Startle modulation by heat pain with varying threat levels in chronic pain patients and pain-free controls

C Horn-Hofmann1
D Wolf1
S Wolff1
M Heesen2
K Knippenberg-Bigge3
P M Lang3
S Lautenbacher1

1Department of Physiological Psychology, University of Bamberg, Germany; 2Department of Anesthesiology and Pain Therapy, Kantonsspital Baden, Switzerland; 3Department of Anesthesiology and Pain Therapy, Sozialstiftung Bamberg, Bamberg, Germany

Background: Empirical evidence suggests that affective responses to pain are changed in chronic pain. The investigation of startle responses to pain might contribute to clarifying whether such alterations also expand to motivational defensive reactions. We aimed at comparing startle responses to tonic heat pain with high threat (HT) or low threat (LT) in patients with chronic musculoskeletal pain and controls. As pain-related anxiety and catastrophizing are typically elevated in chronic pain, we expected to find stronger startle responses in patients specifically under experimental HT.

Methods: Patients with chronic musculoskeletal pain, preferentially, back pain (N = 19) and matched pain-free controls (N = 19) underwent two pain-related threat conditions (high and low) in balanced order. Only, in the HT condition, 50% of the trials were announced to include a short further noxious temperature increase at the end. Startle responses to loud tones were always assessed prior to a potential temperature increase in the phase of anticipation and were recorded by surface electromyogram.

Results: Surprisingly, we observed no differences in startle responses and ratings of emotional and pain responses between patients and controls despite significantly higher pain-related anxiety and catastrophizing in the patients. Overall, startle was potentiated in the HT condition, but only in participants who started with this condition.

Conclusion: Our results suggest that, in general, patients with pain are not more responsive emotionally to experimental threat manipulations despite elevated pain anxiety and catastrophizing. Instead, exaggerated responses in patients might be triggered only by individual concerns relating to pain, which are not sufficiently mirrored by our threat paradigm.

Keywords: musculoskeletal pain, catastrophizing, anxiety, defensive motivation, experimental pain, startle reflex, threat induction

Introduction

Chronic pain is associated with considerable emotional distress which becomes obvious by the high comorbidity with affective disorders.1 Relating to this, it has been shown that compared to healthy controls, patients with pain report higher levels of pain-related anxiety and catastrophizing2,3 and show enhanced activation in emotional brain circuits during noxious stimulation.4–6 These empirical findings suggest that emotional pain processing is altered in patients with chronic pain conditions.

Given these differences in affective responding, the question arises whether the motivational defensive component of pain – that is, the urge to escape the noxious stimulus – is also altered in chronic pain. Such alterations in motivational defensive processes might – although being protective against harm due to acute pain – contribute to the development of maladaptive avoidance behavior in chronic pain if the defensive
system is activated too easily and strongly. One of the most established paradigms for measuring motivational defensive states is the startle paradigm: The startle reflex is reliably potentiated by threatening stimuli, thus capturing the motivational component of negative emotions, that is, activation of the defensive system.7,8

Previous evidence concerning startle responses to experimental pain in healthy individuals has been controversial, suggesting that certain characteristics of the stimulation paradigm might determine whether defensive responses are triggered by pain. Two studies reporting negative findings used constant stimulation over several minutes,9,10 whereas startle potentiation by pain was shown in another study which administered brief painful and non-painful stimuli in a random order.11 These observations indicate that highly predictable pain stimulation might be too low in threat to elicit defensive responding, which is in line with previous research linking unpredictable pain stimulation to high threat (HT).12–15

Based on these considerations, we recently developed a new paradigm encompassing two tonic pain conditions differing in predictability and thus pain-associated threat.16 In the HT condition, participants are led to expect a further increase in stimulation intensity which then occurs in half of the trials; in the low threat (LT) condition, constant stimulation is applied and participants are informed that no further increase will occur. Our previous study with healthy controls yielded an interaction between condition and sequence, with startle potentiation in the HT compared to the LT condition only in the subgroup that completed HT first. In the current study, this paradigm was used to investigate differences between patients with chronic pain and pain-free controls in defensive reactions (ie, startle) due to pain-related threat. Chronic musculoskeletal pain was chosen as the clinical model because of the proven relevance of psychological factors for the development and maintenance of this chronic pain condition.17,18

All the participants underwent both the conditions (LT and HT) in a counter-balanced order; startle responses were elicited by noise bursts and recorded by surface electromyogram (EMG) in the phase of anticipation, that is, always prior to the time points when increases were scheduled in the HT condition and at matched time points in trials without scheduled increases. We also included a standard motivational priming task19 with the presentation of affective pictures in order to check for general differences in startle responding between patients and controls. In addition, we administered a set of questionnaires assessing different emotional and cognitive processing styles relating to pain (eg, pain anxiety and catastrophizing) in order to verify whether our patient sample actually scored higher on these questionnaire measures and to investigate the interrelations between self-report and startle responses during pain stimulation. We assumed that patients with pain would show exaggerated startle responses during pain compared to controls specifically in the HT condition, presumably as a consequence of high pain anxiety or catastrophizing, which is crucially implicated in the development and maintenance of chronic pain.20–22

Materials and methods

Subjects

Patients with pain

Patients with chronic pain attending a multimodal 4-week pain management program at the outpatient unit for pain therapy of the Sozialstiftung Bamberg (Bamberg, Germany) were recruited. This program is run by an interdisciplinary team of medical doctors specialized in anaesthesiology and pain therapy, psychological therapists, physical therapists, occupational therapists, art therapists, and medical assistants and is composed of different therapeutic modules like individual and group therapy with medical or psychological focus, physical therapy, psychological pain coping training, body awareness training, relaxation, creative therapeutic sessions, and outdoor activities (eg, Nordic walking).

Between January 2012 and February 2013, 94 patients attended the program, of whom 41 patients met our inclusion criteria and were approached. Twenty-three patients agreed to participate, representing a recruitment rate of 56%. Four patients dropped out during pain testing because they could not tolerate the pain stimulation (N = 3) or the startle noise (N = 1), resulting in a final sample of 19 patients. Testing took place at a mean of 9.7 days (SD = 5.7) after the start of the pain management program. The age of the chronic pain patients ranged between 23 and 57 years (mean age = 44.0 years, SD = 9.9, 11 women and 8 men).

The inclusion criteria are patients were required to present with a primary diagnosis of musculoskeletal pain (neck pain, upper back pain, and low back pain) or fibromyalgia with pain lasting for a minimum of 6 months before participation. Headaches (migraine, tension-type headache, or non-specified headache) were allowed as secondary diagnosis. The exclusion criteria were other predominant pain diagnosis (eg, patients suffering primarily or exclusively from headache), surgical interventions within the last year and severe mental disorders (mild or moderate affective disorders were allowed due to their high comorbidity with chronic pain). The distribution of pain diagnoses in our final sample is shown
in Table 1. Table 2 provides an overview of patients’ medication. Patients were not asked to pause analgesic medication on the day of testing in order to assure their compliance and minimize interference with the ongoing pain therapy.

Controls

Nineteen healthy, pain-free controls (female: N = 11; age: 25–56 years, M = 44.6, SD = 8.7) were matched with the patients based on sex and age by selecting them from a larger sample of healthy participants. For every patient, we selected an “experimental twin” with the same sex and a minimal difference in age (0–3 years) from our pool of healthy controls; independent t-tests confirmed that the two groups did not differ regarding age. Seventeen of the 19 controls were also part of the healthy sample analyzed in our previous publication.15 Healthy participants were recruited by advertisement at the University of Bamberg and in the local newspaper. None of them had taken any central nervous medication or consumed alcohol at least 24 h prior to the test session according to self-report. Exclusion criteria (assessed by a telephone interview) included all acute or chronic diseases.

Menstrual cycle

Overall, six of the 22 female participants took oral contraceptives; of the remaining 16, three were in the first, four in the second, and three in the third phase of their menstrual cycle, five were post-menopausal, and information was missing for one participant. There were no differences between patients and controls regarding contraceptive intake or distribution of menstrual phases at the time point of assessment.

Ethics

All subjects provided written informed consent and received monetary compensation for their participation. The experimental procedure was approved by the ethics committee of the University of Bamberg.

Materials and procedures

All the experimental sessions were conducted in the afternoon. The duration of the whole session was ~1 h. Participants sat upright in a comfortable chair and were carefully familiarized with all the methods to be used before the start of the experiment.

The testing procedure included: 1) filling out a set of questionnaires assessing different emotional and cognitive processing styles relating to pain; 2) the assessment of startle responses to affective pictures; 3) the psychophysical determination of stimulation intensity; and 4) the assessment of startle responses in two tonic heat stimulation conditions (LT and HT). The sequence of the affective startle task (2) and the pain block (3 and 4) was balanced across subjects.

Table 1 Predominant pain sites in patients with chronic pain (n = 19)

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Total</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic back pain, thereof</td>
<td>10</td>
<td>53</td>
</tr>
<tr>
<td>Neck pain</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Upper back pain</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Low back pain</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Neck pain and low back pain</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Chronic back pain and fibromyalgia, thereof</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Low back pain and fibromyalgia</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Chronic back pain and headache, thereof</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>Neck pain and migraine</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Neck pain, low back pain, and headache (not specified)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Neck pain, low back pain, TTH, and migraine</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Chronic back pain, fibromyalgia, and headache, thereof</td>
<td>2</td>
<td>10.5</td>
</tr>
<tr>
<td>Neck pain, fibromyalgia, and TTH</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Neck pain, low back pain, fibromyalgia, and migraine</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Fibromyalgia</td>
<td>2</td>
<td>10.5</td>
</tr>
<tr>
<td>Fibromyalgia and headache, thereof</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Fibromyalgia, TTH, and migraine</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Note: Diagnoses made by medical doctor trained in anesthesiology and pain therapy.

Abbreviation: TTH, tension-type headache.

Table 2 Consumption of analgesics and antidepressants in the chronic pain patients (n = 19)

<table>
<thead>
<tr>
<th>Medication</th>
<th>Total</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analgesics</td>
<td>17</td>
<td>89</td>
</tr>
<tr>
<td>On demand, thereof</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Nonsteroidal anti-inflammatories</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Nonopioid analgesics</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Nonsteroidal anti-inflammatories and nonopioid analgesics</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Nonsteroidal anti-inflammatories and opioid analgesics</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>On demand in combination with antidepressants, thereof</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Nonsteroidal anti-inflammatories</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>As prescribed, thereof</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Nonsteroidal anti-inflammatories</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Nonsteroidal anti-inflammatories and opioid analgesics</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Nonsteroidal anti-inflammatories, nonopioid analgesics</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>As prescribed in combination with antidepressants, thereof</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Nonsteroidal anti-inflammatories</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Nonopioid analgesics</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Nonsteroidal anti-inflammatories and nonopioid analgesics</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Nonopioid analgesics and opioid analgesics</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Antidepressants</td>
<td>2</td>
<td>11</td>
</tr>
</tbody>
</table>

Note: Analgesics consumed as prescribed and were taken at least once a day.
Pain was induced by the use of a Peltier-based, computerized thermal stimulator with a 3×3 cm² contact probe (Medoc TSA-2001; Medoc Ltd, Ramat Yishai, Israel). The contact probe was attached to the ventral site of the left forearm; the exact position on the forearm was changed after predetermination of stimulation intensity and again between the two threat stimulation conditions in order to avoid local sensitization.

**Questionnaire measures**

As we aimed at a broad coverage of different cognitive and emotional processing styles relating to pain, we used a set of four questionnaires that are related but neither theoretically nor empirically redundant. This set consisted of German versions of the Pain Catastrophizing Scale (PCS), the Pain Anxiety Symptoms Scale (PASS), the Pain Vigilance and Awareness Questionnaire (PV AQ), and the Anxiety Sensitivity Index (ASI). The PCS and PV AQ were translated into German by our workgroup, using a standard “forward–backward” procedure. Translation accuracy was evaluated by an English native speaker based on similarity of the resulting backward English version to the original version. Our German version of the PV AQ has been validated for use in clinical and nonclinical samples.

The PCS was developed as a measure of catastrophizing related to pain. It comprises 13 items that can be divided into three subscales, namely rumination, magnification, and helplessness. The items (eg, “I worry all the time about whether the pain will end”) are rated on a 5-point scale. The PASS is composed of four subscales – cognitive anxiety, escape/avoidance, fearful appraisal, and physiologic anxiety – and is designed to measure fear of pain across cognitive, behavioral, and physiologic domains. The items (eg, “When I feel pain I am afraid that something terrible will happen”) are rated on a 6-point scale. The PV AQ was developed as a comprehensive measure of attention to pain. It consists of 16 items (eg, “I am quick to notice changes in pain intensity”) that are rated on a 6-point scale and that assess awareness, vigilance, preoccupation, and observation of pain. The ASI is a 16-item self-report questionnaire designed to measure anxiety sensitivity, that is, the fear of anxiety-related sensations due to the belief that they might have harmful consequences. It is composed of three subscales: fear of somatic sensations, fear of cognitive dyscontrol, and fear of socially observable anxiety reactions. The items (eg, “It scares me when my heart beats rapidly”) are rated on a 5-point scale. In contrast to pain catastrophizing, pain anxiety, and pain vigilance, anxiety sensitivity is not a pain-specific concept but has been proven to be strongly associated with pain-related fear.

All the questionnaires demonstrated good internal consistency (PCS: Cronbach’s $\alpha = 0.94$; PASS: Cronbach’s $\alpha = 0.92$; PV AQ: Cronbach’s $\alpha = 0.77$; ASI: Cronbach’s $\alpha = 0.93$); values of Cronbach’s $\alpha$ were very similar to those reported for the original English versions. In addition, inter-correlations of the four questionnaires ranged between 0.456 and 0.857 in the current study, which is in accordance with inter-correlations reported for English versions.

**Assessment of startle responses to affective pictures**

In order to check for general differences in startle responding between patients and controls, we administered a standard motivational priming paradigm in which startle responses are recorded during the passive viewing of affective pictures. Commonly, compared to neutral pictures that serve as reference category, startle is potentiated while viewing negative pictures and inhibited while viewing positive pictures. Affective pictures were selected from the IAPS. Based on previous research studies, we chose attack pictures as negative category and erotic pictures as positive category. As we were interested in startle responses to pain stimulation, we added pain-related pictures as second category with negative valence. Neutral pictures were used as reference category. For each of the four valence categories, we chose six representative pictures, resulting in a total of 24 pictures (the IAPS identification numbers were as follows: erotic pictures: 4652, 4659, 4660, 4670, 4687, 4695; attack pictures: 1120, 1300, 1525, 6250.1, 6300, 6510; pain-related pictures: 3010, 3180, 3261, 3350, 9253, 9410; neutral pictures: 2200, 5120, 5534, 7002, 7031, 7150). The sequence of pictures was randomized once and then set for all participants. Each picture was presented for 6 s; a startle noise (white noise bursts of 105 dB and 50 ms duration) was applied during each picture 3, 4, or 5 s after picture onset. After each picture, subjects rated picture valence and arousal as well as the perceived mean intensity of the startle noise (ratings are not reported here). This rating period lasted for 10 s, resulting in a total duration of 6.4 min for the whole task. Data from one patient were missing for this task due to problems with EMG recording.

**Psychophysical determination of stimulus intensity**

Prior to the pain stimulation block, the individual stimulus intensity had to be determined in order to ensure comparable levels of pain sensation in all the participants. For this purpose, we used a two-step procedure (threshold determination and rating procedure) with a total duration of about 15 min.

In the first step, the individual heat pain threshold (PT) was assessed. We decided to use the method of adjustment,
which is particularly reliable due to its independence of reaction time.\textsuperscript{36} Participants were asked to obtain a stimulus level that they perceived as barely painful by adjusting the thermode temperature using heating and cooling buttons, starting from a baseline temperature of 35°C. A constant press of the buttons resulted in heating and cooling rates of 0.5°C/s, respectively. Totally, there were six trials, of which the first was used as familiarization trial and the remaining five were averaged to constitute the threshold estimate.

In the second step, we applied a psychophysical rating procedure, which has been established as standard in our laboratory.\textsuperscript{37–39} The aim of this procedure was to determine two stimulation temperatures that corresponded to a visual analogue scale (VAS) rating of 60 and 70, respectively. Participants were asked to rate eight heat stimuli (5-s duration) with plateaus at temperatures of ±0°C, +1°C, +2°C, +3°C relative to the individual PT. Stimuli were applied in a predetermined pseudo-random order (two stimuli per temperature; rate of change to and from plateau: 2°C/s; inter-stimulus interval [ISI]: 8–10 s). Intensity ratings were provided on a horizontal VAS of 100 mm after each stimulus. Participants were informed about the verbal anchors of the VAS (left end: “no pain”; right end: “extreme pain”) by the experimenter. If the most intense stimulus (PT +3°C) was rated below 50, the series was repeated with an increase by 1°C for all temperature levels.

After completion of this rating procedure, a power function was fitted to the psychophysical data (Table Curve 2D; Scientific Solutions; Lausanne, Switzerland), which provided corresponding temperature intensities for VAS ratings of 60 and 70.

**Assessment of startle responses to pain**

After stimulus intensity had been determined, we applied two pain stimulation conditions. These two conditions were very similar regarding characteristics of the tonic heat stimulus but differed specifically regarding the experience of pain-associated threat. Stimuli in the LT condition were of constant intensity and were also announced as such. In the HT condition, participants were led to expect a further increase in stimulus intensity, which was in fact applied in 50% of the trials. The time interval for presentation of startle probes was identical in both the conditions (18–28 s after stimulus onset); this interval was chosen in order to assure that the startle probe was always presented prior to the further increase in the HT condition. Both the conditions were preceded by a corresponding practice trial, so that participants were informed about the nature of the impending condition. The sequence of the two threat conditions was balanced across subjects and was distributed as follows: 10 patients and 8 controls were assigned to the sequence “low threat – high threat,” 9 patients and 11 controls were assigned to the sequence “high threat – low threat.”

The design of both the conditions is described in detail in the following paragraphs and is also illustrated in Figure 1.

**LT condition**

The LT condition consisted of 12 repetitions of the same stimulus type (stimulus A; Figure 1). Temperature increased from the 35°C baseline to 40°C (rate of rise: 5°C/s) and remained constant for 5 s; after that, it increased to the temperature intensity corresponding to a VAS rating of 60 (rate of rise: 2°C/s) and remained there for 40 s. A startle probe was presented between 18 and 28 s after stimulus onset. After 40 s, temperature decreased to baseline and the subjects provided ratings regarding the startle noise and the temperature (see “Ratings” section for detailed information on ratings). The next stimulus followed after an ISI of 5 s. During each stimulus, the message “Temperature will not increase further” was displayed in blue letters on the computer screen in front of the participant. This message appeared 13 s after onset of the painful temperature, at least 5 s before the startle probe, and remained on the screen until stimulus offset.

**HT condition**

The HT condition consisted of two stimulus types (stimulus A and stimulus B; Figure 1). Stimulus A and B were presented in a predetermined random order (ABBABABBAABAA) with both the stimulus types being repeated six times.

Stimulus A was identical to the 12 stimuli applied in the LT condition (see “LT condition” section). Stimulus B differed from stimulus A only in the respect that the temperature remained on VAS 60 only for 35 s (instead of 40 s); after that, it increased to VAS 70 for 5 s. Again, a startle probe was presented in the interval between 18 and 28 s after stimulus onset (prior to the second increase). During each stimulus (A and B), the message “Temperature may increase further” was displayed in red letters on the computer screen in the same time window as in the LT condition.

Maximum values for tonic stimulation were set to 48.5°C for VAS 60 and to 49.5°C for VAS 70 in order to avoid potential tissue damage. Temperature intensity had to be restricted to these maximum values for four control participants and one patient with pain. However, a simple \textit{t}-test yielded no significant difference in temperature ratings between these five participants and the remaining 33 (valence ratings: T\textsuperscript{[37]} = 1.739, \textit{p} = 0.090; arousal ratings: T\textsuperscript{[37]} = 0.204, \textit{p} = 0.839), indicating that the subgroup stimulated with
restricted intensities did not perceive the stimulation as less intense and aversive.

**Startle probe presentation**

Startle probes were brief acoustic stimuli of 50 ms duration and 105 dB intensity (white noise bursts). The startle probes and a masking background noise (constant white noise, 68 dB) were applied binaurally over headphones. One startle probe was presented during each pain stimulus, resulting in 12 probes per condition. The familiarization trial applied prior to each condition also served as startle habituation trial. The interval of 18–28 s after stimulus onset was used for startle probe presentation in order to ensure that participants had read the message on the screen and were either expecting a second temperature increase (HT condition) or no further increase (LT condition).

**Ratings**

After each pain stimulus, participants were required to provide three ratings regarding the startle noise (1 rating) and the temperature (2 ratings) within the next 22 s. The intensity of the startle noise was rated on a computerized numerical rating scale ranging from 0 (labeled as “no noise”) to 100 (labeled as “extremely loud noise”), which was displayed as a scrollbar. Participants operated the scrollbar via a mouse slide; the starting position of the cursor was determined randomly in each trial. The temperature was rated regarding two affective dimensions, valence and arousal, using a computerized version of the Self-Assessment Manikin (SAM), which is a gold standard for the ratings of emotional responses. We decided on assessing emotional responses instead of pain intensity as the stimulation temperature was already tailored to predefined intensities (VAS 60 and 70, respectively). Participants were asked to provide ratings of valence and arousal by choosing one of five manikins or a box in between via mouse click. Ratings ranged from 1 to 9, with higher ratings corresponding to positive valence and high arousal. Noise ratings of one control participant were invalid due to problems with handling the scrollbar.

**Electromyographic recording and analysis**

Startle blinks were quantified by recording surface EMG activity on the musculus orbicularis oculi beneath the right eye. For that
purpose, Ag/AgCl electrodes filled with electrode paste were used. Prior to application of the electrodes, skin was cleaned with an alcoholic skin detergent to reduce electrode resistance.

EMG raw signals were recorded using the device Sigma PLpro/Type Databox DB 36 including a 16 bit AD convertor with a dynamic range from 0.5 μV to 2 mV. The recording bandwidth of the EMG signal was between 0.2 and 300 Hz; input resistance was >20 mOhm. The signal was sampled at 512 Hz. After recording, the raw signal was analyzed offline using the program “Vision Analyzer” (Brain Products; Munich, Germany). Our algorithm for the analysis of startle responses was based on the recommendations by Blumenthal et al and has been repeatedly applied in previous studies. The raw signal was smoothed, using a 50 Hz notch filter, 20 Hz high-pass filter, and 256 Hz low-pass filter. After that, the signal was rectified and then integrated over a time interval from 0 to 250 ms after startle probe onset. Startle responses were considered as invalid and excluded from analysis if considerable fluctuations in the baseline EMG activity were detected and/or if the peak of activity did not occur in the pre-defined time window (30–100 ms) after the onset of stimulus.

Further statistical analyses were based on two variables, namely peak latency and amplitude of blink responses. Peak latency was defined as time from the onset of startle probe to the maximum value of voltage. Amplitude was defined as voltage difference between the averaged baseline and voltage peak within a time frame of 30–100 ms after the onset of startle probe. For the pain block, peak latency and amplitude were calculated separately for each response, resulting in a total of 24 values (12 values per condition) per subject. Missing values – resulting from the exclusion procedure described earlier – were substituted by means, which were calculated from the data available before and after the missing data point. For the emotion block, mean values of latency and amplitude were calculated for each of the four valence categories based on the available segments, resulting in four values per subject.

Statistical analysis

Descriptive measures

Differences between 1) patients and controls and 2) the two sequence groups (LT–HT and HT–LT) regarding PT, stimulation temperature, and questionnaire measures were explored using independent sample \( t \)-tests.

Startle responses to affective pictures

To evaluate the effects of group and picture valence on startle peak latency and amplitude, we conducted two repeated measurement analyses of variance (ANOVAs) with “picture category” (attack, pain-related, erotic, and neutral) as within-subject factor and “group” (patients vs controls) as between-subject factor.

Adjusting degrees of freedom with Greenhouse-Geisser correction was necessary if sphericity could not be observed. For \( F \)-tests, partial \( \eta^2 \) is reported as an estimate of effect size. For detailed analyses, \( t \)-tests were computed; Cohen’s \( d \) is reported to describe the effect size for paired comparisons.

Startle responses to pain

To evaluate the effects of group and the experimental threat conditions on startle peak latency and amplitude, we conducted two repeated measurement ANOVAs with “condition” (LT vs HT) and “time course” (startles 1–12) as within-subject factors and “group” (patients vs controls) as between-subject factor.

Adjusting degrees of freedom with Greenhouse-Geisser correction was necessary if sphericity could not be observed. For \( F \)-tests, partial \( \eta^2 \) is reported as an estimate of effect size. For detailed analyses, \( t \)-tests were computed; Cohen’s \( d \) is reported to describe the effect size for paired comparisons.

Subjective ratings

To evaluate the effects of group and the experimental threat conditions on valence and arousal ratings of the temperature as well as on startle noise ratings, we conducted three repeated measurement ANOVAs with “condition” (LT vs HT) as within-subject factor and “group” (patients vs controls) as between-subject factor. Adjusting degrees of freedom with Greenhouse-Geisser correction was necessary if sphericity could not be observed. For \( F \)-tests, partial \( \eta^2 \) is reported as an estimate of effect size. For detailed analyses, \( t \)-tests were computed; Cohen’s \( d \) is reported to describe the effect size for paired comparisons.

Correlations between startle and questionnaire measures

To explore the relationship between pain-related psychological variables on the one hand and startle reactions on the other hand, we computed correlations between questionnaire scores and startle amplitude in both the conditions (HT and LT) as well as the difference score (startle amplitude\(_{HT} \) minus startle amplitude\(_{LT} \)).

SPSS 21 (IBM Corporation; Armonk, NY, USA) was used for all calculations; significant effects were assumed at \( \alpha = 0.05 \).

Results

Participant characteristics

Descriptive statistics of scores of each questionnaire are presented in Table 3. Independent sample \( t \)-tests showed that
patients with pain scored significantly higher compared to pain-free controls not only on pain-specific questionnaires but also on anxiety sensitivity (PCS: T \( [36] = 5.039, p < 0.001, d = 1.68 \); PVAQ: T \( [36] = 4.643, p < 0.001, d = 1.55 \); PASS: T \( [36] = 4.705, p < 0.001, d = 1.57 \); ASI: T \( [36] = 3.596, p < 0.001, d = 1.20 \)). In contrast, there were no significant differences between the two sequence groups as regards the temporal position of threat (LT–HT, HT–LT) (PCS: T \( [36] = 7.60 \mu V, SD = 10.68 \mu V, SD = 0.006, d = 0.233, d = 0.05 \). However, there was no significant main effect of group (F\( [1,35] = 0.211; p = 0.649, \eta^2 = 0.006 \)) and no significant interaction (F\( [1,2,41.7] = 0.045; p = 0.987, \eta^2 = 0.001 \)). Thus, we were able to replicate the established motivational priming effect and found no evidence for altered startle response to affective pictures in the patient group.

### PT and stimulation intensities

There were no significant differences in PT between patients with chronic pain and pain-free controls (patients: M = 43.4°C, SD = 2.4°C; controls: M = 44.4°C, SD = 2.65; T \( [36] = 1.173, p = 0.248, d = 0.39 \)). Likewise, there were also no significant differences in the two stimulation intensities, which were determined in the rating procedure. For intensity 1 (corresponding to a VAS rating of 60), mean values were 45.4°C (SD = 2.3°C) for patients and 46.3°C (SD = 2.2°C) for controls (T \( [36] = 1.261; p = 0.215, d = 0.42 \)). For intensity 2 (corresponding to a VAS rating of 70), mean values were 46.1°C (SD = 2.3°C) for patients and 47.1°C (SD = 2.1°C) for controls (T \( [36] = 1.424; p = 0.163, d = 0.47 \)). There were also no significant differences between the two sequence groups (PT: T \( [36] = 0.533, p = 0.597, d = 0.19 \); intensity 1: T \( [36] = 0.337, p = 0.738, d = 0.11 \); intensity 2: T \( [36] = 0.276, p = 0.784, d = 0.09 \).

### Effects of group and threat conditions on startle peak latency and amplitude

#### Startle peak latency

Mean values of peak latency lay in the expected range (patients: M = 76.4 ms, SD = 18.4; controls: M = 78.6 ms, SD = 14.7). There was no main effect of group on startle peak latency (F\( [1,35] = 0.292; p = 0.592, \eta^2 = 0.019 \)). Likewise, we also observed no significant effect of picture category (F\( [1,36] = 1.009; p = 0.322, \eta^2 = 0.008 \)) and no significant interaction (F\( [3,105] = 0.310; p = 0.818, \eta^2 = 0.009 \)).

#### Startle amplitude

The ANOVA yielded a significant main effect of picture category (F\( [1,2,41.7] = 8.903; p = 0.003; \eta^2 = 0.203 \)). As expected, startle amplitude was potentiated during attack compared to neutral pictures (attack: M = 15.85 µV, SD = 26.96; neutral: M = 10.68 µV, SD = 17.78; T \( [36] = 2.944; p = 0.006, d = 0.12 \) and inhibited during erotic compared to neutral pictures (erotic: M = 7.60 µV, SD = 13.61; T \( [36] = 2.876; p = 0.007, d = 0.15 \). Surprisingly, there was no significant difference between pain-related and neutral pictures (pain-related: M = 9.85 µV, SD = 16.60; T \( [36] = 1.214; p = 0.233, d = 0.05 \).

However, there was no significant main effect of group (F\( [1,35] = 0.211; p = 0.649, \eta^2 = 0.006 \)) and no significant interaction (F\( [1,2,41.7] = 0.045; p = 0.987, \eta^2 = 0.001 \)). Thus, we were able to replicate the established motivational priming effect and found no evidence for altered startle response to affective pictures in the patient group.

### Effects of group and picture valence on startle peak latency and amplitude

#### Startle peak latency

Mean values of peak latency lay in the expected range (patients: M = 76.4 ms, SD = 18.4; controls: M = 78.6 ms, SD = 14.7). There was no main effect of group on startle peak latency (F\( [1,35] = 0.292; p = 0.592, \eta^2 = 0.019 \)). Likewise, we also observed no significant effect of picture category (F\( [1,36] = 1.009; p = 0.322, \eta^2 = 0.008 \)) and no significant interaction (F\( [3,105] = 0.310; p = 0.818, \eta^2 = 0.009 \)).

#### Startle amplitude

The ANOVA yielded a significant main effect of picture category (F\( [1,2,41.7] = 8.903; p = 0.003; \eta^2 = 0.203 \)). As expected, startle amplitude was potentiated during attack compared to neutral pictures (attack: M = 15.85 µV, SD = 26.96; neutral: M = 10.68 µV, SD = 17.78; T \( [36] = 2.944; p = 0.006, d = 0.12 \) and inhibited during erotic compared to neutral pictures (erotic: M = 7.60 µV, SD = 13.61; T \( [36] = 2.876; p = 0.007, d = 0.15 \). Surprisingly, there was no significant difference between pain-related and neutral pictures (pain-related: M = 9.85 µV, SD = 16.60; T \( [36] = 1.214; p = 0.233, d = 0.05 \).

However, there was no significant main effect of group (F\( [1,35] = 0.211; p = 0.649, \eta^2 = 0.006 \)) and no significant interaction (F\( [1,2,41.7] = 0.045; p = 0.987, \eta^2 = 0.001 \)). Thus, we were able to replicate the established motivational priming effect and found no evidence for altered startle response to affective pictures in the patient group.

### Effects of group and threat conditions on startle peak latency and amplitude

#### Startle peak latency

Mean values of peak latency were found to be in the expected range (patients: M = 75.1 ms, SD = 15.3; controls: M = 78.6 ms, SD = 12.1). There was no main effect of group on startle peak latency (F\( [1,36] = 0.709; p = 0.405; \eta^2 = 0.019 \)). Likewise, no significant effect of threat condition (F\( [1,36] = 1.009; p = 0.322, \eta^2 = 0.027 \)) and of time course (F\( [7,3,263.2] = 1.814; p = 0.081, \eta^2 = 0.048 \)) was observed. There were no significant interaction effects.

#### Startle amplitude

Due to huge intra-individual variations in startle amplitude from trial to trial, we decided to smooth the time series by computing averages: Trials were grouped into four succeeding blocks (trial 1–3, 4–6, 7–9, and 10–12) and then averaged. By using this procedure, we aimed to achieve higher stability than when using single data points and still allowed for checking for possible trends over time (linear, quadratic, or cubic).

Descriptive statistics of startle amplitude for both the groups in each trial block of the LT and HT condition are given in Table 4. We detected no significant main effect of group (F\( [1,36] = 0.327; p = 0.440; \eta^2 = 0.017 \)) and also no interactions with group. Thus, there was no difference in startle amplitude between patients and pain-free controls.
between patients and controls in startle amplitude as reaction to the noxious stimulation and threat manipulation. The lack of group differences in startle modulation observed in both the experimental conditions – picture viewing and pain stimulation – suggests normal emotional responsiveness in our patient sample.

Importantly, the effect of “time course” also failed to reach significance ($F[2,176.8] = 1.621; p = 0.203; \eta^2 = 0.043$), indicating that there was no decline in startle amplitude due to habituation, for example, in the course of each condition. Correspondingly, no significant trend for “time course” was observed (linear, quadratic, or cubic; all $p > 0.100$). However, there was a significant effect of threat condition ($F[1,36] = 4.655; p = 0.038; \eta^2 = 0.115$); as hypothesized, startle amplitude was overall higher in the HT compared to the LT condition (HT: $M = 21.1 \mu V$, SD = 34.7; LT: $M = 14.8 \mu V$, SD = 19.9; $T[37] = 2.180, p = 0.036, d = 0.35$). All other effects failed to reach significance.

Due to the nonsignificant effect of group, we decided on combining both the groups and running a second ANOVA with “condition” (LT vs HT) and “time course” (startle blocks 1–4) as within-subject factors and “sequence of conditions” (LT–HT vs HT–LT) as between-subject factor. Besides the still significant main effect of “condition” ($F[1,36] = 4.715, p = 0.037; \eta^2 = 0.116$), the ANOVA also yielded a significant main effect of “sequence of conditions” ($F[1,36] = 4.983; p = 0.032; \eta^2 = 0.122$). Most importantly, we found a significant interaction between “condition” and “sequence of conditions” ($F[1,36] = 4.715; p = 0.037; \eta^2 = 0.116$); effects of the threat condition on startle amplitudes were detected mainly in participants who completed the HT condition first. In this subgroup, startle responses were significantly higher in the HT compared to the LT condition (HT: $M = 32.8 \mu V$, SD = 44.7; LT: $M = 20.6 \mu V$, SD = 25.4; $T[19] = 2.340, p = 0.030, d = 0.52$), whereas there was no significant difference between the two threat conditions in participants who started with LT (LT: $M = 8.3 \mu V$, SD = 7.5; HT: $M = 8.2 \mu V$, SD = 6.6; $T[17] = 0.071, p = 0.944, d = 0.02$; Figure 2). In addition, when comparing responses to both the conditions between the two sequence groups (LT–HT and HT–LT), participants starting with HT showed higher responses to both the conditions than those starting with LT (group comparison for HT: $T[36] = 2.306, p = 0.025, d = 0.75$; group comparison for LT: $T[36] = 1.983, p = 0.050, d = 0.64$).

Again, there was no main effect of or interaction with “time course” (all $p > 0.05$).

Effects of group and threat conditions on temperature and noise ratings

As we obtained no difference between patients and controls in temperature and noise ratings (all $p > 0.50$), we again combined both the groups and conducted three ANOVAs with “condition” (LT vs HT) as within-subject factor and “sequence of conditions” (LT–HT vs HT–LT) as between-subject factor. Results of these ANOVAs are shown in Table 5 and summarized in the following paragraphs.

Temperature valence ratings

The ANOVA conducted on temperature valence ratings yielded no significant main effect of “condition” or “sequence of conditions” (Table 5). However, there was a significant interaction: participants who started with the LT condition rated the HT condition as significantly more aversive than the LT condition (LT: $M = 3.66$, SD = 1.57; HT: $M = 3.04$, SD = 1.6; $T[17] = 2.66, p = 0.016, d = 0.39$). In contrast, there was no difference between conditions in participants who started with the HT condition (HT: $M = 3.99$, SD = 1.47; LT: $M = 3.95$, SD = 1.89; $T[19] = 0.228, p = 0.822, d = 0.03$).

Temperature arousal ratings

The ANOVA conducted on temperature arousal ratings yielded no significant main effect of “condition” or “sequence of conditions” and also no significant interaction (Table 5). Overall, the thermal stimulation was rated as moderately arousing ($M = 4.2$, SD = 2.1).

Startle noise ratings

For startle noise ratings also, we detected no significant main effect of “condition” or “sequence of conditions” and also no significant interaction (Table 5). Overall, ratings for the noises burst were – as expected – in the medium range ($M = 47.5$, SD = 17.3).

Table 4 Descriptive statistics of startle amplitude (µV) for patients and controls for each trial block (1–4) in both the threat conditions (high threat and low threat)

<table>
<thead>
<tr>
<th>Trial block</th>
<th>M (SD) Patients with pain (N = 19)</th>
<th>M (SD) Pain-free controls (N = 19)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High threat condition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block 1 (Trials 1–3)</td>
<td>19.33 (21.10)</td>
<td>28.71 (48.83)</td>
</tr>
<tr>
<td>Block 2 (Trials 4–6)</td>
<td>18.37 (32.66)</td>
<td>25.56 (45.33)</td>
</tr>
<tr>
<td>Block 3 (Trials 7–9)</td>
<td>18.08 (26.45)</td>
<td>25.10 (46.00)</td>
</tr>
<tr>
<td>Block 4 (Trials 10–12)</td>
<td>12.25 (13.91)</td>
<td>21.63 (36.25)</td>
</tr>
<tr>
<td>Low threat condition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block 1 (Trials 1–3)</td>
<td>11.47 (14.23)</td>
<td>19.00 (23.06)</td>
</tr>
<tr>
<td>Block 2 (Trials 4–6)</td>
<td>12.30 (14.23)</td>
<td>16.85 (20.68)</td>
</tr>
<tr>
<td>Block 3 (Trials 7–9)</td>
<td>11.44 (17.21)</td>
<td>15.61 (24.56)</td>
</tr>
<tr>
<td>Block 4 (Trials 10–12)</td>
<td>12.98 (20.30)</td>
<td>18.35 (32.57)</td>
</tr>
</tbody>
</table>
Correlations between questionnaire scores and startle responses

In chronic pain patients, we found a significant positive correlation between PASS score and startle amplitude in the LT condition \( (r = 0.458; p = 0.049) \); in the HT condition, this correlation was still positive but far from significance \( (r = 0.243; p = 0.316) \). Thus, patients reporting higher levels of pain-related anxiety showed stronger startle reactions to pain in case of low situational threat.

Conversely, in pain-free controls, there was a significant negative correlation between PVAQ score and startle amplitude in both the conditions (HT condition: \( r = -0.508, p = 0.026 \); LT condition: \( r = -0.528, p = 0.020 \)), suggesting that pain-free subjects who reported higher vigilance to pain showed smaller startle reactions irrespective of threat level.
All the other correlations failed to pass the level of significance (all $p > 0.05$). Thus, only three of 24 correlations (12 for each group) reached significance.

**Discussion**

Our study aimed at investigating startle responses to tonic pain with varying threat levels in patients with chronic pain compared to matched healthy controls. Pain-associated threat was manipulated by applying either constant, predictable stimulation (LT condition) or constant stimulation including announced but unpredictably timed further increases in temperature (HT condition). We hypothesized that pain patients would show higher reactions than controls particularly in the HT condition. This specific assumption was derived from previous research showing effects of trait catastrophizing on performance in a painful task being most pronounced under high situational threat,\(^4\) which were interpreted within the framework of theories postulating the activation of stable cognitive schemas relating to threat by threat-associated situational cues.\(^43\)–\(^46\)

Contrary to our hypotheses, startle reactions during noxious stimulation as well as valence and arousal ratings were comparable in patients and controls despite significantly higher pain-related anxiety, pain catastrophizing, pain vigilance, and anxiety sensitivity in the patient group. Thus, patients showed neither stronger activation of the motivational defense system (startle responses) nor stronger emotional responding on a subjective level (ratings).

This finding is somewhat surprising given the different experiences with pain in these two groups. According to a catastrophizing hypothesis of chronic pain,\(^20\) one would expect stronger motivational and emotional responding to pain stimuli in patients with chronic pain due to heightened levels of pain-related anxiety and pain catastrophizing, which have been frequently reported.\(^5\)\(^,\)\(^6\) In line with this, fMRI studies have shown enhanced activity in affective brain regions during experimental pain stimulation\(^4\)\(^,\)\(^6\) and disturbed resting-state connectivity in frontolimbic networks\(^5\)\(^,\)\(^47\)\(^,\)\(^48\) in chronic back pain patients. Furthermore, chronification of back pain seems to be accompanied by a shift in brain activity from nociceptive to emotional circuits.\(^49\) Some studies suggest that patients with pain also evaluate experimental pain as more unpleasant than controls;\(^4\)\(^,\)\(^50\) however, this view has been challenged by several negative findings.\(^51\)\(^–\)\(^53\)

One possible explanation for the lack of group differences might be that individual differences in pain-related anxiety between pain patients and controls might not become visible in unambiguous experimental situations, relating to the concept of the strong situation.\(^54\) This concept has been proposed to explain the effects of manipulations in anxiety experiments; here, strong situations, that is, situations providing unambiguous threat cues, tend to trigger uniform responding in all participants independent of their individual level of trait anxiety, whereas individual differences become obvious in weak situations which offer some degree of ambiguity.\(^55\) Our experimental paradigm might be regarded to create strong situations as the HT condition was clearly announced by situational cues and participants knew exactly what to expect.

An alternative explanation for the lack of group differences might be that pain patients show exaggerated responding only to more individualized or stronger threat manipulations. For example, it might be necessary to induce concerns that the painful stimuli might actually cause bodily harm; this has been done in several studies by using threatening instructions concerning the stimulation paradigm.\(^42\)\(^,\)\(^56\)\(^,\)\(^57\) Direct comparison of different types of experimental threat manipulations and their effects on healthy participants and patients with chronic pain should be targeted in future studies.

The observation of a significant main effect of threat condition is in line with the observation that startle potentiated during the uncertain expectation of aversive stimulation which has been replicated numerous times in anxiety research.\(^58\)\(^,\)\(^59\) However, replicating the results of our previous study conducted with a healthy pain-free sample,\(^16\) we also detected a significant interaction between threat condition and sequence: The difference between the two conditions, that is, higher startle reactions in the HT compared to the LT condition, was significant only in the subgroup completing the HT condition first. This pattern of results might indicate that the effect of threat is strong (and thus easily detectable) only under conditions of novelty, whereas it is more subtle (and thus harder to detect with small sample sizes) in case of familiarity with the painful stimulus. Alternatively, the subjective appraisal of pain-associated threat as low or high might have been partly transferred from the first to the second condition. Similar to context conditioning, having the first experience with the painful stimulus in a non-threatening context might have led to a less-threatening appraisal of the HT condition afterwards and vice versa despite a change in instructions. Importantly, the observation that this sequence effect was equally pronounced in patients with pain and controls despite higher pain-related catastrophizing and anxiety in the patient group suggests that it is based on experiences with a specific situation rather than on general experiences with pain as more or less threatening.
Limitations
There are some limitations in our study. First, our sample of patients with chronic pain was not perfectly homogeneous with ~40% of the patients presenting with two or three pain diagnoses. However, we deemed it as most important that 1) patients had a primary diagnosis of neck or back pain or fibromyalgia and 2) our sample was representative for patients with musculoskeletal pain admitted to an outpatient unit. Second, most of the patients took analgesic and/or antidepressant medication which might have influenced responses to our experimental manipulation. Third, we did not include direct ratings of subjective threat in order to avoid response bias by making the subjects aware of the key relevance of threat experiences; instead, threat was indirectly assessed using the arousal subscale of the SAM which is a standard instrument for the evaluation of emotional stimuli. However, future studies might benefit from including a manipulation check of threat at the end of the experiment and/or the assessment of state anxiety in the course of the experimental session. Finally, due to the pilot-like character of our study and the limited sample size, we decided on being less conservative as regards control of type 1 error and using a relatively simple analytical strategy (ie, inclusion of only one between factor at a time) as well as skipping Bonferroni correction in the correlational analyses; future studies investigating larger samples should be more conservative in this respect.

Conclusion
The similarity between patients and controls of the motivational priming effect observed in the viewing task for emotional pictures indicates that, in general, emotional responsiveness is not necessarily altered in patients with chronic pain. More importantly, we also observed no differences between patients and controls in startle responding to pain in two conditions with varying threat levels despite significantly higher pain-related catastrophizing and anxiety in the patient group. Altogether, emotional regulation seemed to be preserved in patients with chronic musculoskeletal pain. Differences between pain patients and healthy individuals might become obvious only in more ambiguous situations or in situations where specific health concerns of the patients are provoked by the threat manipulation. In addition, the fact that sequence effects were similarly pronounced in patients and controls indicates that just previous experiences with specific pain stimuli might be more decisive for the subjective appraisal of pain-associated threat than the remote pain history.

Acknowledgment
This study was supported by a research grant from the Deutsche Forschungsgemeinschaft (La 685/6-3).

Disclosure
The authors report no conflicts of interest in this work.

References