

Trial and Error (-Related Negativity): An Odyssey of Integrating Different Experimental Paradigms

Juliane Traxler^{1,2†}, Roxane V. Philips^{1†}, Andreas von Leupoldt¹,
Johan W. S. Vlaeyen^{1,2}

¹Research Group Health Psychology, KU Leuven, Leuven, Belgium

²Experimental Health Psychology, Maastricht University, Maastricht, The Netherlands

†Equally contributing authors.

Correspondence: Juliane Traxler, Research Group Health Psychology, University of Leuven, Tiensestraat 102, Box 3726, 3000 Leuven, Belgium

Email: juliane.traxler@kuleuven.be

Funding: This work was supported by the 'Athenes' long-term structural funding Methusalem grant (METH/15/011) from the Flemish Government, Belgium, and by an infrastructure grant from the Herculesstichting, Belgium (AKUL/13/07).

Acknowledgments: The authors wish to thank Mathijs Franssen and Jeroen Clarysse for the brainstorming and their technical support in programming the experiment. Furthermore, the authors would like to extend their gratitude and acknowledgements to all study participants for their time and critical reflections.

Companion Piece: This paper has an accompanying reflection piece, [Derksen \(2020\)](#). Read the abstract at the end of this article.



Abstract

Pain can be considered as a signal of “bodily error”: Errors – discrepancies between the actual and optimal/targeted state – can put organisms at danger and activate behavioral defensive systems. If the error relates to the body, pain is the warning signal that motivates protective action such as avoidance behavior to safeguard our body’s integrity. Hence, pain shares the functionality of errors. On the neural level, an important error processing component is the error-related negativity (ERN), a negative deflection in the electroencephalographic (EEG) signal generated primarily in the anterior cingulate cortex within 100 ms after error commission. Despite compelling evidence that the ERN plays an important role in the development of various psychopathologies and is implicated in learning and adjustment of behavior, its relation to pain-related avoidance has not yet been examined. Based on findings from anxiety research, it seems conceivable that individuals with elevated ERN amplitudes are more prone to engage in pain-related avoidance behavior, which may, under certain conditions, be a risk factor for developing chronic pain. Consequently, this new line of research promises to contribute to our understanding of human pain. As in most novel research areas, a first crucial step for integrating the scientific fields of ERN and pain is developing a paradigm suited to address the needs from both fields. The present manuscript presents the development and piloting of an experimental task measuring both ERN and avoidance behavior in response to painful mistakes, as well as the challenges encountered herein. A total of 12 participants underwent one of six different task versions. We describe in detail each of these versions, including their results, shortcomings, our solutions, and subsequent steps. Finally, we provide some advice for researchers aiming at developing novel paradigms.

KEYWORDS

error-related negativity, avoidance behavior, chronic pain, experimental psychopathology, experimental paradigm

Take-home message

Developing a new experimental paradigm is challenging, time-consuming and requires thorough testing. To make this process most efficient, we advise to clearly define the requirements at the beginning, to keep record of all decisions and adjustments made in the process, and to investigate encountered problems and failures which may help to gradually improve the paradigm.

Purpose

Avoidance behavior is part of a natural defense mechanism that is activated when individuals are experiencing chronic pain. In order to shed more light on the neural underpinnings of pain avoidance behavior, it is our long-term aim to investigate its association with the error-related negativity (ERN), an important component of neural error processing. So far, the ERN has been predominantly examined in relation to anxiety disorders and has been suggested as an endophenotype of obsessive-compulsive disorder (Riesel et al., 2011). Initially, ERN research focused on its relationship with the tendency

to learn best from negative feedback (Frank et al., 2005), whereas a few more recent studies also observed greater disorder-specific defensive motivation and avoidance behavior in individuals with elevated ERN amplitudes (Riesel, 2019; Riesel et al., 2019; Weinberg et al., 2016). These findings support the assumption that the ERN may be similarly implicated in pain-related avoidance, but this possibility remains to be tested. Investigating this theoretically plausible association may generally improve our understanding of the neural mechanisms underlying avoidance behavior and help to identify individuals at risk of developing chronic pain. The novelty of this endeavor requires the integration of knowledge of both research strands, as well as the development of an experimental paradigm that is well-suited to their respective requirements. Here, we describe the process of developing and piloting such a paradigm and provide an overview of problems encountered herein. We conclude with a tutorial section on the development of novel experimental paradigms based on our experience with this specific study, which we hope is also broadly applicable.

Introduction

Chronic pain affects as many as 20% of adults in Europe (of Pain, 2017). This poses a considerable health care challenge and causes tremendous individual suffering. Many chronic pain conditions are best understood from a biopsychosocial perspective which acknowledges the complex interplay of organic, cognitive, emotional, behavioral, and social factors. For example, on the neurobiological level, the sensitization model states that following acute pain, hypersensitization of local nerves occurs, causing exaggerated pain perception, or a lowered pain threshold, which may initiate a spiral in which hypersensitization triggers pain, which in turn triggers sensitization (Meeus & Nijs, 2007; van Wilgen & Keizer, 2012). Although the sensitization model is very intuitive, and often used as a metaphor in pain education with patients, it has also been criticized (e.g., van den Broeke & Van den Bergh, 2019). On the cognitive and behavioral level, one of the most prevalent explanations for the transition from a common acute pain episode to chronic disabling pain is the fear-avoidance model (Vlaeyen & Crombez, 2020; Vlaeyen et al., 2016; Vlaeyen & Linton, 2000). It posits that pain-related avoidance behavior, often fueled by catastrophic (mis)interpretations of pain, contributes to individuals entering a downward spiral of fear, avoidance, inactivity, disability, and negative affect. Pain-avoidance refers to individuals avoiding stimuli that are predictive of pain or pain exacerbations. For instance, persons who have sustained an injury to their back, may avoid lifting objects for fear that it will lead to further pain, or that lifting may create bodily harm that is signaled by pain. Typically, avoidance occurs in anticipation of pain, and hence individuals have little opportunity to test and, if necessary, correct their beliefs about the threat associated with such activities. This failure to correct mistaken beliefs maintains pain-related fear and further increases mood disturbances and pain itself (van Vliet et al., 2018; Vlaeyen & Linton, 2000).

The neural underpinnings of pain-related avoidance behavior are poorly understood. A new and promising avenue to understanding pain-related avoidance is to study its relation to error-related negativity (ERN). The ERN is an event-related potential (ERP), which occurs within ~100ms after error commission (Hajcak, 2012). It can be measured as a negative deflection in the electroencephalogram (EEG) over fronto-central scalp positions, originating

in the anterior cingulate cortex (Falkenstein et al., 2000; Gehring et al., 1993; Miltner et al., 2003; Taylor et al., 2007). Differences in the amplitude of the ERN are affected by both contextual and individual factors such as motivational aspects and error salience, and are considered to reflect individual error sensitivity (Hajcak, 2012; Hajcak & Foti, 2008; Taylor et al., 2007). For instance, it has been observed that the ERN is larger when errors are perceived to have worse consequences (Hajcak, 2012). Additionally, the ERN is believed to activate a defensive motivational system, in order to adjust the erroneous behavior and protect the organism (Hajcak & Foti, 2008). In line with these findings, increased ERN amplitudes have consistently been associated with anxiety disorders (Hanna et al., 2020; Weinberg et al., 2012) as persons affected arguably perceive their own mistakes in relation to the respective object of fear as highly threatening. In addition, avoidance behavior (including experiential avoidance) is a typical feature of anxiety disorders (Aupperle & Paulus, 2010; Berman et al., 2010; Dikman & Allen, 2000; Dymond & Roche, 2009).

Given the commonalities between anxiety disorders and chronic pain (Asmundson & Katz, 2009), and the high prevalence of negative affect in chronic pain (Geisser et al., 2000), it is theoretically plausible that the ERN¹ may be a biomarker for the development of pain-related avoidance behavior as well. In the case of anxiety, the particular feared object is considered threatening, whereas individuals with chronic pain perceive pain to be highly threatening. Because pain is such a salient and aversive stimulus, which conceptually signals bodily error and is often interpreted as such, avoidance responses constitute a defense mechanism to prevent the feared outcome (Leeuw et al., 2007). As such we may see an association between the ERN amplitude and the level of avoidance exhibited by an individual, with avoidance representing a defensive behavior in response to a motivationally salient and aversive stimulus.

Our long-term goal is to investigate whether individual differences in the ERN amplitude are related to differences in pain-related avoidance behavior, which may help identify individuals more at risk for the development of chronic disabling pain. To this end, we aimed to create an experimental paradigm that would allow us to measure the ERN and avoidance responses simultaneously, the process of which is described in the present manuscript. Previous experiments that studied this relationship in our lab (Traxler et al., in preparation) used the visual Flanker task unrelated to pain and had participants coming in on a separate day to measure pain avoidance behavior using a robotic arm (HapticMaster). A task integrating both the ERN and pain avoidance seems more efficient than measuring them in separate tasks. It also ensures that there is a relation between error commissions and pain-related avoidance, as it is for instance not clear how inhibition errors on a visual task relate to pain-related avoidance. Additionally, it puts less burden on the participant. Such a paradigm would need to meet requirements for both measures (ERN and pain-related avoidance), namely: (1) evoke an ERN; (2) yield a reliable estimate of the ERN, requiring a minimum of six errors (Olvet & Hajcak, 2009; Steele et al., 2016); and (3) be able to detect individual differences in avoidance behavior

1. Readers familiar with the error processing literature may wonder why we do not consider the Pe (error positivity), a later component of neural error processing. The reason is that we have no particular hypotheses about its relationship with pain-related avoidance, as it has not been as extensively studied in anxiety research as the ERN, especially with regard to avoidance behavior. We do aim to explore it in upcoming studies, but our main target of this report is the ERN, therefore we focused on developing a paradigm suitable to measure the ERN.

by providing sufficient opportunities to avoid (at least 20)². Although there is some debate as to whether one needs to be aware of having committed an error to evoke the ERN (Nieuwenhuis et al., 2001), awareness is essential in this task as participants would otherwise be unable to perform the avoidance response. In addition, the task ought to evoke inhibition errors rather than errors due to lack of skill or knowledge. For that purpose, we decided to integrate two experimental tasks, one of which is frequently used to measure the ERN (the Eriksen Flanker task; Eriksen & Eriksen, 1974) and the other is a common assessment of avoidance behavior (the avoidance task by Vervliet & Indekeu, 2015). We piloted this basic paradigm and continuously adjusted it based on the pilot results, leading to a total of six task variations. The present article focusses on the design and execution of these different tasks, what we can learn from them, and on suggestions for future research.

Methods

Participants

Twelve participants (8 females) took part in the pilot study between January and March of 2020. They were recruited by word of mouth, had a mean age of $M = 29.25$ ($SD = 10.64$), and gave autonomous consent. The exclusion criteria were self-reported: (a) pregnancy, (b) diagnosis of a psychiatric disorder (e.g. depression, anxiety, etc.), (c) a serious medical illness (e.g. diabetes mellitus, cardiovascular disease, neurological problems, etc.), (d) skin disease or condition where electrocutaneous stimuli could cause damage (e.g. an operation scar at the site where the electrodes would be placed), (e) pain: acute (e.g. due to an injury) or chronic, (f) electronic implant (e.g. pacemaker), (g) having been asked by a doctor to avoid stressful situations, and (h) poor vision or hearing that is not corrected. Participants gave written informed consent at the onset of the experimental session. Participants did not receive compensation for participation and mostly consisted of colleagues and acquaintances who were blinded with respect to the main aim of the tasks. The procedures conformed to the Helsinki Declaration and were approved by the institutional ethics committee (Social and Societal Ethics Committee; approval reference number: G-2019 12 1905).

Apparatus

Electroencephalography

Event-related potentials in response to errors and correct trials were measured by means of electroencephalography (EEG) which was continuously recorded from the scalp using a 129-channel system (HydroCel Geodesic Sensor Net, Philips Electrical Geodesics Inc., Eugene, USA). The sampling rate was set at 250 Hz (Hajcak & Foti, 2008; Hajcak et al., 2005; Trujillo & Allen, 2007), electrode impedances were kept below 50 k Ω and the vertex electrode was used as reference point (Sucec et al., 2019; Tan et al., 2019).

Stimulation

Vibrotactile stimuli were administered to the left side of participants' lower back through two vibrotactors (Dancer Design, St Helens, England) placed 2 cm apart (center to center; Figure 1). The stimulus duration was 112 ms and the vibration intensity, which was set to a clearly perceptible level that remained non-painful throughout the experiment, was kept constant across

participants. Tactors were activated in a semi-randomized order.

Electrocutaneous stimuli (e-stim) were generated by a constant-current stimulator (DS7A; Digitimer, Welwyn Garden City, England) and delivered for a duration of 2 ms through two 4mm Ag/AgCl reusable snap electrodes filled with K-Y gel attached 1 cm above the vibrotactors. The intensity of these stimuli was individually calibrated to a level that was "painful and demanding some effort to tolerate". For that purpose, a series of stimuli of ascending intensity was administered, each of which participants were instructed to verbally rate on a scale ranging from 0 (no sensation) to 10 (worst imaginable pain). The calibrated stimulus intensity was kept constant throughout the task.

The lower back was selected as stimulus location for two reasons: (1) the sensory acuity at the back is relatively low compared to more typical distal stimulation sites, such as the arms or hands (Weissman-Fogel et al., 2012), which was considered advantageous with regard to task difficulty; and (2) a large, even surface area was required to attach the vibrotactors and electrodes.

Software

The experimental task was programmed and presented in Affect 5 (Spruyt et al., 2009), which was run on a Windows 7 Professional (Microsoft Corporation Redmond) 64 bit Dell OptiPlex 780 (Dell Inc, Round Rock, TX) with 4-GB RAM, Duo-CPU at 3.10 GHz.

Measures

The purpose of the pilot study was to create a task that allows to measure the ERN and avoidance behavior simultaneously. In each task version vibrotactors emitted vibrations on participants' left lower back, EEG in response to correct and error trials was recorded, and painful e-stim were applied when participants made an erroneous response which they could omit by pressing the space bar. Avoidance was operationalized as the number of button presses to cancel an e-stim.

After performing the task participants were asked open questions to receive their perspective of the task and to ensure that requirements were met. Specifically, they were asked whether they were able to perceive when they made errors, how difficult they found the task as well as initiating the avoidance response, and any additional comments they would like to provide.

General procedure

All task versions were based on a tactile task that was inspired by Eriksen's Flanker task (Eriksen & Eriksen, 1974) and a fear-avoidance paradigm established by Vervliet and Indekeu (2015). The Eriksen Flanker task typically consists of presentations of arrow stimuli, (e.g. « > «), and participants are asked to indicate the direction in which the center arrow points by pressing one of two buttons as quickly as possible. The flanker arrows can be congruent

2. This is a somewhat arbitrary criterion reflecting the problem that too few error trials leave only little room for observing individual differences in avoidance. In the absence of empirical data regarding such a criterion, we based our decision on recent avoidance studies in our and others' labs that used acquisition phases of avoidance that usually consisted of around 20 trials. This makes it possible to detect avoidance patterns across participants (Meulders et al., 2016; Vervliet & Indekeu, 2015).

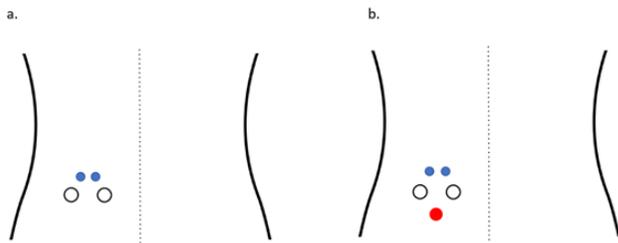


FIGURE 1 Electrode and tactor placements, with electrodes in blue. (a) Two tactor placement. (b) Three tactor placement with the red tactor indicating the distractor tactor.

or incongruent and thus potentially cause inhibition errors³. In Vervliet & Indekeu’s avoidance task, participants view an office space on a monitor, in which a desk lamp can take one of three colors. Each color is indicative of whether or not an e-stim will be delivered. Participants may press a button to avoid the e-stim and learn to avoid based on the color of the lamp.

In the tactile task participants were instructed to distinguish between locations of vibrations emitted on their left lower back. We opted for tactile stimuli because these stimuli are more relevant and ecologically valid for pain responses, and more strongly resemble the tactile and proprioceptive input that frequently precedes or co-occurs with pain. Much like the arrow version of the Eriksen Flanker task in which the direction of arrows is the cue for a correct response, participants used left and right mouse buttons to indicate whether they felt a vibration at the left or right location respectively. Participants were instructed to respond as quickly and as accurate as possible. In addition, participants were informed that upon error commission they would receive an electrocutaneous stimulus, which they could cancel by pressing the space bar on the keyboard. This discrimination task was used in all task versions. Tasks employed either two or three vibrotactors (Figure 1). The left and right tactors will be referred to as “target tactors”, as they emit the target stimuli that are to be discriminated by the participant. The third tactor is referred to as the “distractor tactor” as it is meant to make the discrimination task more difficult.

A typical trial proceeded as follows (see Figure 2): a vibration was emitted (112ms), followed by a response window, to indicate the location of the vibration (200-1000ms)⁴. After a fixation period (300ms), participants could provide an avoidance response (1000ms). If participants chose not to avoid the e-stim upon error commissioning, an e-stim was delivered following this response window. The trial concluded in a jittered intertrial interval (600-1000ms). To prevent ‘better safe than sorry’ reasoning for performing the avoidance response and to mimic the costs that real-life avoidance typically bears, participants were told that five, and later two, trials would be added to the duration of the experiment every time they cancelled an e-stim. In reality no trials were added. As this proved to be highly aversive to the participants, leading to an unwillingness to avoid, we removed this clause after the eighth participant. The base settings for the tasks are as follows: vibrations are set to the same intensity and the third (distractor) tactor coactivates with every target stimulation. Task versions replace one another rather than being superimposed. Any deviations from the general procedure will be indicated, with Table 1 giving an overview of the task versions. In general, the task versions were

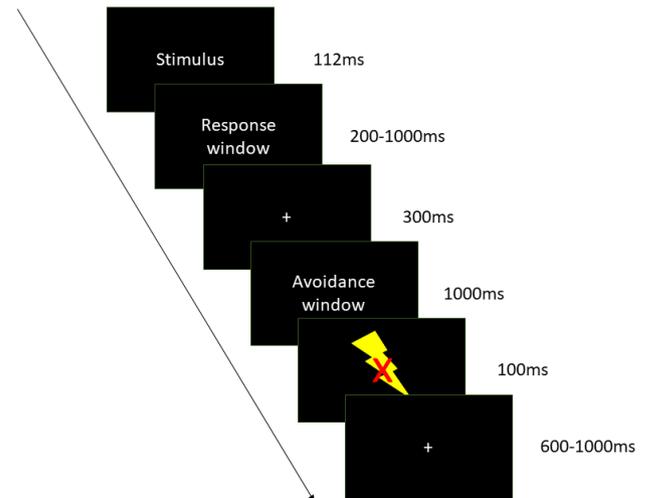


FIGURE 2 Schematic representation of the basic trial flow with exact timings depending on the task version.

composed of four or five blocks of 60 trials each, Table 2 indicates the total number of participants and trials per task version. In total, six different versions of the task were tested: *Two Tactor task*, *Distractor task*, *100% coactivation task*, *Different intensities task*, *50% coactivation task*, and a *Sequence task*.

Note. Number of tasks refers to whether participants had to perform any additional tasks to the discrimination task. *Tasks are listed in the order in which they were tested.*

Two Tactor task

Procedure

In this first version of the task, target tactors were attached to the left lower back of the participants (Figure 1, a). Participants indicated via left and right mouse button press whether they felt a vibration at the left or right location, respectively. Upon error commission participants were free to cancel the e-stim by pressing the space bar. One participant completed four blocks of this task version.

Evaluation

After running one participant it was quickly established that this task was too easy. This was reflected by participant feedback as well as the number of errors ($M_{error} = 2$). The participant was aware of having made mistakes but chose not to give any avoidance responses.

3. We seek to elicit inhibition errors, as this is how the ERN is traditionally measured (Ribes-Guardiola et al., 2020) with participants being aware of having committed such an error.

4. Non-responses or late responses did not bear any consequences as other studies conducted in our lab showed that participants do not simply cease to respond, even if this could be an alternative way of avoiding the painful stimulation.

Task	Number of Factors	Stimulation intensity	Number of tasks	When does the distractor factor activate?
Two Tactor	2	equal	1	NA
Distractor	3	equal	2	Plays a song
100% coactivation	3	equal	1	With every target stimulus
Different Intensities	3	3rd tactor more intense	1	With every target stimulus
50% Coactivation	3	equal	1	With 50% of target stimuli
Sequence	3	equal	1	With every target stimulus

TABLE 1 Overview of manipulations over task versions.

| Distractor task

Procedure

In order to increase task difficulty, three tactors were attached to the left lower back, with the two target tactors at their original locations and the third, the distractor tactor, placed centrally below them (Figure 1, b). This distractor tactor emitted vibrations to the beat of a widely known song (e.g., “Seven Nation Army” by The White Stripes). Participants were still expected to perform the same discrimination task (i.e. indicate on each trial whether the left or right tactor became active). Additionally, at the end of each block they had to choose, in a multiple choice manner, which song the distractor tactor had been vibrating to. One participant completed four blocks of 60 trials.

Evaluation

There were several issues with this task version. The task appeared to be difficult, leading the participant to be exclusively focused on the discrimination task ($M_{error} = 10$). Additionally, even though participants wore earplugs to prevent them from distinguishing vibrations based on the sound they made, the distractor tactor acted as a small speaker, and the participant became aware of the song because they could hear it. Due to these issues we did not run any further participants on this task.

| 100% coactivation task

Procedure

This task version used three tactors. The distractor tactor emitted vibrations simultaneously with the target tactors, in an effort to make their activations harder to discriminate. The first participant for this task completed four blocks; the subsequent six participants completed five blocks of this task. By increasing the number of blocks we sought to establish whether increasing task length would lead to more consistently elevated error commissions. Consistently higher error commissions across participants would ensure that less participants are excluded from analysis, were the paradigm to be used in the future. It would also enable us to make more reliable estimates of the ERN and individual differences in avoidance responses. One participant only completed two blocks of this task as they then completed two blocks of the 50% coactivation task. All but one participant were told that additional trials would be added if they chose to cancel the e-stim. In an attempt to balance the trade-off and increase the cost of not avoiding, the painful stimulation was made more threatening by

stating that a “stimulus with a slightly higher intensity than the one previously calibrated” could be delivered occasionally for four participants in the task (this manipulation was discontinued simultaneously with the added trials cost).

Evaluation

On average 22.5 mistakes were made across eight participants. Participant feedback indicated that the task was not too difficult and that they were conscious of having committed an error as it occurred.

| Different intensities task

Procedure

This task used the 100% coactivation schedule, with the distractor tactor set to a higher intensity than the target tactors. The goal, once more, was to make the task more difficult, in an attempt to raise the number of error commissions across participants. One participant performed five blocks of this task.

Evaluation

Although this task did raise the number of error commissions ($M_{error} = 33$), according to participant feedback, the participant was less aware of having made a mistake and therefore unable to avoid the e-stim.

| 50% coactivation task

Procedure

In this task the distractor tactor only coactivated with target tactors 50% of the time. This adaptation was expected to make the task more difficult as participants may habituate less to the sensation of the third tactor. Two participants completed respectively two and one block on this task.

Evaluation

This task proved to be easier than its 100% coactivation counterpart, with two participants making 18.5 mistakes on average. Participants reported having difficulty recognizing error commission, as such further task versions were explored.

| Sequence task

Procedure

In this last task, participants were instructed to recreate sequences of three vibrations using button presses. The target factors went off in a sequence (e.g. left-right-right), and participants recreated this using the left and right mouse buttons. The distractor factor followed a 100% coactivation schedule, i.e. coactivated with the target vibrations on each trial. Two participants performed five blocks of the task. We expected that this task would be more difficult as it required a more speeded response from the participants. Additionally, we thought it might lead to more inhibition errors: In the arrow version of the Flanker task, incongruent trials (i.e., those in which the middle arrow points in the opposite direction of the flanking arrows) tend to evoke more errors. Similarly, on the present task, two left activations, for instance, may mislead the participant to expect another left activation, instead of a right.

Evaluation

This task provided inconsistent results as the deviations in the number of error commissions varied widely between participants (pilot_a: $M_{error} = 73$, pilot_b: $M_{error} = 12$). Though individual differences in performance are to be expected in any task, the task should be reliable in evoking the needed number of error commissions. Participants were aware of their error commissions.

| Statistical Analysis

The EEG data was processed in Brain Electrical Source Analysis Research 6.0 (BESA GmbH, Gräfelfing, Germany), using a high-cut filter of 30 Hz, a low-cut filter of 0.1 Hz and a notch filter of 50 Hz (Gorka et al., 2019). Ocular artifact removal was performed using the BESA algorithm, which does not require the measurement of an additional electrooculogram (EOG), but uses fronto-facial sensors to detect ocular movements. Bad channels were interpolated or deleted upon visual inspection, with a maximum of 12 bad channels [10% of the total number of electrodes; (Keil et al., 2014)]. Response-locked epochs of 1500 ms (500 ms pre- and 1000 ms post-response) were extracted and averaged per participant, using the 500 - 300 ms pre-response interval as baseline (Gorka et al., 2019; Jackson et al., 2015; Tan et al., 2019). Data were re-referenced to the average reference.

The ERN was operationalized as the mean amplitude in the 0 to 100 ms time window after error commission at the fronto-central site FCz (Geodesic net electrode 6) (Hajcak et al., 2019). In addition, the correct response negativity (CRN) - a similar yet smaller negative deflection in the EEG signal following correct responses (0 - 100 ms post-response) - was computed in the same way. The difference score between the ERN and the CRN, known as the Δ ERN, will be reported. Previous research suggests that the validity of the Δ ERN is higher than that of the mean amplitude of the ERN alone (Gorka et al., 2019; Riesel et al., 2013). Both the preprocessed EEG and behavioral data were analyzed using R (Team, 2013). Due to the small group sizes across tasks and resulting insufficient statistical power to make reliable inferential conclusions, the analysis consists mostly of descriptive statistics.

Results

EEG data were recorded for 10 of the 12 participants, including both *Sequence task* pilots, the *Different intensities* pilot, one *50% coactivation* pilot, and six *100% coactivation* pilots. One participant on the 100% coactivation task committed less than six errors and was therefore excluded from the EEG analyses. No EEG data were recorded for the *Distractor* task or *Two Tactor* task. Given the small sample size, the results presented below need to be read with caution and should only be evaluated with regard to the pre-specified task criteria.

| Task performance and avoidance behavior

On average, participants committed 24.07 errors ($SD = 18.33$, $range = 0-73$, 10.59% of all accepted responses) and gave 203.21 correct responses ($SD = 98.82$, $range = 1-297$) on the discrimination task. A Related Samples Wilcoxon Signed Rank Test revealed that reaction times were significantly faster during correct trials ($M = 527.35$ ms, $SD = 88.39$) compared to error trials ($M = 619.53$ ms, $SD = 114.75$; $Z = 3.059$, $p = .002$).

Table 2 shows the results for the behavioral responses. The *Two Tactor* task yielded few mistakes ($M = 2$, $SD = 0$) and no avoidance responses. The *100% coactivation* task produced more error commissions ($M = 22.5$, $SD = 20.61$), and some avoidance responses ($M = 7.12$, $SD = 16.25$). The *Different intensities* task version led to a higher number of mistakes ($M = 33$, $SD = 0$), but did not produce any avoidance responses. Moving on to the *50% coactivation* task, we observe a comparable average error commission as for the *100% coactivation* task ($M = 18.5$, $SD = 19.09$). However, the frequency of avoidance is low ($M = 0.5$, $SD = 0.7$). This might indicate either that participants are unaware of having committed an error and thus are unable to avoid, or that they were unwilling to bear the costs introduced with this task version. The *Sequence* task produced the highest average error commissions ($M = 42.5$, $SD = 43.13$), but these scores are highly inconsistent across participants. The same applies to the avoidance responses in this task ($M = 27.5$, $SD = 19.09$). The *Distractor* task shows only $M = 10$ error commissions ($SD = 0$) and no avoidance responses. Note that results for each task version are based on unequal and few data points (Table 2).

We further compared when the most error commissions occurred in each task, in order to determine whether mistakes are more likely to be attributable to the novelty of the task or whether the task manages to consistently evoke error commissions. Across tasks, 56% of errors occurred in the first two blocks of the experiment. Note that for participants for whom the total number of trials was lower than four blocks, only the first block was considered. The task that evoked errors the most consistently throughout the blocks was the 100% coactivation task (45.8% of errors in the first two blocks) and the least was the *Different intensities* task (81.8% of errors in the first two blocks).

| Error-related negativity

Figure 3 shows the response-locked ERPs and topographies for error and correct responses across the four task versions (*Sequence*, *Different intensities*, *50% coactivation*, *100% coactivation*) separately. All four tasks do show the potential for evoking an ERN with negative going deflections being observed at the expected fronto-central locations. Across all tasks, the average difference

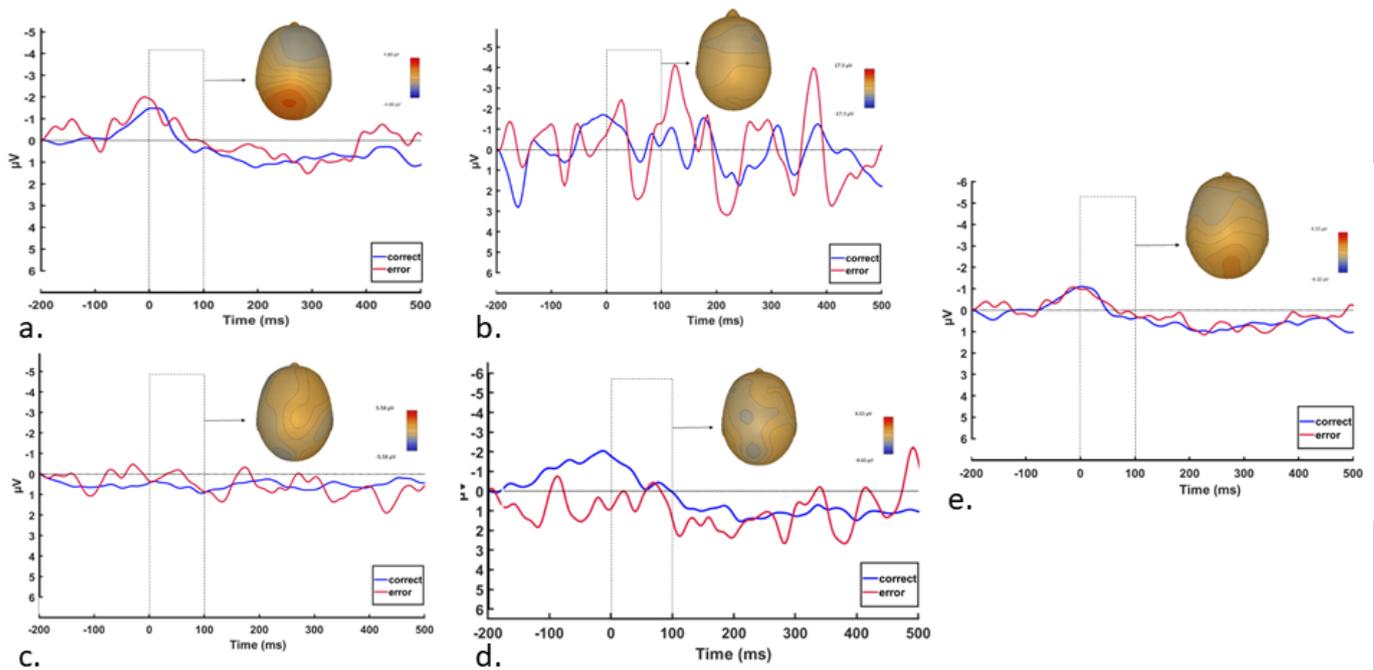


FIGURE 3 Response-locked grand average waveforms for ERN and CRN at Cz with topography, averaged for (a) the 100% coactivation task (N = 5); (b) the 50% coactivation task (N = 2); (c) the Sequence task (N = 1); (d) the Different intensities task (N = 1); and (e) averaged across the four task versions (N = 9).

Task	N	Total		Errors			Mean		Peak (ms)
		trials	M	SD	Range	Avoided	Δ ERN		
Two Factor	1	240	2	0	2	0	NA	NA	
Distractor	1	240	10	0	10	0	NA	NA	
100% Coactivation	8	2100	22.5	20.61	0-64	7.12	-3.97	0	
Different Intensities	1	300	33	0	33	0	-0.43	68	
50% Coactivation	2	180	18.5	19.09	5-32	0.5	-0.02	28	
Sequence	2	600	42.5	43.13	12-73	27.5	-0.05	52	

TABLE 2 Results of behavioral and EEG data

score between ERN and CRN was Δ ERN = -2.27 μ V. The task showing the most pronounced Δ ERN at FCz is the 100% coactivation task (Δ ERN= -3.97 μ V) and the least pronounced Δ ERN was observed in the 50% coactivation task (Δ ERN= -0.02 μ V). Though Δ ERN scores are variable across studies (Barker et al., 2015; Buzzell et al., 2017; Gorka et al., 2019), our results do suggest smaller difference scores than observed using traditional reaction time tasks. Importantly, given the small sample size for each task, no comparative conclusions can be drawn with regards to the ERN and its relation to pain-related avoidance responses across tasks.

Discussion

Here we describe the piloting of a novel paradigm based on the integration of two experimental tasks stemming from different fields of psychology. Our aim was to develop an instrument suited to measure both the neural activity during error commission that is followed by a negative consequence, namely pain, and the avoidance tendencies in direct response to these pain-evoking errors. As described above, this task is required to evoke an ERN and allow for its reliable measurement (≥ 6 errors; Olvet & Hajcak, 2009; Steele et al., 2016), evoke conscious errors of inhibition in order to allow participants to avoid the painful stimulation, and create sufficient variability in avoidance behavior (\geq potential avoidance trials, i.e., error trials, and adequate balance between costs and benefits of avoiding vs. not avoiding). Altogether, six different task versions were put to the test, which are discussed in the following.

The first task using only two factors proved to be too easy, based on the low number of mistakes made and the reported feedback of the participant. Despite the low perceptual acuity at the lower back, the vibrations were reportedly clearly distinguishable and after a first orientation during the practice phase, the participant had no difficulty completing the task. Given the small number of error commissions, this task did not allow for a reliable measurement of the ERN. With the aim of increasing task difficulty and thus the likelihood of error commissions, it was decided to add a third vibrotactor to the paradigm to distract from the target stimulation, which resembles the distracting flankers of the Eriksen Flanker task (Eriksen & Eriksen, 1974).

The Distractor task, the first version using the distractor factor, was programmed to emit vibrations to the rhythm of famous songs which the participant

was instructed to identify whilst responding to the left and right vibrations. This task turned out highly demanding, leaving the participant focusing entirely on the discrimination task. Although this version successfully elicited a considerable number of errors, according to the participant the task load interfered with performing the avoidance response, which was perceived as disruptive, such that none of the e-stim were avoided.

In order to decrease the demands of the task, we switched to the *100% coactivation task*, in which the distractor factor coactivated with the target factor on each trial. On average, participants made a sufficient number of errors for ERN analysis, as well as showing a reasonable variability in avoidance responses. Notably, the introduction of different types of avoidance costs during this task version led to a strong decline in avoidance responses and was therefore discontinued. While this version in principle did meet the pre-specified requirements, we decided to further modify this task in an attempt to ensure that the majority of participants would commit sufficient errors.

First, the *vibration intensity* of the distractor factor was slightly increased with the aim to better mask the vibrations of the target factors. This led to the participant being unsure about their performance and thus unable to perform the avoidance behavior. Second, after setting the vibration intensity of the distractor factor back to its original level, we changed the *coactivation to 50%* of the trials with the goal of increasing unpredictability. Contrary to our expectations, participants experienced this version as easier than the 100% coactivation version and made fewer mistakes.

Lastly, in the *Sequence task*, participants were asked to perform a sequence of responses repeating after the vibrotactors. Given the speed of the task, it was hypothesized that participants would be easily led into committing errors of inhibition. This approach was successful for one participant; nevertheless, it appeared that the activation of both target factors within one trial helped the other participant to orient towards the stimulation locations, causing a large difference in the amount of error commissions.

Overall, the tasks succeeded in eliciting an ERN at fronto-central sites when participants committed errors, which resembled the typical ERN as observed in more classic tasks such as the Eriksen Flanker task (Imburgio et al., 2020). Bearing in mind that the reported results are based on small sample sizes, it is noteworthy that the ERNs produced by the different task versions were slightly less pronounced and some peaked earlier than typically found with the Flanker task (Hajcak & Foti, 2008). Moreover, whereas commonly reaction times for error trials are found to be shorter than those for correct trials (Aarts & Pourtois, 2010; Imburgio et al., 2020), the opposite pattern was found in all of the piloted versions of this paradigm. However, a study that compared the ERN elicited by errors on the visual Flanker task with an interoceptive discrimination task using respiratory stimuli reported that the interoceptive ERN (intERN) did peak significantly earlier than the exteroceptive ERN (Tan et al., 2019). The authors of this article argue that interoceptive stimuli might be more relevant for survival, making the detection of interoceptive errors a priority. It is conceivable that somatosensory errors which indicate proximal threat carry this relevance, too, and thus evoke a similarly early peaking ERN, allowing for fast adaptation of behavior. A possible explanation for the less pronounced waveforms is that participants were less aware of having made a mistake than is commonly the case on the visual Flanker task. In addition,

deviations of the present ERN waveforms from those observed in previous studies are most likely also related to the small number of averaged trials and/or participants (Kappenman & Luck, 2011) and therefore need to be interpreted with caution. Future well-powered studies will have to show whether these results are reliable and clarify the origins of potential differences between the processing of interoceptive and exteroceptive errors.

In conclusion, out of the six versions tested, the *100% coactivation* task proved the most reliable and useful for investigating whether individuals with larger ERN amplitudes are more prone to engage in pain-related avoidance behavior. Nevertheless, this task does have several limitations: *Firstly*, error commissions may remain low for a fair number of participants, so that it can be expected that a study employing this paradigm will lose many participants on the exclusion criterion of 20 errors that are needed to obtain a useful assessment of avoidance behavior. Besides, as stated above, this criterion is somewhat arbitrary and future pilots will have to show if this number of error trials is sufficient to detect individual differences in pain avoidance. Hence, it is possible that even more errors would be required. *Secondly*, the operationalization of avoidance behavior as a single button press may be considered simplistic. Persons with chronic pain tend to engage in a broad variety of frequently subtle alternative behaviors and withdrawal from activities that are thought to evoke or worsen their pain (Asmundson et al., 1999; Volders et al., 2015). Capturing this complexity in an experimental paradigm is highly challenging, and even though better approximations than button presses do exist (e.g., Meulders et al., 2016; Claes et al., 2014), the constraints of the Flanker task, a speeded reaction time task, do not allow for longer, more complex avoidance behaviors in between trials. Even though other tasks can be used to measure the ERN apart from the Flanker task, such as the Stroop task and the Go/No-Go task (Riesel et al., 2013), the constraints of a speeded response task remain and would not allow for the integration of such avoidance responses either. *Thirdly*, in line with this limitation, engaging in avoidance behavior bears no cost in this paradigm, which causes a discrepancy with real-life avoidance where patients are faced with considerable costs to their social life, physical functioning, and personal well-being (Martinez-Calderon et al., 2019; Prkachin et al., 2007; Vlaeyen & Linton, 2012; Wiech & Tracey, 2013). Yet, when exploring options to add an avoidance cost, in the trade-off between these costs on the one hand and receiving a painful stimulus on the other, participants clearly favored the latter: Participants were instructed that engaging in avoidance behavior would have the consequence of trials being added at the end of the task. As this did not only prolong the duration of the experiment but also further increased the likelihood of error commission, participants appeared to prefer to endure the e-stim. In an attempt to balance the trade-off and increase the cost of not avoiding, the painful stimulation was made more threatening by stating that a "stimulus with a slightly higher intensity than the one previously calibrated" could be delivered occasionally. Surprisingly, this change also did not increase the number of avoidance responses. Adjusting the painful stimulus in length or quality was not feasible given the timing of the trial flow. Based on the low levels of avoidance behavior throughout the pilots, possibly due to the cognitive demand of switching from the discrimination task to avoiding the painful stimulus, a surge in avoidance responses is not expected. Moreover, low-cost clinical avoidance, such as constantly carrying pain medication, is

common and may be problematic as it impedes treatment and maintains fear (van Vliet et al., 2018; Vervliet & Indekeu, 2015; Volders et al., 2012). Hence, studying these low-cost avoidance behaviors is clinically relevant, and it was decided not to integrate any further avoidance costs.

Apart from these task-specific limitations, two limitations of the piloting process need to be mentioned: On the one hand, the small and unequal sample sizes for the different task versions do not allow extensive statistical comparisons. Some decisions were based on just one or two participants who could be outliers, rendering the conclusions invalid. Even though the impressions gained from the pilot participants' experiences on the Two Tactor and the Distractor tasks were supported by informal pilots conducted on the authors themselves, more pilots per task version would be necessary to provide certainty on the utility of each version. Given that the aim of this study was not to test hypotheses and that a large number of within-subject trials was used, we believe that our observations are nonetheless informative. On the other hand, some changes to the paradigm were made simultaneously, making it impossible to disentangle the individual effects. The decisions to further modify the paradigm were made as soon as clear converging evidence was obtained from behavioral performance and self-report that a given version was either too easy or too complex, often after just one or two pilots, in order to save time and participants. While this allowed us to try out various ideas and establish a functional paradigm within a few weeks' time, the paradigm clearly requires validation in future studies.

Altogether, developing this experimental task based on the integration of two well-established paradigms was demanding, despite the relative simplicity of the original paradigms. Although the Flanker and the avoidance task seemed compatible on account of their short response intervals and general set-up, many problems arose with regard to timing, task difficulty, and interference of responses, and possible solutions often caused unexpected side-effects to participants' behavior, such as the introduction of an avoidance cost preventing participants from avoiding altogether. Therefore, a thorough and extensive piloting was (and remains) pivotal in identifying various pitfalls and re-adjusting the task. While the 100% coactivation task seems the most promising, we aim to conduct a second pilot study solely focusing on this task version to ensure (a) that our criteria are consistently met, and (b) that the criterion of minimally 20 error commissions is sufficient to observe inter-individual variability in avoidance, before implementing this paradigm in a larger study. In the following section, we have compiled a list of recommendations based on our experiences for those planning to create a novel experimental paradigm.

Tutorial

This pilot study was an interesting, highly instructive journey. Based on this experience, we propose five key elements to consider when developing a new experimental paradigm or merging existing ones.

1. **Clarity:** Prior to designing the experimental paradigm, it is advised to clearly define the requirements the paradigm ought to meet. One should consider any known necessities or restrictions inherent to the subject or to the original tasks. Relevant questions in this step are “What is the ultimate goal of this paradigm?”, “How do I operationalize the independent and outcome variables?”, “What does this operationalization require in terms of equipment and measurement?” and “Which other variables need to be controlled for?”.
2. **Thoroughness:** Developing a novel experimental paradigm, even if it seems straightforward, requires thorough piloting, perhaps of several different approaches, and is often a process of trial and error. Keeping a detailed record of decisions, adjustments, and their motivation as well as problems encountered, is crucial to understand what works and what does not, and to compare different versions of a paradigm. We recommend considering the use of an open-source repository, for example the OSF platform (<https://osf.io/>), which may encourage researchers to be more transparent in their decision-making. This would not only be informative for colleagues but also benefit the researchers themselves by stimulating the a-priori formulation of well-founded arguments. Furthermore, it is advised to check all relevant data as seemingly small adjustments of a task may have major effects on participants' perception and behavior, as illustrated by the apparent drop in avoidance responses upon the introduction of an associated cost in the paradigm described above. Next to documentation, transparency, and data checking, another important aspect of thoroughness is to implement changes subsequently rather than simultaneously, as this allows one to disentangle the respective effects. While we largely adhered to these principles, we introduced the avoidance cost at the first pilot of the 100% coactivation task and terminated both the cost and the pain threat at the same time. Moreover, we did not consistently use the same number of blocks throughout the process, which makes the comparison of task versions more difficult. In hindsight, these changes were hasty, and we eventually reverted to the pure version of the task without cost and increased pain threat.
3. **Curiosity:** When it becomes clear what does not work, the follow-up question should be “Why does it not work?” Exploring the problem more closely may lead to its solution and to an improvement of the paradigm. Thus, it may prove useful to include additional measures during the pilots: For example, in studies with human subjects the researchers may choose to use a post-experimental questionnaire that addresses participants' experience of the task as a whole and of any manipulations, as well as their understanding of instructions and materials. An example from the current pilot study is the avoidance cost that was introduced to better reflect real-life avoidance: We initially expected that participants would readily avoid the painful stimulus and hence decided to use penalties to avoid a ceiling effect. However, by adding trials for each avoidance response, we apparently did not meet a good balance between the pain threat and the avoidance cost as it almost entirely prevented participants from avoiding, even after increasing the pain threat. The post-experimental questions revealed that participants experienced the cost as too aversive as it would prolong the duration of the experiment even further. This insight led us to discontinue the use of this avoidance cost. Preparing such questions ahead of the pilots, asking them systematically and adjusting them based on new insights may provide valuable information from a different perspective. Again, it may be valuable to share these experiences on OSF, which may aid others in replicating, extending or modifying the paradigm.
4. **Creativity & Collaboration:** Sometimes, a problem in the paradigm cannot be easily fixed. Researchers are well-advised to seek inspiration in directly

and indirectly related literature, to reach out to or work together with researchers experienced in the respective domains, and to test the paradigm both on themselves and colleagues to broaden their perspective. Hence, we made sure to first test each task version on ourselves, and shared our ideas and experiences at lab meetings, which inspired some of the task versions. Moreover, joining forces with other labs may facilitate the testing of multiple task versions simultaneously, allowing for larger samples to be tested.

5. Perseverance: Developing new paradigms or integrating existing ones takes time. We strongly advise to allow for both sufficient time and sufficient participants to run several pilots, especially if a version of the paradigm seemingly meets all requirements. A well-tested paradigm will minimize the risks during data collection. In the present case, due to limited resources we opted for small samples for each task version to establish their potential feasibility and utility. Yet, as a next step, we hope to conduct a more extensive pilot study on the most promising task (100% coactivation) in a larger sample ($N = 10$) before implementing it in a hypothesis-testing study aimed at investigating our initial research question. In fact, whereas a power calculation for a novel paradigm is likely to be difficult, it may be advisable to determine a minimum number of pilot participants a priori.

Conclusion

Developing a novel experimental paradigm is a challenging endeavor that necessitates good preparation and thorough evaluation through extensive piloting. We advise to clearly specify the task requirements in advance, record all problems and applied solutions, and remain open-minded for unconventional approaches and inspirations from various sources. As with conventional research projects, the development of a novel paradigm may greatly benefit from using an open source repository.

Peer review

JOTE aims to make the peer review process accessible to its readers. Therefore, the initial submission with integrated peer review comments is available [here](#).

References

- Aarts, K., & Pourtois, G. (2010). Anxiety not only increases, but also alters early error-monitoring functions. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(4), 479–492. <https://doi.org/10.3758/CABN.10.4.479>
- Asmundson, G. J., & Katz, J. (2009). Understanding the co-occurrence of anxiety disorders and chronic pain: State-of-the-art. *Depression and Anxiety*, *26*(10), 888–901. <https://doi.org/10.1002/da.20600>
- Asmundson, G. J., Norton, P. J., & Norton, G. R. (1999). Beyond pain: The role of fear and avoidance in chronicity. *Clinical Psychology Review*, *19*(1), 97–119. [https://doi.org/10.1016/S0272-7358\(98\)00034-8](https://doi.org/10.1016/S0272-7358(98)00034-8)
- Aupperle, R. L., & Paulus, M. P. (2010). Neural systems underlying approach and avoidance in anxiety disorders. *Dialogues in Clinical Neuroscience*, *12*(4), 517–531. <https://doi.org/10.31887/DCNS.2010.12.4.raupperle>
- Barker, T. V., Troller-Renfree, S., Pine, D. S., & Fox, N. A. (2015). Individual differences in social anxiety affect the salience of errors in social contexts.

- Cognitive, Affective, & Behavioral Neuroscience*, *15*(4), 723–735. <https://doi.org/10.3758/s13415-015-0360-9>
- Berman, N. C., Wheaton, M. G., McGrath, P., & Abramowitz, J. S. (2010). Predicting anxiety: The role of experiential avoidance and anxiety sensitivity. *Journal of Anxiety Disorders*, *24*(1), 109–113. <https://doi.org/10.1016/j.janxdis.2009.09.005>
- Buzzell, G. A., Richards, J. E., White, L. K., Barker, T. V., Pine, D. S., & Fox, N. A. (2017). Development of the error-monitoring system from ages 9–35: Unique insight provided by MRI-constrained source localization of EEG. *Neuroimage*, *157*, 13–26. <https://doi.org/10.1016/j.neuroimage.2017.05.045>
- Claes, N., Karos, K., Meulders, A., Crombez, G., & Vlaeyen, J. W. S. (2014). Competing goals attenuate avoidance behavior in the context of pain. *Journal of Pain*, *15*(11), 1120–1129. <https://doi.org/10.1016/j.jpain.2014.08.003>
- Derksen, M. (2020). Reflection on Trial and Error (-related negativity). *Journal of Trial and Error*, *1*(1), 39–42. <https://doi.org/10.36850/r2>
- Dikman, Z. V., & Allen, J. J. B. (2000). Error monitoring during reward and avoidance learning in high- and low-socialized individuals. *Psychophysiology*, *37*(1), 43–54. <https://doi.org/10.1111/1469-8986.3710043>
- Dymond, S., & Roche, B. (2009). A contemporary behavior analysis of anxiety and avoidance. *The Behavior Analyst*, *32*(1), 7–27. <https://doi.org/10.1007/BF03392173>
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. <https://doi.org/10.3758/BF03203267>
- Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: A tutorial. *Biological Psychology*, *51*(2-3), 87–107. [https://doi.org/10.1016/S0301-0511\(99\)00031-9](https://doi.org/10.1016/S0301-0511(99)00031-9)
- Frank, M. J., Woroach, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, *47*(4), 495–501. <https://doi.org/10.1016/j.neuron.2005.06.020>
- Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*(6), 385–390. <https://doi.org/10.1111/j.1467-9280.1993.tb00586.x>
- Geisser, M. E., Roth, R. S., Theisen, M. E., Robinson, M. E., & Riley, J. L. (2000). Negative Affect, Self-Report of Depressive Symptoms, and Clinical Depression: Relation to the Experience of Chronic Pain. *The Clinical Journal of Pain*, *16*(2), 110–120. <https://doi.org/10.1097/00002508-200006000-00004>
- Gorka, S. M., Lieberman, L., Kreutzer, K. A., Carrillo, V., Weinberg, A., & Shankman, S. A. (2019). Error-related neural activity and alcohol use disorder: Differences from risk to remission. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *92*, 271–278. <https://doi.org/10.1016/j.pnpbp.2019.01.011>
- Hajcak, G. (2012). What we've learned from mistakes: Insights from error-related brain activity. *Current Directions in Psychological Science*, *21*(2), 101–106. <https://doi.org/10.1177/0963721412436809>
- Hajcak, G., & Foti, D. (2008). Errors are aversive: Defensive motivation and the error-related negativity. *Psychological Science*, *19*(2), 103–108. <https://doi.org/10.1111/j.1467-9280.2008.02053.x>

- Taylor, S. F., Stern, E. R., & Gehring, W. J. (2007). Neural systems for error monitoring: Recent findings and theoretical perspectives. *The Neuroscientist*, 13(2), 160–172. <https://doi.org/10.1177/1073858406298184>
- Team, R. C. (2013). *R: A language and environment for statistical computing*. <https://www.r-project.org/>
- Trujillo, L. T., & Allen, J. J. (2007). Theta EEG dynamics of the error-related negativity. *Clinical Neurophysiology*, 118(3), 645–668. <https://doi.org/10.1016/j.clinph.2006.11.009>
- van den Broeke, E. N., & Van den Bergh, O. (2019). Central sensitization in humans: Popular phrase or useful concept? *Journal of Psychosomatic Research*, 119, 51–52. <https://doi.org/10.1016/j.jpsychores.2019.02.007>
- van Vliet, C. M., Meulders, A., Vancleef, L. M. G., & Vlaeyen, J. W. S. (2018). The opportunity to avoid pain may paradoxically increase fear. *The Journal of Pain*, 19(10), 1222–1230. <https://doi.org/10.1016/j.jpain.2018.05.003>
- van Wilgen, C. P., & Keizer, D. (2012). The sensitization model to explain how chronic pain exists without tissue damage. *Pain Management Nursing*, 13(1), 60–65. <https://doi.org/10.1016/j.pmn.2010.03.001>
- Vervliet, B., & Indekeu, E. (2015). Low-cost avoidance behaviors are resistant to fear extinction in humans. *Frontiers in Behavioral Neuroscience*, 9, Article 351. <https://doi.org/10.3389/fnbeh.2015.00351>
- Vlaeyen, J. W. S., & Crombez, G. (2020). Behavioral conceptualization and treatment of chronic pain. *Annual Review of Clinical Psychology*, 16(1), 187–212. <https://doi.org/10.1146/annurev-clinpsy-050718-095744>
- Vlaeyen, J. W. S., Crombez, G., & Linton, S. J. (2016). The fear-avoidance model of pain. *Pain*, 157(8), 1588–1589. <https://doi.org/10.1097/j.pain.0000000000000574>
- Vlaeyen, J. W. S., & Linton, S. J. (2000). Fear-avoidance and its consequences in chronic musculoskeletal pain: A state of the art. *Pain*, 85(3), 317–332. [https://doi.org/10.1016/S0304-3959\(99\)00242-0](https://doi.org/10.1016/S0304-3959(99)00242-0)
- Vlaeyen, J. W. S., & Linton, S. J. (2012). Fear-avoidance model of chronic musculoskeletal pain: 12 years on. *Pain*, 153(6), 1144–1147. <https://doi.org/10.1016/j.pain.2011.12.009>
- Volders, S., Boddez, Y., De Peuter, S., Meulders, A., & Vlaeyen, J. W. S. (2015). Avoidance behavior in chronic pain research: A cold case revisited. *Behaviour Research and Therapy*, 64, 31–37. <https://doi.org/10.1016/j.brat.2014.11.003>
- Volders, S., Meulders, A., De Peuter, S., Vervliet, B., & Vlaeyen, J. W. (2012). Safety behavior can hamper the extinction of fear of movement-related pain: An experimental investigation in healthy participants. *Behaviour Research and Therapy*, 50(11), 735–746. <https://doi.org/10.1016/j.brat.2012.06.004>
- Weinberg, A., Meyer, A., Hale-Rude, E., Perlman, G., Kotov, R., Klein, D. N., & Hajcak, G. (2016). Error-related negativity (ERN) and sustained threat: Conceptual framework and empirical evaluation in an adolescent sample. *Psychophysiology*, 53(3), 372–385. <https://doi.org/10.1111/psyp.12538>
- Weinberg, A., Riesel, A., & Hajcak, G. (2012). Integrating multiple perspectives on error-related brain activity: The ERN as a neural indicator of trait defensive reactivity. *Motivation and Emotion*, 36(1), 84–100. <https://doi.org/10.1007/s11031-011-9269-y>
- Weissman-Fogel, I., Brayer-Zwi, N., & Defrin, R. (2012). Spatial resolution of the pain system: A proximal-to-distal gradient of sensitivity revealed with psychophysical testing. *Experimental Brain Research*, 216(2), 181–190. <https://doi.org/10.1007/s00221-011-2924-4>
- Wiech, K., & Tracey, I. (2013). Pain, decisions, and actions: A motivational perspective. *Frontiers in Neuroscience*, 7, Article 46. <https://doi.org/10.3389/fnins.2013.00046>

Companion Piece

Abstract. ‘Trial and Error (-related negativity)’ is a fascinating paper detailing the attempt to develop a new experimental paradigm to study the role of error-related negativity in the development of avoidance behavior. In my comments on this paper I will focus on the interaction between experimenters and participants as the former investigate various ways of designing the experiment, aiming to elicit the right kind of behavior from the participants. As in many psychological experiments, there is a fundamental tension here that experimenters must find a way to deal with: they must guide the subject to the proper performance, without the subject responding to the guidance as such. The performance must be natural, but within tight constraints. Recalcitrance or resistance of the subject must be prevented. Ultimately, the authors of ‘Trial and Error (-related negativity)’ failed in their attempt to do this. Their reflections on their failure are thorough and illuminating, but I will argue that they can be pushed slightly further.

Derksen, M. (2020). Reflection on Trial and Error (-related negativity). *Journal of Trial and Error*, 1(1), 39–42. <https://doi.org/10.36850/te>

License



Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020