{quadratic} + \beta_{14} \text{group*trial}_{inverse} + \beta_{15} \text{group*intensity}_{linear} \text{of previous trial} + \beta_{16} \text{group*pain threshold} + \beta_{17} \text{group*sensory threshold} + \epsilon_t + u_{1t} \text{intensity}_{linear} + u_{2t} \text{trial}_{linear} + u_{3t} \text{trial}_{inverse} + u_{4t} \text{trial}_{quadratic} \]  

(2)

5.3 Preliminary results

First, some general group characteristics are presented in Table 1. The CLBP patients report much more pain and pain interference. In addition, as might be expected, their mood is negatively affected, which is expressed in a higher depression score.

<table>
<thead>
<tr>
<th>Group</th>
<th>Pain-free control (N = 76)</th>
<th>CLBP patients (N = 75)</th>
<th>T-value/(\chi^2)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>34.68</td>
<td>42.11</td>
<td>3.15</td>
<td>0.002</td>
</tr>
<tr>
<td>Gender male</td>
<td>26</td>
<td>34</td>
<td>2.26</td>
<td>0.09</td>
</tr>
<tr>
<td>Gender female</td>
<td>50</td>
<td>41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain magnitude (SF-36)</td>
<td>1.65</td>
<td>3.51</td>
<td>13.73</td>
<td>0.000</td>
</tr>
<tr>
<td>Pain interference (SF-36)</td>
<td>1.18</td>
<td>2.23</td>
<td>9.84</td>
<td>0.000</td>
</tr>
<tr>
<td>Depression (BDI)</td>
<td>3.35</td>
<td>7.09</td>
<td>5.8</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 1. Characteristics of the two experimental groups.

Because of the preliminary aspect of the analyses, a full description of all results is beyond the scope of this chapter. Therefore, the focus will be on the robust and salient results of the analyses.
We present the results of four independent variables, applied in the first multilevel model (see equation 1): intensity, the intensity of the previous trial, habituation (analyzed with a linear contrast), and group (CLBP patients versus pain-free controls). The results for the 20-ms periods between 0 and 500 ms are shown in Figures 6-9.
On the vertical axis, the T-value of the variable is depicted, and the horizontal axis represents latency (i.e., the 25 consecutive 20-ms ERFIAs). The variable intensity, habituation (linear), and previous trial intensity show profound and long-lasting, significant effects, as shown in the figure. Note a significant negative effect of stimulus intensity on all three electrodes from 60 to 120 ms and a very strong positive effect from 160 to 400 ms. Remarkably, two positive intensity processes in the latency range between 100 and 400 ms are apparent (from 140-200 ms and from 200-400 ms). In addition, an asymmetry can be observed between C4 and C3/Cz in the 140-to-200-ms range, where C4 demonstrates no
clear significance. In contrast to the results presented by Vossen and colleagues (2011), where no main effects of the previous trial intensity variable were found in peak amplitude analyses, the ERFIs show a large and consistent negative effect from 320 to 500 ms at all three electrodes. There are some small, significant, positive effects in 180-to-260-ms latency. With respect to linear habituation, large and long-lasting significant t-values can also be identified. The linear habituation emerges from the 100-140-ms range as a significant positive effect and becomes significantly negative in the range from 200-380 ms. Although less pronounced, the linear habituation t-curve appears to be opposite compared to the intensity t-curve; whereas intensity has an amplitude-inflating effect in the range from 140 to 400 ms, linear habituation has the opposite effect. Although not displayed, significant effects were observed for the inverse variables habituation (1/trial) and quadratic habituation. Inverse habituation had a clear, significant (T-values between -2 and -4) amplitude-reducing effect during a latency period of 360-500 ms, and quadratic habituation showed strong, significant amplitude-inflating effects (T-values up to 4) in the range of 200-300 ms. No convincing main effect of group, independent of the effect of all other variables in the model, could be demonstrated. The only area of interest appears to be located on C4 and is situated between 140 and 200 ms, but taking the number of tests into account (3x25=75), this effect is questionable.

Finally, we investigated whether there were significant interaction effects with the group variable (see equation 2). A priori (see Vossen et al., 2011), we expected a group*habituation interaction showing CLBP patients to have a reduced habituation. We were also interested in whether the intensity effect was modified by group. In order to limit the number of figures, we present the results on the group by intensity and group by linear habituation interactions (note that group was coded “1” for CLBP patients and “0” for pain-free controls). As can be seen from Figures 10 and 11, the group effect may depend on stimulus intensity (especially in the 200-300-ms range) as well as linear habituation (320-440 ms, especially on Cz). There were also clear significant effects (T-values between 2 and 3 between 320 and 440 ms) for the group by quadratic habituation (no graphs included).

Fig. 10.
5.4 Discussion of the preliminary findings

In paragraph 4.2, several critical points were discussed in relation to ERP peak measures, followed by the proposal of an alternative approach based on AUC, analyzed with multilevel random regression techniques. A number of explorative analyses were performed. We are aware of the fact that the results are merely explorative, since we only analyzed three cranial locations and restricted the analyses to a 500-ms post-stimulus time range. However, given these limitations, we think the analyses show promising results and illustrate proof of concept.

As a first step, we investigated whether the ‘new’ method would yield comparable results, with respect to already established relationships between peak-ERP, stimulus intensity, and habituation. Reviewing our results, the answer seems to be affirmative. Consistent results were found for all three central electrodes. A large number of the 20-ms ERFIA segments were significantly related to stimulus intensity and habituation. When examining the areas in which we would expect significant results for the intensity and habituation variables a priori, the alternative method produced results comparable to other studies (e.g., Bromm & Meier, 1984; Stowell, 1977). For example, the main effect of the variable intensity was very significant in the N2 range and the P3 range. A similar observation could be made for linear habituation in the N2 range (De Tommaso, 2011).

In addition to these basic ‘validating’ analyses, we were also interested in whether ERP-pain processing of CLBP patients differs from pain-free controls. No clear main group effect could be observed. However, the results strongly suggest that the effects on intensity and linear and quadratic habituation depend on being a CLBP patient or not. An interesting observation was that the group*intensity interaction took place at an earlier latency range compared to the group*habituation interaction in the ERP. The group*intensity interaction effect is situated in the latency range of the P3 and probably can be replicated using peak amplitudes. The group*habituation effect, however, is situated after the P3 and, as a
consequence, most likely can not be found in peak analyses. These off-peak effects may be valuable in the search for chronification mechanisms.

All group interaction effects are based on a contrast of a subclinical CLBP population to pain free-controls. The choice of this CLBP group may be disputed, since this subclinical group is likely to be heterogeneous with respect to underlying pathology. Nonetheless, the CLBP group clearly differs from the control group with regard to the key variable pain (see Table 1). The analyses of the ERFIA segments seem to produce more pronounced and significant results, compared to the peak results published by Vossen (2010c). This can be concluded not only from the very large T-values (up to 14) but also from the prolonged latency effects. A typical example is the very broad latency window (from 160-420 ms) of the main effect on intensity. Also, effects of linear habituation are significantly embedded in a large range of consecutive ERFIA segments. Interestingly, intensity of a previous stimulus, indicative for a 'memory' of painful events, showed a significant long-lasting influence (see Figure 5). Since no apparent peaks emerge after 300 ms in the averaged pain-ERP (see Figure 1), the ERFIA method seems to be more useful to detect such late effects than the peak method. This is demonstrated by the fact that we could not find a main effect of the previous stimulus intensity in earlier analyses (Vossen et al., 2011). By plotting the T-values of consecutive ERFIAs, we observed another advantage in the interpretation of the results. In time, variables become more significant and reach a 'peak significance,' followed by a decrease. This information gives insight into the start and end of an influential effect of a variable. To illustrate, intensity seems to have two main effects in the latency range of 140-400 ms. One could speculate that this T-value graph (Figure 3) is indicative for two 'intensity' processes.

Some critical aspects need to be considered in the application of the proposed ERFIA method. First, a large number of consecutive ERFIA segments may result in an unacceptable number of statistical tests. In our view, a rigid correction method for multiple testing, such as the Bonferroni correction, would increase the risk of rejecting 'real' effects in this early, explorative phase of the study. However, an appropriate correction for multiple testing is required. Another critical point concerns the optimal width of ERFIAs. In the present study, we used fixed segments of 20 ms. In order to get a general impression of effects within a relatively large window (500 ms or more), we judge the 20-ms criterion to be appropriate. When investigating small effects, one could argue for the use of smaller areas. Enlargement of the width would reduce the number of tests but may introduce more noise. A third note is related to EOG rejection. In the present analyses, we used a ±25-μV criterion to reject AUCs. It remains to be investigated whether this criterion is optimal. In handling confounded EOG segments, the use of multilevel analyses is especially worthwhile, since all valid, analyzable segments are included, whereas in analysis of variance, a whole subject would have been excluded in the case of too many invalid segments. Finally, one major disadvantage of both peak and AUC measures is the poor spatial resolution. Other techniques in analyzing ERP have been developed to overcome this problem. As an example, probabilistic independent component analysis (PICA) has been applied to gain more insight in the source of underlying multimodal and modality-specific neural activities (Jung et al., 2001; Makeigh et al., 1997; Mouraux et al., 2009). Also, many fMRI studies and magnetoencephalography studies are emerging, with high spatial resolution (Bromm, 2001; Makeig, 004; Stancak et al., 2011; Peyron et al., 2000). Combining ERP methods with fMRI will allow investigation of pain processing in a temporal as well as a spatial superior fashion.
6. Conclusion

In conclusion, without doubting the importance of maximized data derived from peak analyses, one could express doubt whether this approach represents too large a reduction and oversimplification of the post-stimulus cortical processing. In this respect, the present alternative method seems to be more appropriate. A direct comparison of the methods is difficult, if not impossible, since the ERFIA method is based on fixed latency intervals for all trials, whereas in the peak method, the latency of the maximum amplitude differs per trial. Future research has to clarify when to use peak amplitude analysis and in which situations a fixed-AUC method is more suitable.

Using ERP measures, many interesting insights in cortical processing of pain are emerging, such as habituation processes, genetic influences, and influences of personality. These phenomena may contribute to finding explanations for the transition of acute pain to chronic (low back) pain states. Nevertheless, longitudinal research designs are necessary to study this process in detail, as well as a combination of ERP with other methods, such as fMRI, and magnetoencephalography. Furthermore, the application of mixed regression will enable a better understanding of the variance in the pain-ERP. Once the pain-ERP and its underlying cortical processes are understood more completely, the path to remediation in clinical practice is open. Then, development of diagnostic tools could be in reach.

7. Acknowledgments

We are grateful to Jacco Ronner and his colleagues, Department of Instrumentation, Faculty of Psychology and Neuroscience, Maastricht University, for their technical assistance and programming.

8. References


