Journal of Pain 37 (2025) 105557

Contents lists available at ScienceDirect

The Journal of Pain

journal homepage: www.jpain.org

Original reports

Assessing the reliability and association of pain ratings and skin conductance responses: Insights from habituation and sensitization to pain☆

Maite M. van der Miesen [a,*], Catherine J. Vossen [a,b], Judith Eck [c,d], Sophie Kühne [e], Elbert A.J. Joosten [a,b], David E.J. Linden [f], Judith C. Peters [d]

[a] Department of Anesthesiology and Pain Management, Mental Health and Neuroscience Research Institute, Maastricht University, Maastricht, the Netherlands
[b] Department of Anesthesiology and Pain Medicine, Maastricht University Medical Centre, Maastricht, the Netherlands
[c] Brain Innovation B.V., Maastricht, the Netherlands
[d] Department of Cognitive Neuroscience, Maastricht University, Maastricht, the Netherlands
[e] Department of Psychiatry, Psychotherapy and Psychosomatics, Faculty of Medicine, RWTH Aachen, Aachen, Germany
[f] Department of Psychiatry and Neuropsychology, Mental Health and Neuroscience Research Institute, Maastricht University, Maastricht, the Netherlands

ARTICLE INFO

Keywords:
Habituation
Sensitization
Pain
Reliability
Skin conductance

ABSTRACT

Repeated painful stimulation results in substantial inter-individual differences in habituation and sensitization. The extent to which these responses reflect state versus trait characteristics remains unclear, highlighting the need to assess the reliability of these differences over time. Furthermore, the association between subjective pain ratings and skin conductance responses (SCR) has not been examined in this context. This preregistered study investigated profiles of habituation and sensitization to pain across two sessions using pain ratings and SCR. Participants underwent repeated painful electrical stimulation over two sessions separated by four weeks, receiving 75 stimuli across three runs per session. Pain intensity was rated after each stimulus, and continuous SCRs were recorded. Our results demonstrated moderate test-retest reliability of both pain ratings and SCRs, with within-run measures showing greater consistency than across-run measures. Remarkably, participants displaying sensitization exhibited higher reliability than those with habituation or no-change patterns. High test-retest reliability could suggest a trait-like response with reduced adaptability to repeated stimulation, while a higher variability (and thus low test-retest reliability) indicates state-dependent flexibility and adaptability. Our results suggest that interventions to modulate pain could be targeted at changing such sensitization patterns and promoting habituation. Furthermore, pain ratings showed diverse trajectories of habituation and sensitization, whereas SCRs predominantly habituated. This dissociation between subjective pain perception and autonomic responses challenges the prevailing view that higher pain ratings correspond to elevated SCRs. Together, these results underscore the importance of considering habituation and sensitization dynamics, with subjective and physiological measures providing complementary insights into the multidimensional pain response.

Perspective: Repeated painful stimulation resulted in patterns of habituation and sensitization, with large individual variability. Test-retest reliability was moderate, with higher consistency for individuals who sensitize. A dissociation between ratings and SCR was demonstrated, with diverse response patterns of ratings and mostly habituation of the SCR.

Introduction

Pain is a subjective experience which is often reported using individual self-report pain ratings. Autonomic measures such as electrodermal activity (EDA) can provide a more objective measure of arousal and salience during painful stimulation. The skin conductance response (SCR), which is a short-term change in EDA due to a stimulus, and the skin conductance level (SCL), representing the average trace of

☆ The authors confirm that the data supporting the findings of this manuscript are available on: https://osf.io/2emsq.

* Correspondence to: Maastricht University, Department of Anesthesiology and Pain Management, Universiteitssingel 50, Maastricht 6229 ER, the Netherlands.
E-mail address: m.vandermiesen@maastrichtuniversity.nl (M.M. van der Miesen).

https://doi.org/10.1016/j.jpain.2025.105557
Received 20 June 2025; Received in revised form 19 August 2025; Accepted 8 September 2025
Available online 11 September 2025
1526-5900/© 2025 The Author(s). Published by Elsevier Inc. on behalf of United States Association for the Study of Pain, Inc This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

skin conductance, are most often investigated.

Repeated painful stimulation can lead to habituation (a decrease in pain response), or sensitization, an increased response.[1,2] Habituation and sensitization are non-associative forms of learning and are shown across phylogeny.[3,4] They can be quantified with measures of self-report, EEG and fMRI.[5,6] as well as autonomic measures like SCL[7] and SCR.[8–10] Our previous work has shown large interindividual differences in pain reports of habituation and sensitization.[11] These patterns emerged in response to repeated painful stimulation at the same intensity, both within and across runs of stimulation. However, it is unclear whether these individual patterns are stable over time. A consistent habituation or sensitization profile would suggest a trait-like characteristic, enabling prediction of future responses. In contrast, fluctuating response profiles across sessions would imply a state-dependent mechanism that is affected by situational factors. Test-retest reliability can assess the temporal stability of these patterns for both across-run and within-run measures. One study reported moderate reliability for across-run measures of SCL, while pain ratings within-runs did not demonstrate good test-retest reliability.[12] To systematically evaluate the stability of individual habituation and sensitization profiles, we quantified in this study the test-retest reliability of pain ratings and SCRs within and across runs between two sessions. Several studies have investigated the association between pain ratings and skin conductance. Overall, both measures vary as a function of stimulus intensity, indicating that higher pain ratings correlate with higher EDA.[13–17] However, one study reported that SCRs failed to discriminate between electrical stimulus intensities.[14] Furthermore, studies reported mixed findings on whether EDA is more closely related to stimulus intensity,[18,19] or more to perceived pain.[17,20] In habituation and sensitization research, a constant stimulus intensity is used, eliminating the effect of stimulus intensity in the relationship between pain reports and EDA. Here, we use this approach to evaluate the association between subjective self-reports and objective autonomic measures in the context of habituation and sensitization patterns independent of stimulus intensity.

Individual factors such as gender, age, catastrophizing and coping strategies have been investigated to predict habituation and sensitization.[5,21] So far, the coping strategy "positive self-statements" showed a relationship with habituation,[11] but this needs further replication. In addition, the role of anxiety has not yet been established.

Therefore, in this preregistered study we investigated how pain ratings and SCRs change across sessions of repeated painful stimulation and quantified the test-retest reliability of these measures. Furthermore, we evaluated the association between pain ratings and SCRs in the context of habituation and sensitization. Finally, we examined the influence of anxiety and pain coping strategies on habituation and sensitization.

Methods

This study is preregistered at the Open Science Framework (OSF; https://osf.io/8a7qn). Data and code can be found at https://osf.io/2emsq. Additions or changes to the preregistration are described in the supplemental material. The research protocol was approved by the local ethics committee of the Faculty of Psychology and Neuroscience at Maastricht University and conducted in accordance with the Declaration of Helsinki. There was no patient and public involvement for the aims of this specific study.

Hypotheses

We were interested in habituation and sensitization patterns and their stability, as assessed by behavioral and physiological measures. To this end, we studied 1) general changes across sessions (hypotheses H1a, H1b, H2a, H2b, H2c), 2) test-retest reliability of self-report ratings (H3a, H3b) and skin conductance (H3c, H3d) measures, 3) the relation between questionnaire reports and habituation and sensitization response profiles (H4), and 4) the relation between self-report ratings and skin conductance measures in the context of habituation and sensitization (H5). The specific formulated preregistered hypotheses are shown in Table 1.

Participants

60 healthy participants between 18 and 30 years old were recruited based on our preregistered sample size estimation (see https://osf.io/8a7qn). Participants were recruited through the Maastricht University research participation system and through local advertisements. Exclusion criteria for participation consisted of chronic pain (lasting three months or longer), recurrent pain in the last three months, diabetes, cardiovascular diseases, pregnancy, reported psychiatric (e.g., depression, anxiety) or neurological disorders, current use of psychotropic medication, alcohol or drug abuse or a skin condition affecting the left hand. A survey assessing exclusion criteria was sent out when the participant had signed up for the experiment. The survey was then repeated at the start of the experiment to check that no exclusion criteria were fulfilled. Furthermore, participants were asked to limit alcohol to two units the evening before, refrain from recreational drugs for a

Table 1
Overview of preregistered hypotheses for the study. Note that the hypotheses are the same as in the preregistration but in different order for clarity. ICC = intraclass correlation coefficient.

Aim	Hypothesis	Content
1. General changes across sessions	H1a	A significant increase in pain threshold (i.e., habituation) is expected over the course of two sessions (i.e., 28 days later).
	H1b	Anxiety is expected to be significantly lower for session 2.
	H2a	No significant change in average pain ratings is expected over the course of two sessions.
	H2b	No significant change in skin conductance response is expected over the course of two sessions.
	H2c	An increase in non-response trials for skin conductance (i.e., habituation) is expected within a session but not between sessions.
2. Test-retest reliability	H3a	No significant change in slopes across runs (reflecting habituation/sensitization) of self-report ratings will be expected over the course of the two sessions.
	H3b	The expected test-retest reliability (expressed with the ICC) for reliability of self-reported pain is moderate to good (0.5–0.9).
	H3c	No significant change in slopes across runs (reflecting habituation/sensitization) of skin conductance responses will occur over the course of two sessions.
	H3d	The expected test-retest reliability (expressed with the ICC) for reliability of skin conductance responses is moderate to good (0.5–0.9).
3. Relation with questionnaire reports	H4	Positive self-instructions will be predictive of the slope of the ratings across run, replicating previous work [11]
4. Association between self-report ratings and skin conductance measures	H5	Self-report ratings and skin conductance will be moderately correlated within a session and this correlation will not be significantly different between sessions.

minimum of one week and painkillers for at least 12 h prior to each experimental session. Participants that fulfilled any of the following preregistered data exclusion criteria during the experiment were excluded and replaced: 1) a pain threshold exceeding 5 mA during the first session (based on [22]) or 2) lacking comprehension of the use of the rating scales or 3) failing equipment leading to more than 20% of data loss 4) when there was a substantial amount in time between runs (i.e., more than 15 min). Participants were excluded from data analysis if more than 20% of rating data was missing (for other reasons than failing equipment) or if in skin conductance data, the average stimulus-locked amplitude was lower than 0.02 μS per session (based on recommendations for electrodermal measurements[23]). All participants received an information letter and signed informed consent before the start of the experiment. Participants were paid €22,50 for two sessions or received course credits as part of the Psychology curriculum at Maastricht University.

Experimental procedures

The study consisted of two sessions (S1 and S2) that were on average four weeks apart, in which participants rated 75 painful stimulation trials per session while their skin conductance was measured.

Experimental procedures were identical in both sessions (Fig. 1). Several questionnaires were administered (see section Questionnaires) and participants practiced the use of a numeric rating scale (NRS) to provide pain intensity ratings. The rating scale was presented on a computer screen and with their right hand, participants moved a slider on the scale between 0 and 100 with several anchors ("How much pain did you feel?" 0: No sensation, 20: Moderate sensation, 40: Pain threshold, 60: Moderate pain, 80: Severe pain, 100: Worst pain imaginable). The anchors, which allow for habituation below the pain threshold, have been used in previous studies.[11,24,25] Participants were carefully explained what each of the anchors meant, especially the pain threshold, which was explained as the moment the stimulus starts to feel painful and/or uncomfortable.

Participants received painful electrical stimulation on the left middle finger (in line with our previous research[11,22]). Participants were asked to wash their hands with warm water, without soap. The skin on the middle finger was disinfected with an alcohol pad and the upper layer of the skin (about 1 mm) at the distal and middle phalanges was scrubbed to reduce resistance for electrical stimulation. Small amounts of electrode gel (Spectra 360, Parker, USA) were put on two electrodes (Biopac, EL509) which were attached to the prepared skin and secured using adhesive tape. For SCR measurements, the participant's hand palm was scrubbed using Nuprep gel (Weaver and company, USA) to increase skin conductivity. Two additional Biopac electrodes with skin conductance electrode paste (Biopac, Gel101) were attached to the hand palm.[23,26] The stimulation electrodes were connected via leads (Biopac, LEAD108B) to an emergency stop button which was interfaced with a Digitimer DS5 constant current stimulator (Digitimer Ltd, UK) and National Instruments card (NI, USA). The skin conductance electrodes were linked to a GSR module which was connected to a BrainAmp ExG amplifier with AUX box (BrainVision LLC, USA).

Participants then performed two calibration sets in which their sensation and pain threshold was established. During calibration, electrical stimuli (10 ms) were delivered starting at 0 mA using steps of 0.1 mA until the participant reported the first sensation, and then continued until the pain threshold was reached. Participants indicated their rating using the pain rating NRS scale. The procedure was repeated twice, and values were averaged. Participants then performed the main task (see section Task) and filled in further questionnaires afterwards.

The second session (S2) followed the same protocol and was scheduled on average 28 days after S1, at approximately the same time of the day. The data was collected by three different experimenters following standardized instructions, always ensuring that the same experimenter collected both sessions per participant. The rating scale practice was briefly repeated with some possibilities for exploration, the electrodes were placed at the same positions, and the calibration procedure was repeated. Habituation might be more pronounced with a lower (perceived) stimulus intensity.[1] Therefore, we decided that it is important to have a similar perceived stimulus in the second session, contrary to previous work which investigated test-retest reliability based on the same applied stimulus intensity.[12] The main task was identical to the first session.

Questionnaires

At the start of the first session, participants provided general information about age, gender, medication use, and use of hormonal birth control. Participants completed the Fear of Pain Questionnaire-9 (FoP)[27]

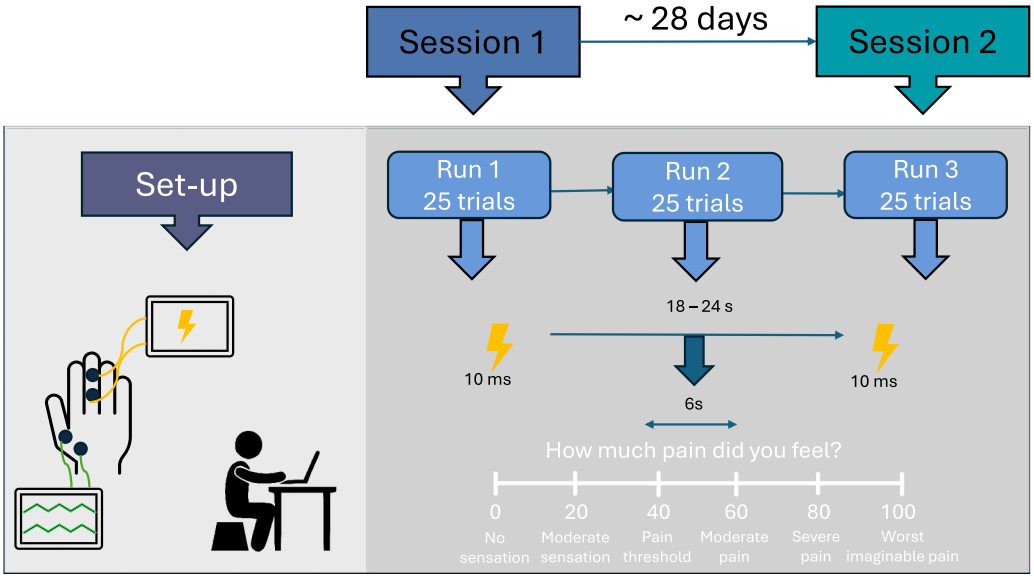

Fig. 1. Overview of the task design. Participants received electrical stimulation on their finger while skin conductance was measured on their palm during two sessions (S1 and S2), which were on average 28 days apart. Each session consisted of three runs with 25 trials each. Time between trials was 18–24 s. After each electrical stimulus, participants rated the intensity of experienced pain using a NRS scale.

and the Coping Strategies Questionnaire (CSQ) to assess strategies for coping with pain (as part of the preregistered hypotheses).[28,29] The CSQ consists of seven subscales, targeting the use of different strategies. To replicate our previous work,[11] we only assessed the relationship between the positive self-statements subscale and habituation. In the beginning of each session, participants rated their current pain using the NRS (0−10) and were asked whether they had experienced any painful, mental or physical events during the past four weeks that could impact their wellbeing. Furthermore, the self-assessment Manikin (SAM)[30] including control, happiness and excitement was assessed on a scale of 1–5, and an NRS on anxiety and energy were administered before and after the experiment (recalling their experience during the experiment) in both sessions. At the end of the second session, participants were asked to describe their change in pain ratings as increasing ("sensitization to pain"), decreasing ("habituation to pain") or "no change".

Task

Participants received three runs of brief electrical stimulations (10 ms). Each run consisted of 25 trials with 18–24 s inter-trial intervals (Fig. 1). Between each run, there was a break of 4 min in which the experimenter talked to the participant. Importantly, all stimuli were of the same intensity, calculated as the pain threshold + 0.50*(pain threshold – sensation threshold). This intensity was chosen to aim for a mild painful stimulus (between 40 and 60 on the scale), leaving room for both habituation and sensitization. After each stimulus, the participant provided their pain rating using the NRS during an interval of 6 s. A fixation cross was displayed on the screen, except during the presentation of the NRS. Participants were neither informed that the stimulation would always be the same, nor that we were interested in habituation/sensitization. The stimuli and rating scale were presented using Presentation software (Neurobehavioral systems Inc, USA).

Skin conductance processing

For skin conductance data, the phasic stimulus-locked component was investigated. Data was filtered using a low-pass filter (5 Hz) and down-sampled to 10 Hz. Event-related activations were extracted using *Ledalab* (Matlab-based software).[31] The phasic driver activity, hereafter referred to as SCR, was extracted using continuous decomposition analysis in a time-window of 1–6 s post-stimulus. According to best practice recommendations, data was transformed using a log transformation to account for the skewed distribution.[23] Non-response trials were defined using a threshold of 0.02 μS.[23]

Statistical analyses

All analyses were performed in RStudio (v 2024.04.1) and R (version 4.3.3). Deviations from the preregistered analyses can be found in the supplemental material. For hypothesis H1a, pain thresholds were compared between sessions using a one-sided paired t-test. In addition, we explored the correlation between the change in pain threshold and change in average pain ratings between sessions using Pearson's r. For H1b, anxiety scores at the start of each session were compared using the non-parametric Wilcoxon signed rank test (as assumptions for the paired t-test were violated). As part of exploratory preregistered analyses, we investigated whether NRS scores on general pain changed across sessions using a Wilcoxon signed rank test. Furthermore, we explored whether scores on the SAM changed from S1 to S2 and from pre to post (we retrospectively assessed their experience during the experiment) using repeated measures ANOVAs.

Changes in pain ratings (H2a) across sessions were assessed using a linear mixed-effects model with trials nested in runs and participants using R package *nlme*.[32] An interaction between trial and run was included to describe effects over time on the group level, and the main effect of session was included to investigate whether pain ratings changed over the sessions. To compare our results with previous studies, we also investigated the main change in pain ratings across sessions using a paired t-test. For H2b, changes in SCRs across sessions (using the phasic stimulus-locked component) were assessed using a similar mixed-effects model analysis as for H2a, with trials nested in runs and participants. For H2c, the number of non-response trials was calculated per participant (defined as a response lower than 0.02 μS) and analyzed using a repeated measures ANOVA. However, due to the non-normality of the data and the limited effect of transformations, we additionally ran a multilevel model with a Poisson distribution to account for the skewed distribution, using R package *lme4*.[33]

For H3 assessing test-retest reliability, individual linear regression slopes were fitted to define habituation and sensitization. Slopes were either fitted across all three runs and compared between sessions (i.e., across-run slope reliability) or fitted within-runs (yielding three slopes per session) and compared between sessions (i.e., within-run slope reliability). A significant positive slope compared to 0 was categorized as 'sensitization', a significant negative slope as 'habituation' and a non-significant slope as 'no change'. Across-run slopes between sessions were compared at the group level (H3a) with a paired t-test. In a separate exploratory analysis, we analyzed the correlation between the change in pain threshold across sessions and the change in across-run slopes with Pearson's r. For H3b, reliability of the across-run slopes between sessions was assessed using the ICC with a two-way random model and type 'agreement', R package *irr*.[34] ICC < 0.4 was classified as 'poor', $0.4 <$ ICC < 0.6 as 'fair', $0.6 <$ ICC < 0.8 as 'good' and ICC > 0.8 as 'excellent'.[35,36] Next, for exploratory purposes to find the best fit to the data and compare with previous research, different measures of variation across sessions were compared. We examined the ICC of the log slope (using individual regression slopes with a log function), the coefficient of variation (CV; a measure of dispersion around the mean), and the repeatability coefficient indicating the closeness of repeated measurements. Additionally, the mean percentage change of ratings between runs, the ICC of the mean ratings, the ICC of the SD, and the group-level correlation between ratings were investigated. These measures were analyzed for both the across-run slope and within-run slopes, always comparing the two sessions. For the group-level correlation, individual ratings of session 1 and 2 were correlated using Spearman's rank correlation (using all data or per run), transformed using Fisher's Z transformation with R package *DescTools*,[37] averaged and transformed back to Spearman's rank correlation. Fisher's Z transformation was employed to account for the skewed sampling distribution of correlations.[38]

For H3c and H3d skin conductance slopes across runs were compared between sessions at the group level using a paired t-test and the ICC was calculated to assess individual reliability, following the procedures described above. Finally, we explored the within-run slopes and ICC of these slopes between sessions.

For H4, the sum score of the positive self-instructions' subscale was correlated with the across-run slope of the pain ratings of S1 (when the questionnaire was administered) using Pearson's r. Furthermore, as part of exploratory preregistered analyses, the total score of the Fear of Pain questionnaire, the anxiety rating and the SAM ratings were correlated to the across-run slopes. P-values were corrected for multiple comparisons using the Bonferroni correction.

For H5, the association between SCRs and ratings was investigated. Spearman rank-order correlation was calculated for each participant per session. The resulting correlation coefficients were transformed using Fisher's Z transformation. Their absolute values were then taken to account for habituation and sensitization effects, and the coefficients were subsequently averaged.[38] The average coefficient per session was tested against zero using Fisher's Z-Test (due to non-normality of the distribution of coefficients this replaced the preregistered one-sample t-test). Then, the average coefficients were compared between S1 and S2 using Zou's approach for dependent non-overlapping correlations,[39] using R package *cocor*.[40]

Three additional (non-preregistered) control analyses were

performed, focusing on the rating categories habituation, sensitization and no change, which were based on the across-run slopes. First, correlation coefficients were compared between each rating category (based on the across-run slope) using a one-way ANOVA. Pairwise comparisons were computed using Tukey's 'Honest Significant Difference' method. Second, skin conductance slopes were compared between each rating category using a one-way ANOVA. Third, the skin conductance slopes were correlated with the individual stimulus intensity using Pearson's r, after log-transforming the data to account for non-normality.

Results

Overall participant characteristics

In total n = 66 participants were recruited. According to preregistered exclusion criteria, n = 6 participants were excluded and replaced (n = 1 for a pain threshold exceeding 5 mA, n = 2 for equipment failure, n = 1 for not comprehending the rating scale, n = 1 for reporting exclusion criteria-relevant information following completion of the study, n = 1 from discontinuing the experiment). Furthermore, n = 6 participants were excluded from analyses as they did not attend the second session, resulting in a total sample of n = 54 (Table 2). Participants (34 female, 19 male, 1 undisclosed) were on average 22.20 years old (range = 18–30, SD = 3.06). Time between sessions was 27.98 days (range 25–33, SD = 1.64). For skin conductance analyses, n = 4 participants were excluded because their average stimulus-locked amplitude was lower than 0.02 μS per session, resulting in a sample of n = 50 for all analyses including SCR.

General changes across sessions

Across the two sessions (S1 and S2), the pain threshold increased for 38 participants, decreased for 15 and was the same for one participant. As expected (H1a), the pain threshold significantly increased (mean S1 = 1.29 mA, mean S2 = 1.83 mA), indicating habituation across sessions ($t(53) = -4.10$, $p < 0.001$; Fig. 2). An increase in pain threshold between sessions was not related to participant's change in mean pain ratings between sessions (Pearson's $r = 0.042$, $p = 0.76$). Anxiety did not differ between the start of session S1 as compared to the start of session S2 (mean S1 = 2.72, mean S2 = 2.91, V = 332.5, p = 0.72, see H1b). Self-reported current experienced pain did not differ between sessions (mean S1 = 0.24, mean S2 = 0.30, V = 46.5, p = 0.72). SAM scores decreased from pre- to post-test for both happiness ($F(1,53) = 17.38$, $p < 0.001$) and feelings of control ($F(1,53) = 9.91$, $p = 0.003$). SAM excitement scores were lower in the second compared to first session ($F(1,53) = 5.76$, $p = 0.020$).

The multilevel model demonstrated that in general participants showed a significant increase in pain intensity ratings over trials ($\beta = 0.37$, SE = 0.07, $p < 0.001$), i.e., sensitization, but not over runs ($\beta = -0.59$, SE = 0.45, $p = 0.197$). The interaction between trials and runs ($\beta = -0.07$, SE = 0.024, $p = 0.005$) indicated that this sensitization over trials became smaller over runs (Fig. 3). In contrast to H2a, pain ratings were on average lower in S2 than S1 ($\beta = -1.15$, SE = 0.17, $p < 0.001$). However, as mean ratings in both sessions were similar using a conventional paired t-test (mean S1: 45.73, mean S2: 44.58; $t(53) = 0.80$, $p = 0.43$), this decrease in mean rating appears to be relatively small.

The multilevel model for SCRs showed a significant decrease in SCR over trials ($\beta = -0.078$, SE = 0.007, $p < 0.001$), i.e., habituation, as well as across runs ($\beta = -0.79$, SE = 0.053, $p < 0.001$), with an interaction between trials and runs ($\beta = 0.015$, SE = 0.003, $p < 0.001$) indicating that this decrease became smaller over runs (Fig. 4). Responses in S2 were significantly lower than in S1 ($\beta = -0.19$, SE = 0.025, $p < 0.001$, see H2b).

There was a significant increase in non-response SCR trials across runs ($F(1.69, 82.77) = 46.59$, $p < 0.001$), but no difference between sessions ($F(1,49) = 0.30$, $p = 0.59$) nor an interaction between session and run $F(1.61, 79.02) = 2.09$, $p = 0.14$, see H2c). However, the non-preregistered analysis accounting for the skewed distribution did reveal an increase across runs ($\beta = 0.67$, SE = 0.053, $p < 0.001$), sessions ($\beta = 0.62$, SE = 0.18, $p < 0.001$) and an interaction between run and session ($\beta = -0.23$, SE = 0.07, $p = 0.002$). This suggests that there were more non-response trials in S2 and that the rate of increase in non-response trials across runs was lower in S2 than S1 (Fig. 5).

Test-retest reliability of habituation and sensitization

Pain ratings

There was large inter-individual variability in response patterns in both sessions. In S1, 21 participants showed habituation (H), 17 no change (No), and 16 sensitization (S). Their individual pattern was correctly estimated by 40.7% of the participants in the exit questionnaire. In S2, 22 participants showed habituation, 21 no change, and 11 sensitization, correctly estimated by 50.0% of participants. At the group level, slopes did not differ across sessions ($t(53) = -0.21$, $p = 0.83$; Fig. 6 and H3a). Participants who showed a larger increase in pain threshold from S1 to S2 were more likely to show a higher (i.e., more positive) slope in S2 compared to S1 ($r = 0.31$, $p = 0.023$).

We assessed reliability of pain ratings and habituation/sensitization across sessions using various measures at the individual level (summarized in Table 3). The mean ratings had a good reliability (ICC=0.69). The reliability of the across-run slope between sessions was poor (ICC=0.34; CI(95%): $0.081 < ICC < 0.558$; H3b). Slopes ranged between −0.35–0.24 in S1 and their repeatability coefficient was 0.28, which also indicated a high variance between individual repetitions. The log slope and CV demonstrated fair reliability (ICC 0.45 and 0.50, respectively). Exploratory analyses further demonstrated a higher, yet still moderate, reliability for within-run slopes (ICC 0.45–0.55; see Table 3 and Fig. 7A). Participants categorized as sensitizers stayed in the same category in 73% of the cases, while this was lower for no change (48%) and habituators (12.5%). In summary, reliability estimates varied across measures, with ICC values ranging from 0.34 to 0.62 for habituation measures, and higher ICC values for mean ratings (0.65–0.69, 'good'). The group-level correlation of run 1 ratings showed the highest and most reliable results for participants who were categorized as sensitizers, (Fig. 7B, sensitizers: $r = 0.71$, $p < 0.001$; habituators: $r = 0.18$, $p = 0.19$; no change: $r = 0.18$, $p = 0.20$). Figures of within-run slope reliability for run 2 and 3 can be found in the supplemental material

Table 2
Descriptives of the participants.

	N	Gender	Age	Pain threshold (mA)	Stimulation intensity (mA)	Mean pain intensity	Current pain	FoP	Anxiety (pre/during)	SAM – happiness (pre/post)	SAM – excitement (pre/post)	SAM – control (pre/post)
Session 1	54	19 M, 34 F, 1 undisclosed	22.20	1.29	1.82	45.73	0.24	23.31	2.72/3.30	3.78/3.31	2.94/2.87	3.61/3.32
Session 2				1.83	2.60	44.58	0.30	-	2.91/3.02	3.63/3.35	2.72/2.54	3.43/3.26

FoP = fear of pain; SAM = self-assessment Manikin

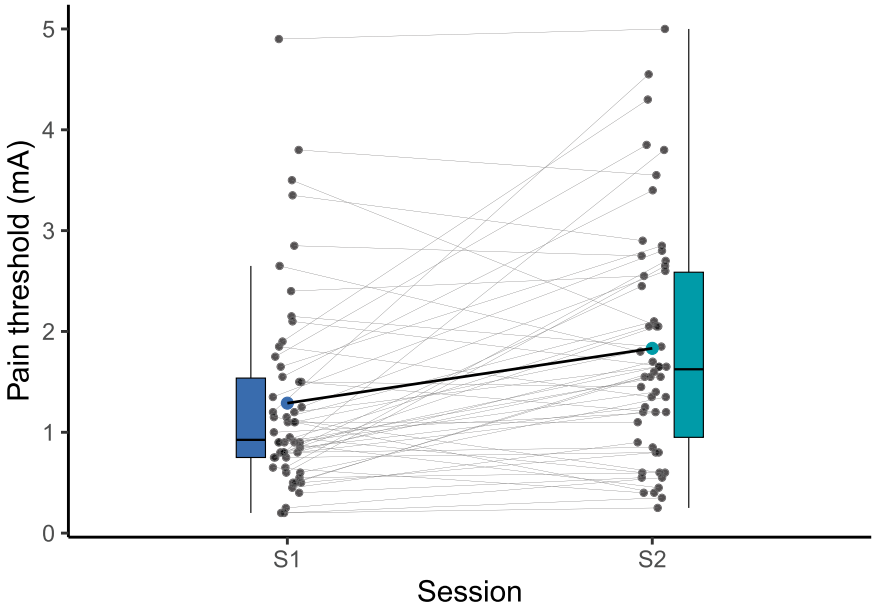

Fig. 2. Change in pain threshold across sessions. Individual participants are reflected by the connected dots across sessions. Boxplots represent the spread and median across participants.

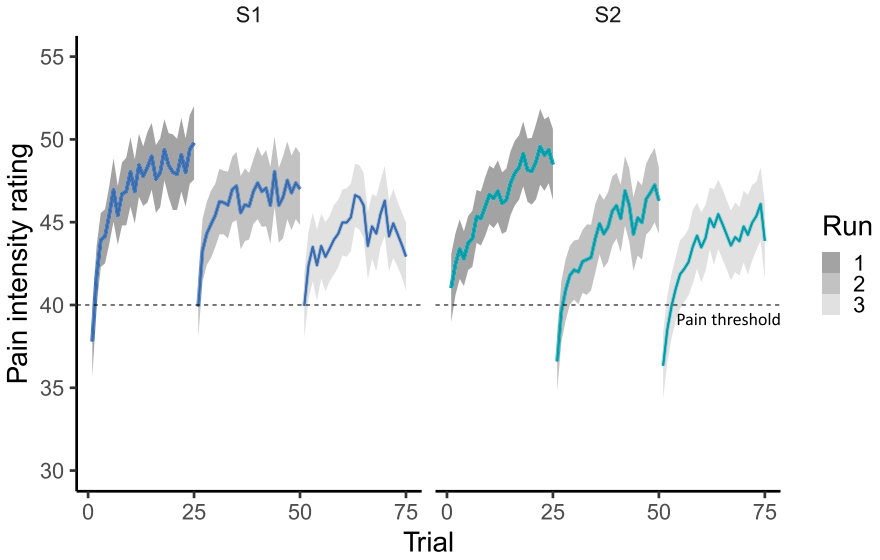

Fig. 3. Average changes of pain intensity for session S1 (left) and session S2 (right) across trials and runs. Gray shading indicates the standard error of the mean, and the gray dashed horizontal line indicates the pain threshold. Recovery of the response is shown during the breaks (i.e., at trials 25 and 50).

(Figure S1 and S2).

Skin conductance responses

Individual analyses demonstrated that in the majority of participants the SCR habituated (Fig. 8). That is, out of 50 participants 44 participants habituated in S1 and 43 participants in S2 (with 6 and 7 participants showing no change, respectively). None of the participants sensitized. On the group level, slopes did not differ across sessions ($t(49) = -0.78$, $p = 0.44$; H3c). Individual reliability of the across-run slope compared between S1 and S2 was poor ($ICC = 0.37$, 95% CI $0.099 < ICC < 0.58$, see H3d). The reliability of within-run slopes across sessions was generally higher (run 1 $ICC = 0.52$, 95% CI $0.29 < ICC < 0.70$; run 2 $ICC = 0.40$, 95% CI $0.14 < ICC < 0.61$; run 3 $ICC = 0.32$, 95% CI $0.046 < ICC < 0.55$), albeit lower for each consecutive run, likely related to an increase in non-response trials.

Relation of questionnaire data with habituation and sensitization of pain ratings

Contrary to our previous work,[11] the use of positive-self statements coping strategies was not predictive of habituation or sensitization ($r = -0.034$, $p = 0.81$, see H4). Anxiety scores at the start of the experiment or during the main task (rated post-experiment) did not correlate with the habituation and sensitization slope (all p's > 0.05). Fear of Pain scores (mean = 23.31, range 9–45) were related to habituation in session 1 ($r = 0.29$, $p = 0.033$), but this was not the case in session 2 ($r = -0.04$, $p = 0.77$). The scores of happiness, excitement and control of the self-assessment Manikin did not show any relation with habituation/sensitization in either session (all p's > 0.05 after correction for multiple comparisons).

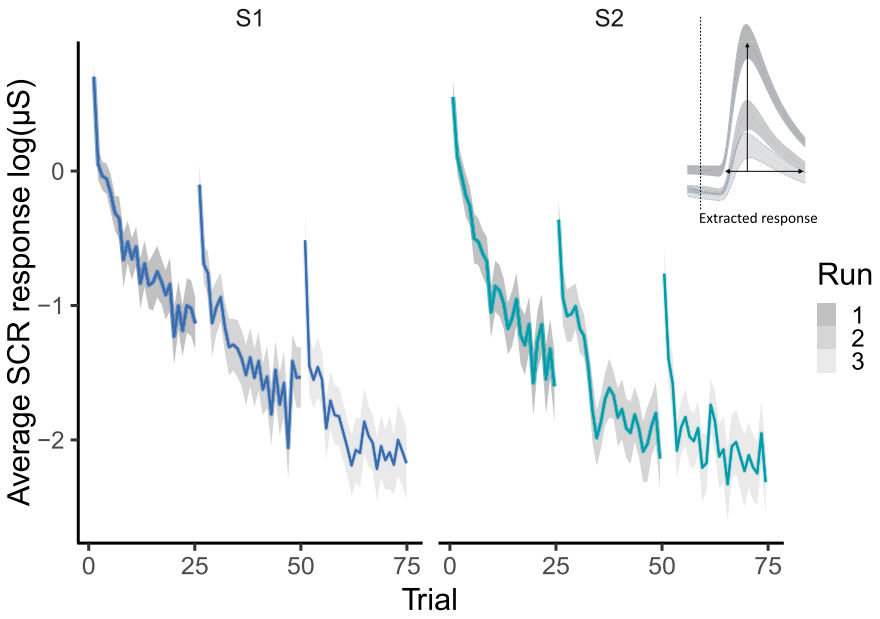

Fig. 4. Average skin conductance response (SCR) on the group level in S1 and S2. Individual responses are calculated in a window 1–6s post-stimulus and then averaged across participants (inset).

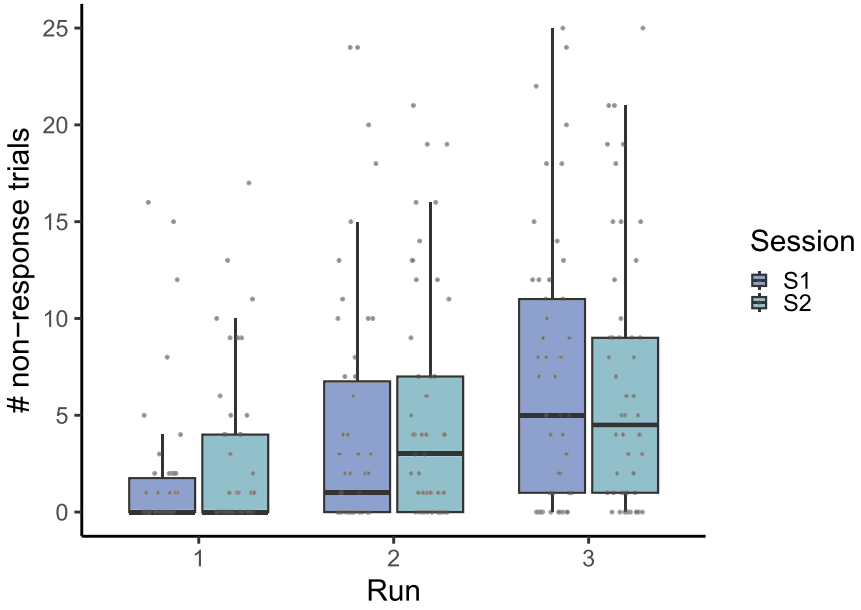

Fig. 5. The number of non-response trials per run for S1 and S2. Individual data is represented by dots. Boxplots represent the spread and median across participants.

Association between self-report ratings and skin conductance responses

Hypothesis H5 assessed the association between self-report ratings and skin conductance responses. Individual data demonstrated that SCRs mostly habituated (H) and, in a few cases, showed no change (No), independent of the ratings (see Fig. 9A for examples of ratings and Fig. 9B for SCRs). Individual correlations per session ranged from −0.77–0.67 (see supplemental Figure S3) and participants who habituated showed more positive correlation coefficients (higher pain rating = higher SCR), while participants who sensitized (S) showed more negative coefficients (higher pain rating = lower SCR). As a control analysis, correlation coefficients were compared across rating categories, revealing significant differences between categories in session S1 ($F(2,47) = 22.76$, $p < 0.001$). Post-hoc comparisons indicated significant differences between S-H ($p < 0.001$) and S-No ($p < 0.001$). The same effects were observed in S2 ($F(2,47) = 7.58$, $p = 0.0014$; categories S-H: $p < 0.001$; S-No: $p = 0.019$). To account for these differences when computing the average group correlation, the absolute values of the individual coefficients were used. After this, SCRs and self-report ratings showed a moderate correlation (S1 $r = 0.27$, $p = 0.06$, S2 $r = 0.21$, $p = 0.14$) and did not differ between S1 and S2 (Zou's CI(95%) −0.29 – 0.40).

As an additional control analysis, we compared the skin conductance slopes for each rating category. These did not differ in either S1 ($F(2,47) = 0.20$, $p = 0.82$) or S2 ($F(2,47) = 0.003$, $p = 0.997$), indicating independent processing of SCRs and habituation/sensitization on the behavioral level. Finally, to rule out the influence of the stimulus intensity (mA), the skin conductance slopes were correlated with the stimulus intensity, which did not result in significant effects (S1: $r = 0.14$, $p = 0.32$; S2: $r = 0.039$, $p = 0.79$).

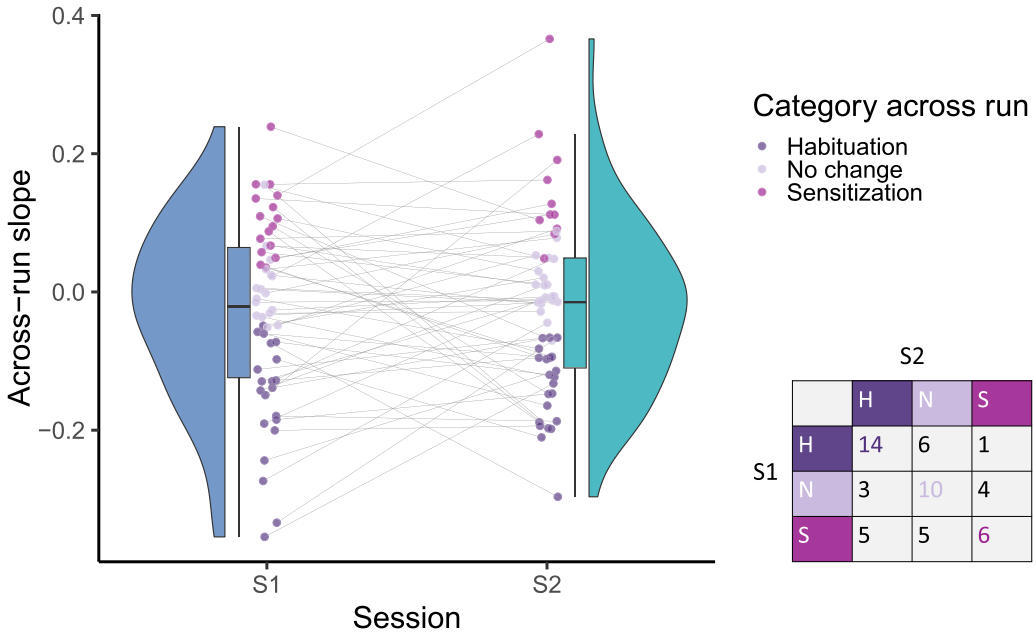

Fig. 6. Change of individual across-run slopes over pain ratings in sessions S1 and S2. Colored points reflect the individual category for each session. Boxplot represents the spread and median of the group level. The table indicates the number of participants per category (S1 rows, S2 columns). H = habituation, N = no change, S = sensitization.

Table 3
Summary of reliability and repeatability estimates for pain ratings. All ICC values reflect reliability across sessions, either for across-run or within-run measures at the participant level.

	ICC linear slope	Repeatability coefficient	ICC log slope	ICC CV	ICC Mean % change	Group-level correlation	ICC mean rating	ICC SD
Across-run	0.34	0.28	0.45	0.50	0.36 1–3 0.49 1–2 −0.07 2–3	0.35	0.69	0.47
Within-run 1	0.55	1.07	0.49	0.35	-	0.47	0.65	0.37
Within-run 2	0.46	0.95	0.48	0.44	-	0.45	0.67	0.44
Within-run 3	0.45	1.02	0.48	0.62	-	0.36	0.69	0.58

ICC = intra-class correlation coefficient, CV = coefficient of variation, SCR = skin conductance response

Discussion

This study investigated changes in self-report and autonomic measures over time and assessed test-retest reliability of habituation and sensitization patterns during repeated painful stimulation. Test-retest reliability was moderate with higher consistency for sensitizing individuals. Pain ratings and SCRs dissociated, with SCRs mainly habituating and pain ratings showing diverse patterns of habituation and sensitization.

General changes across sessions

The pain threshold increased from S1 to S2, which suggests habituation across sessions, corroborating previous findings.[24,41–44] Furthermore, participants reported sensitization across trials, but this effect diminished across runs within a session. Decreased sensitization over time was also previously shown in a similar paradigm[11] and with heat stimulation across days,[45,46] in alignment with early theories of sensitization.[47] SCRs significantly decreased over trials, runs and across sessions, consistent with prior findings.[9,10] While the reduction across sessions was systematic, its magnitude was modest, in line with the observed decrease in pain ratings.

Test-retest reliability

At the group-level, habituation and sensitization slopes did not differ between sessions. However, large individual differences were present, and individual reliability as assessed using the ICC, ranged from poor to fair. The ICC is based on the between-person to within-person variability and can be influenced by experimenter factors, training and homogeneity of the sample.[36] In the current paradigm, we used strict instructions, a script for communication, training for the protocol and the same experimenter across sessions. These measures decrease the between-person variance, leading to relatively higher intraindividual variability and thus potentially lower ICCs.[36] Variance in across-run slopes may have been elevated due to breaks between runs and the resulting recovery of responses. The repeatability coefficient, which is less affected by between-person variance, also indicated large variability between measurements. The mean pain ratings showed good reliability across sessions, consistent with previous findings.[35,36,48–50]

Within-run slope reliability and correlations were high only for participants who sensitized. The variability across sessions shown by participants with a habituating and no change pattern may reflect the ability to adapt to changing stimuli and environments. In this case, low test-retest reliability could thus be a positive sign of well-functioning adaptability and flexibility. In contrast, consistent sensitization, i.e., less variability across sessions, might be a sign of less adaptability. Recent research on cognitive and psychological flexibility showed that

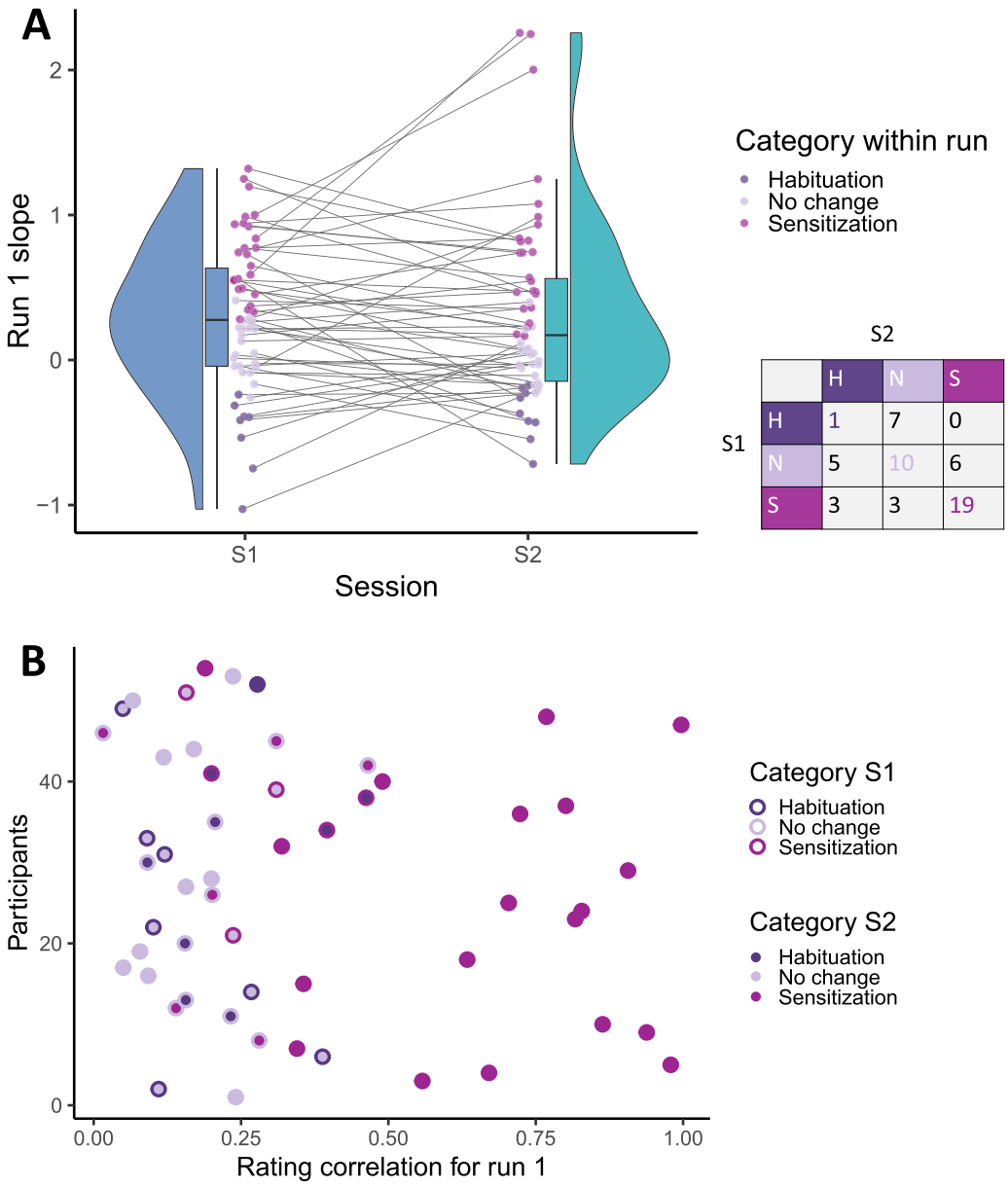

Fig. 7. Within-run 1 reliability. A) Slopes for run 1 across sessions S1 and S2 (left) and number of participants per category (right). H = habituation, N = no change, S = sensitization. B) Correlation between pain ratings of S1 run 1 and S2 run 1. Participants who sensitize (shown in pink color with the outer circle representing S1 and the filled inner circle representing S2) have higher correlations between sessions than the habituation and no change categories, mean r = 0.71 for sensitizers.

some people might be 'stuck' in their pain response, which could predict worse outcomes of chronic pain recovery.[51,52] Patients with acute low back pain that showed restricted corticomotor excitability were also more at risk of developing chronic or recurrent low back pain at 6 months.[53] Patients with chronic pain also demonstrated more variable pain thresholds across sessions than healthy controls.[54] Thus, the mechanisms of flexibility and adaptability in the context of sensitization can be further explored. Studies can also investigate the therapeutic potential of targeting sensitization, although the relation of chronic pain and sensitization remains under debate.[55,56]

A more pessimistic view might be that our measures of linear habituation and sensitization do not fully capture the variance, resulting in low reliability findings. In general, and given the broad range of factors that influence pain, error variance was tried to be minimized. However, we cannot fully exclude that other factors (such as measurement error) may have affected the results. For example, recalibrating may reduce reliability, but it was preferred in our study to maintain a consistent pain percept across sessions.[57] Intra-individual variability in pain responses is generally large, although it is often not reported due to the common practice of averaging responses. Notably, such variability may hold predictive value for chronic pain.[58]

Test-retest reliability for SCRs was poor to fair. The observed ICC values are comparable to previous studies on sympathetic skin responses, despite methodological differences such as stimulation type (heat), intersession interval (10 days), and use of a fixed rather than individualized stimulus.[12] These findings also align with the generally lower reliability of EDA measures compared to pain reports.[59,60]

Predictive factors of habituation and sensitization

Several correlations with habituation were investigated. Our previous study indicated that a higher score on positive self-statements predicted more habituation,[11] but this was not replicated in the current study. Furthermore, the significant correlation with the Fear of Pain

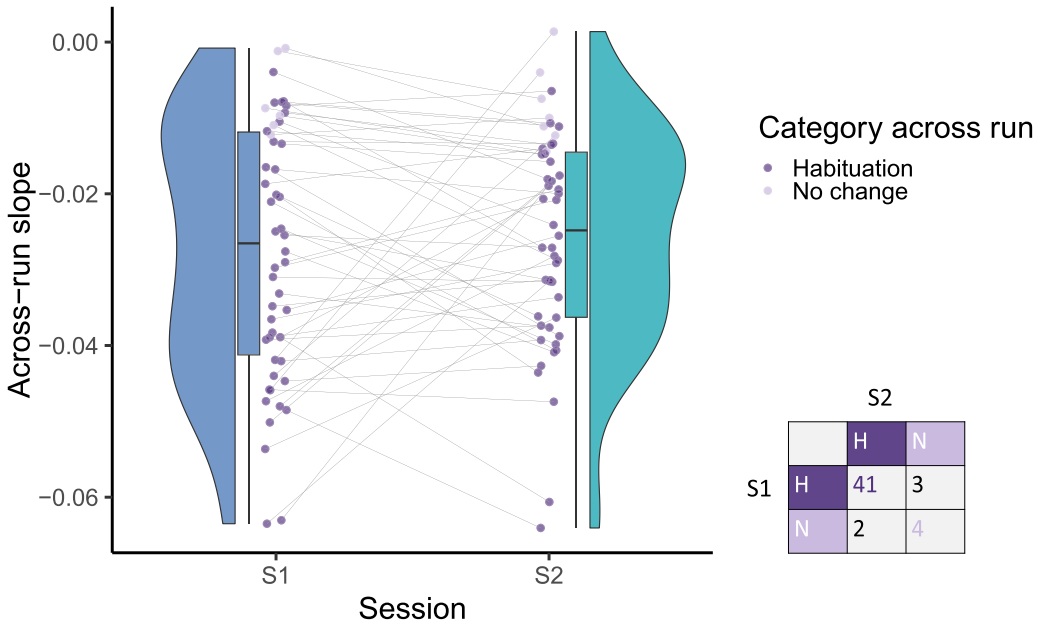

Fig. 8. Change of individual across-run slopes for skin conductance responses over sessions. Colored points reflect the individual category for each session. Boxplot represents the spread and median of the group level. None of the participants sensitized. H = habituation, N = no change.

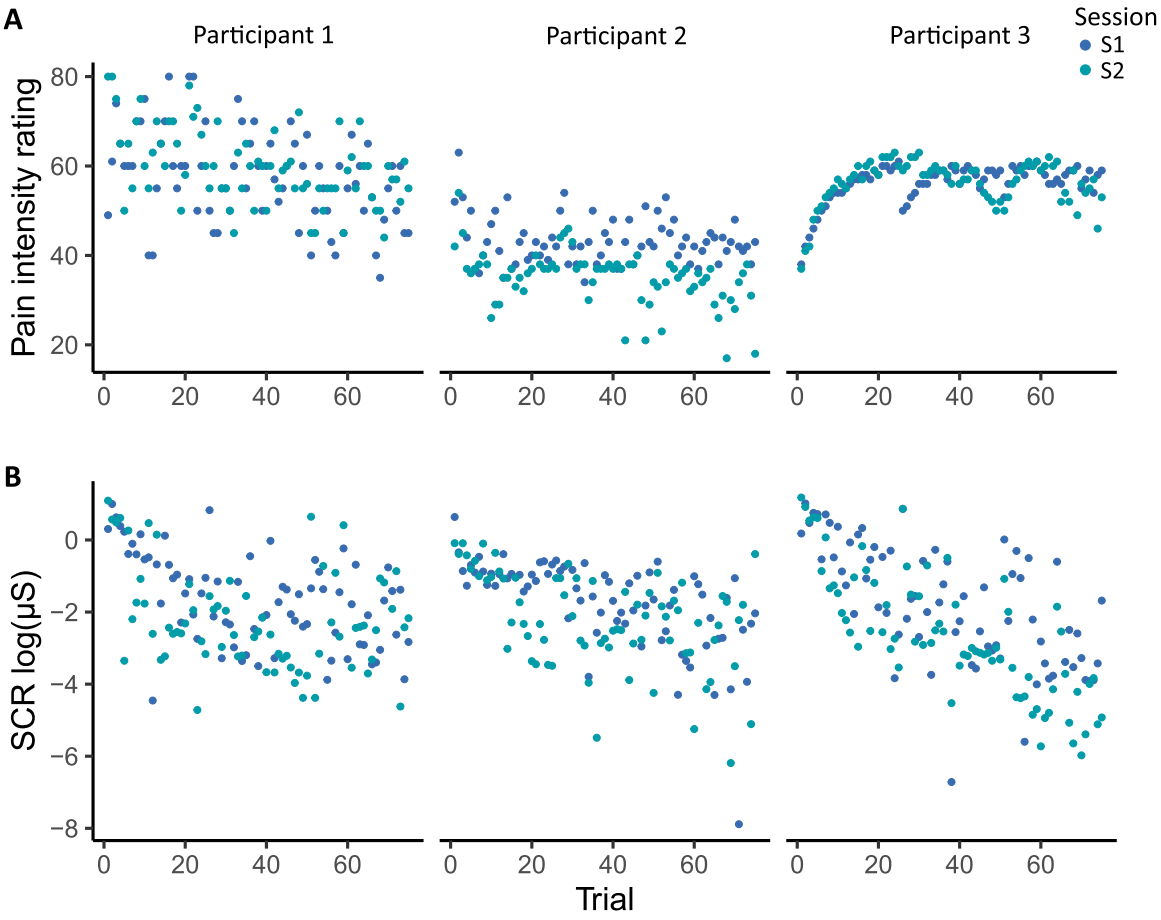

Fig. 9. Individual differences in responses to repeated painful stimulation across sessions for three exemplary participants (columns). A) Pain intensity ratings show habituation (left), no change (middle) or sensitization (right) across runs. B) All SCR responses show habituation across runs.

questionnaire and sensitization was only shown in S1, and not in S2. This could partly be because of the timing of administering the questionnaire (before the task in S1). Habituation of the pain threshold did show a correlation with the slope, i.e., a larger increase in pain threshold indicated a more positive slope in S2 compared to S1. This is in line with the habituation characteristic that a stronger stimulus results in less

pronounced habituation.[1] However, only 9,6% of variance in the slopes was explained by this effect. Habituation/sensitization and the change in slopes from S1 to S2 is likely a complex interplay between multiple variables, including the pain threshold, fear of pain, arousal, expectations and the current state of the participant.

Habituation and sensitization have been proposed to relate to alertness and readiness to respond.[3,61] Current evidence, including our results, remains inconclusive about the verification of these theories. Further studies could explore predictive factors such as alertness, flexibility and adaptability, as well as resilience, which has been associated with habituation of the pain threshold.[62,63]

Association between SCR and ratings

When taking habituation and sensitization differences into account, our findings revealed moderate correlations between SCRs and pain ratings. This result is in line with some earlier findings,[17] although stronger associations have been reported as well.[13,17] Previous work did not investigate the association while considering patterns of habituation and sensitization, or applied only a limited number of stimuli, restricting the ability to detect these processes. Our results show that the correlation is heavily influenced by habituation and sensitization effects, resulting in either positive correlations (habituation of both ratings and SCR) or negative correlations (sensitization of ratings and habituation of the SCR). Two earlier studies also reported negative correlations between pain reports and SCRs,[13,15] which could have been a result of sensitization. Some studies showed a stronger association of SCRs with stimulus intensity than stimulus perception,[18,19] but this was not observed in our study. Our findings suggest that the autonomic nervous system generally exhibits a habituating response to repeated stimulation, while the subjective pain experience may follow a different pattern, including sensitization, indicating a dissociation between both processes. Previous studies have reported weaker or absent associations between SCRs and lower stimulus intensities compared to higher stimulus intensities.[15,17,64] However, our stimuli were calibrated at 50% above the individual pain threshold, initially leading to a painful sensation. Although further evidence is limited, some studies reported a dissociation between perceived pain intensity and brain activation[65] and between habituation of brain activation and stable pain reports.[66] Our study contributes to this dissociation, by revealing that pain reports and SCRs are also uncoupled in the context of habituation and sensitization. Even with a reduction in autonomic arousal, the subjective experience may still show sensitization. A neurophysiological explanation might be that pain-related brain pathways (e.g., somatosensory cortices, cingulate cortex, thalamus)[67] could remain active and affected by cognitive and emotional pathways, influencing pain perception, while the autonomic nervous system shows distinct patterns[68] signaling reduced threat leading to a habituating response.

Limitations and challenges

This study has a few limitations. First, our participants were mostly English-speaking and completed the English version of the coping strategy questionnaire, whereas our previous study used the Dutch version.[28,29] This may have introduced inconsistencies across studies and contributed to the divergent findings. In addition, the questionnaire was completed at the beginning of the experiment and therefore was not directly linked to participants' experienced pain. Second, we did not assess pain-related expectations or specific response patterns, to avoid response bias when introducing these concepts to participants. Third, identifying an appropriate metric to characterize patterns of habituation and sensitization remains a challenge. Common approaches include using linear slopes and mean percentage changes,[69] which differ depending on whether across-run or within-run analyses are employed.[11] The present findings indicate that no single measure provides a comprehensive account of these processes. Finally, our sample was very homogeneous in terms of being healthy, young, and not diagnosed with any disorders. A more representative sample might show more variability in responses. Especially since factors like depression, sleep and anxiety have strong relations to experiencing pain. The lack of change in anxiety across sessions, for example, might be due to the homogenous sample. Therefore, the generalizability of our findings should be evaluated in more diverse and representative samples.

Conclusion

Repeated painful stimulation results in large individual differences with, at best, moderate test-retest reliability. Varying response profiles across sessions, i.e., habituation and sensitization, could reflect the ability to adapt to threatening stimuli under different circumstances. Sensitization patterns were more robust and showed a dissociation with the habituating SCR. Given these diverging patterns, it is important to measure both subjective and physiological responses, especially when habituation and sensitization occur.

Study preregistration

The study was preregistered prior to data collection and the registration can be found at: https://osf.io/8a7qn.

CRediT authorship contribution statement

Maite van der Miesen: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – Original Draft, Writing – Review & Editing, Supervision, Visualization; Carine Vossen: Methodology, Writing – Review & Editing, Supervision; Judith Eck: Methodology, Writing – Review & Editing, Supervision; Sophie Kühne: Investigation, Writing – Review & Editing, Visualization; Elbert Joosten: Methodology, Resources, Writing – Review & Editing, Supervision; David Linden: Conceptualization, Methodology, Resources, Writing – Review & Editing, Supervision; Judith Peters: Conceptualization, Methodology, Writing – Review & Editing, Supervision.

Declaration of Competing Interest

No external funding was provided. The authors have no conflict of interest to declare.

Acknowledgements

We would like to thank Dr. Inge Timmers for guidance with the design and skin conductance preprocessing and Dr. Rachel Sjouwerman for input on skin conductance analyses. We thank Eleonora Marazita and Cristina Bonet Olivares for assisting with data collection.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jpain.2025.105557.

Data availability

The data and analysis code are available at: https://osf.io/2emsq.

References

1. Rankin CH, Abrams T, Barry RJ, et al. Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. *Neurobiol Learn Mem.* 2009;92(2):135–138. https://doi.org/10.1016/j.nlm.2008.09.012.
2. van den Broeke EN, Crombez G, Vlaeyen JW. Reconceptualizing sensitization in pain: back to basics. *Pain Rep.* 2024;9(1), e1125. https://doi.org/10.1097/PR9.0000000000001125.

3. Eisenstein EM, Eisenstein D, Smith JC. The evolutionary significance of habituation and sensitization across phylogeny: A behavioral homeostasis model. *Integr Physiol Behav Sci*. 2001;36(4):251–265. https://doi.org/10.1007/BF02688794.
4. Christoffersen G. Habituation: events in the history of its characterization and linkage to synaptic depression. A new proposed kinetic criterion for its identification. *Prog Neurobiol*. 1997;53(1):45–66.
5. van der Miesen MM, Joosten EA, Kaas AL, et al. Habituation to pain: self-report, electroencephalography, and functional magnetic resonance imaging in healthy individuals. A scoping review and future recommendations. *PAIN*. 2024;165(3):500–522. https://doi.org/10.1097/j.pain.0000000000003052.
6. De Paepe AL, Williams ACC, Crombez G. Habituation to pain: a motivational-ethological perspective. *Pain*. 2019;160(8):1693–1697. https://doi.org/10.1097/j.pain.0000000000001533.
7. Breimhorst M, Hondrich M, Rebhorn C, et al. Sensory and sympathetic correlates of heat pain sensitization and habituation in men and women. *Eur J Pain*. 2012;16(9):1281–1292. https://doi.org/10.1002/j.1532-2149.2012.00133.x.
8. Arntz A, Dreessen L, Merckelbach H. Attention, not anxiety, influences pain. *Behav Res Ther*. 1991;29(1):41–50. https://doi.org/10.1016/s0005-7967(09)80006-5.
9. Mobascher A, Brinkmeyer J, Warbrick T, et al. Fluctuations in electrodermal activity reveal variations in single trial brain responses to painful laser stimuli - A fMRI/EEG study. *Neuroimage*. 2009;44(3):1081–1092. https://doi.org/10.1016/j.neuroimage.2008.09.004.
10. Mobascher A, Brinkmeyer J, Warbrick T, et al. Brain activation patterns underlying fast habituation to painful laser stimuli. *Int J Psychophysiol*. 2010;75(1):16–24. https://doi.org/10.1016/j.ijpsycho.2009.10.008.
11. van der Miesen MM, Vossen CJ, Compère VN, et al. Individual differences in response to repeated painful stimulation: Habituation, sensitization, and nocebo effects. *J Pain*. 2025, 105377.
12. De Schoenmacker I, Leu C, Curt A, Hubli M. Pain-autonomic interaction is a reliable measure of pain habituation in healthy subjects. *Eur J Pain*. 2022;26(8):1679–1690. https://doi.org/10.1002/ejp.1990.
13. Chapman CR, Nakamura Y, Donaldson GW, et al. Sensory and affective dimensions of phasic pain are indistinguishable in the self-report and psychophysiology of normal laboratory subjects. *J Pain*. 2001;2(5):279–294. https://doi.org/10.1054/jpai.2001.25529.
14. Breimhorst M, Sandrock S, Fechir M, et al. Do intensity ratings and skin conductance responses reliably discriminate between different stimulus intensities in experimentally induced pain? *J Pain*. 2011;12(1):61–70. https://doi.org/10.1016/j.jpain.2010.04.012.
15. Loggia ML, Juneau M, Bushnell CM. Autonomic responses to heat pain: Heart rate, skin conductance, and their relation to verbal ratings and stimulus intensity. *PAIN*. 2011;152(3).
16. Treister R, Kliger M, Zuckerman G, et al. Differentiating between heat pain intensities: The combined effect of multiple autonomic parameters. *PAIN*. 2012;153(9):1807–1814. https://doi.org/10.1016/j.pain.2012.04.008.
17. Geuter S, Gamer M, Onat S, Büchel C. Parametric trial-by-trial prediction of pain by easily available physiological measures. *PAIN*. 2014;155(5):994–1001. https://doi.org/10.1016/j.pain.2014.02.005.
18. Nickel MM, May ES, Tiemann L, et al. Autonomic responses to tonic pain are more closely related to stimulus intensity than to pain intensity. *PAIN*. 2017;158(11):2129–2136. https://doi.org/10.1097/j.pain.0000000000001010.
19. Posada–Quintero HF, Kong Y, Chon KH. Objective pain stimulation intensity and pain sensation assessment using machine learning classification and regression based on electrodermal activity. *Am J PhysiolRegul Integr Comp Physiol*. 2021;321(2):R186–R196. https://doi.org/10.1152/ajpregu.00094.2021.
20. Mischkowski DP-B, E E, Banker L, et al. Pain or nociception? Subjective experience mediates the effects of acute noxious heat on autonomic responses. *Pain*. 2018;159(4):699–711. https://doi.org/10.1097/j.pain.0000000000001132.
21. De Schoenmacker I, Scheuren PS, Sirucek L, et al. Understanding inter-individual variability of experimental pain habituation and conditioned pain modulation in healthy individuals. *Sci Rep*. 2024;14(1), 22070. https://doi.org/10.1038/s41598-024-73158-5.
22. Vossen CJ, Luijcks R, van Os J, et al. Does pain hypervigilance further impact the lack of habituation to pain in individuals with chronic pain? A cross-sectional pain ERP study. *J Pain Res*. 2018;11:395–405. https://doi.org/10.2147/jpr.S146916.
23. Society for Psychophysiological Research, Ad Hoc Committee on Electrodermal Measures, Boucsein W, et al. Publication recommendations for electrodermal measurements. *Psychophysiology*. 2012;49(8):1017–1034. https://doi.org/10.1111/j.1469-8986.2012.01384.x.
24. Bingel U, Schoell E, Herken W, et al. Habituation to painful stimulation involves the antinociceptive system. *Pain*. 2007;131(1-2):21–30. https://doi.org/10.1016/j.pain.2006.12.005.
25. Bauch EM, Andreou C, Rausch VH, Bunzeck N. Neural Habituation to Painful Stimuli Is Modulated by Dopamine: Evidence from a Pharmacological fMRI Study. *Front Hum Neurosci*. 2017;11:630. https://doi.org/10.3389/fnhum.2017.00630.
26. Payne AFH, Schell AM, Dawson ME. Lapses in skin conductance responding across anatomical sites: Comparison of fingers, feet, forehead, and wrist. *Psychophysiology*. 2016;53(7):1084–1092. https://doi.org/10.1111/psyp.12643.
27. McNeil DW, Kennedy SG, Randall CL, et al. Fear of Pain Questionnaire-9: Brief assessment of pain-related fear and anxiety. *Eur J Pain*. 2018;22(1):39–48. https://doi.org/10.1002/ejp.1074.
28. Rosenstiel AK, Keefe FJ. The use of coping strategies in chronic low back pain patients: relationship to patient characteristics and current adjustment. *Pain*. 1983;17(1):33–44. https://doi.org/10.1016/0304-3959(83)90125-2.
29. Spinhoven P, Ter Kuile MM, Linssen ACG, Gazendam B. Pain coping strategies in a Dutch population of chronic low back pain patients. *Pain*. 1989;37(1):77–83. https://doi.org/10.1016/0304-3959(89)90155-3.
30. Bradley MM, Lang PJ. Measuring emotion: The self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry*. 1994;25(1):49–59. https://doi.org/10.1016/0005-7916(94)90063-9.
31. Benedek M, Kaernbach C. A continuous measure of phasic electrodermal activity. *J Neurosci Methods*. 2010;190(1):80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028.
32. 2023. ⟨https://CRAN.R-project.org/package=nlme⟩. Linear Nonlinear Mixed Eff Models R Package Version 3 1164.
33. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw*. 2015;67(1):1–48. https://doi.org/10.18637/jss.v067.i01.
34. 2019. irr Var Coeff Inter Reliab Agreem R Package Version 0 84 1, 2019. ⟨https://CRAN.R-project.org/package=irr⟩.
35. Letzen JE, Sevel LS, Gay CW, et al. Test-Retest Reliability of Pain-Related Brain Activity in Healthy Controls Undergoing Experimental Thermal Pain. *J Pain*. 2014;15(10):1008–1014. https://doi.org/10.1016/j.jpain.2014.06.011.
36. Woo C-W, Wager TD. What reliability can and cannot tell us about pain report and pain neuroimaging. *PAIN*. 2016;157(3):511–513. https://doi.org/10.1097/j.pain.0000000000000442.
37. 2024. DescTools Tools Descr Stat R Package Version 0 99 54, 2024. ⟨https://CRAN.R-project.org/package=DescTools⟩.
38. Corey DM, Dunlap WP, Burke MJ. Averaging Correlations: Expected Values and Bias in Combined Pearson rs and Fisher's z Transformations. *J Gen Psychol*. 1998;125(3):245–261. https://doi.org/10.1080/00221309809595548.
39. Zou GY. Toward using confidence intervals to compare correlations. *Psychol Methods*. 2007;12(4):399.
40. Diedenhofen B, Musch J. cocor: A comprehensive solution for the statistical comparison of correlations. *PLOS ONE*. 2015;10(4), e0121945. https://doi.org/10.1371/journal.pone.0121945.
41. Smith BW, Tooley EM, Montague EQ, et al. Habituation and sensitization to heat and cold pain in women with fibromyalgia and healthy controls. *Pain*. 2008;140(3):420–428. https://doi.org/10.1016/j.pain.2008.09.018.
42. Rennefeld C, Wiech K, Schoell ED, et al. Habituation to pain: further support for a central component. *Pain*. 2010;148(3):503–508. https://doi.org/10.1016/j.pain.2009.12.014.
43. Agostinho CM, Scherens A, Richter H, et al. Habituation and short-term repeatability of thermal testing in healthy human subjects and patients with chronic non-neuropathic pain. *Eur J Pain*. 2009;13(8):779–785. https://doi.org/10.1016/j.ejpain.2008.10.002.
44. Rodriguez-Raecke R, Doganci B, Breimhorst M, et al. Insular cortex activity is associated with effects of negative expectation on nociceptive long-term habituation. *J Neurosci*. 2010;30(34):11363–11368. https://doi.org/10.1523/jneurosci.2197-10.2010.
45. May A, Rodriguez-Raecke R, Schulte A, et al. Within-session sensitization and between-session habituation: a robust physiological response to repetitive painful heat stimulation. *Eur J Pain*. 2012;16(3):401–409. https://doi.org/10.1002/j.1532-2149.2011.00023.x.
46. Ellerbrock I, Wiehler A, Arndt M, May A. Nocebo context modulates long-term habituation to heat pain and influences functional connectivity of the operculum. *Pain*. 2015;156(11):2222–2233. https://doi.org/10.1097/j.pain.0000000000000297.
47. Groves PM, Thompson RF. Habituation: a dual-process theory. *Psychol Rev*. 1970;77(5):419. https://doi.org/10.1037/h0029810.
48. Letzen JE, Boissoneault J, Sevel LS, Robinson ME. Test-retest reliability of pain-related functional brain connectivity compared with pain self-report. *PAIN*. 2016;157(3):546–551. https://doi.org/10.1097/j.pain.0000000000000356.
49. Colloca L, Benedetti F, Pollo A. Repeatability of autonomic responses to pain anticipation and pain stimulation. *Eur J Pain*. 2006;10(7):659–665. https://doi.org/10.1016/j.ejpain.2005.10.009.
50. Bi Y, Hou X, Zhong J, Hu L. Test–retest reliability of laser evoked pain perception and fMRI BOLD responses. *Sci Rep*. 2021;11(1):1322. https://doi.org/10.1038/s41598-020-79196-z.
51. Borsook D, Youssef AM, Simons L, et al. When pain gets stuck: the evolution of pain chronification and treatment resistance. *PAIN*. 2018;159(12):2421–2436. https://doi.org/10.1097/j.pain.0000000000001401.
52. Åkerblom S, Perrin S, Rivano Fischer M, McCracken LM. Predictors and mediators of outcome in cognitive behavioral therapy for chronic pain: the contributions of psychological flexibility. *J Behav Med*. 2021;44(1):111–122. https://doi.org/10.1007/s10865-020-00168-9.
53. Jenkins LC, Chang W-J, Buscemi V, et al. Cortical function and sensorimotor plasticity are prognostic factors associated with future low back pain after an acute episode: the Understanding persistent Pain Where it ResiDes prospective cohort study. *PAIN*. 2023;164(1):14–26. https://doi.org/10.1097/j.pain.0000000000002684.
54. Lund Rpt I, Lundeberg T, Kowalski J, et al. Evaluation of variations in sensory and pain threshold assessments by electrocutaneous stimulation. *Physiother Theory Pract*. 2005;21(2):81–92. https://doi.org/10.1080/09593980590922307.
55. De Schoenmacker I, Costa Marques D, Scheuren PS, et al. Novel neurophysiological evidence for preserved pain habituation across chronic pain conditions. *Clin Neurophysiol*. 2024;166:31–42. https://doi.org/10.1016/j.clinph.2024.07.007.
56. van der Miesen MM, Vossen CJ, Joosten EA. Habituation to Pain in Patients with Chronic Pain: Clinical Implications and Future Directions. *J Clin Med*. 2023;12(13):4305. https://doi.org/10.3390/jcm12134305.

57. Adamczyk WM, Szikszay TM, Nahman-Averbuch H, et al. To Calibrate or not to Calibrate? A Methodological Dilemma in Experimental Pain Research. *J Pain*. 2022; 23(11):1823–1832. https://doi.org/10.1016/j.jpain.2022.07.007.
58. Madden VJ, Kamerman PR, Catley MJ, et al. Variability in experimental pain studies: nuisance or opportunity? *Br J Anaesth*. 2021;126(2):e61–e64. https://doi.org/10.1016/j.bja.2020.11.005.
59. Cooper SE, Dunsmoor JE, Koval KA, et al. Test–retest reliability of human threat conditioning and generalization across a 1-to-2-week interval. *Psychophysiology*. 2023;60(6), e14242. https://doi.org/10.1111/psyp.14242.
60. Klingelhöfer-Jens M, Ehlers MR, Kuhn M, et al. Robust group- but limited individual-level (longitudinal) reliability and insights into cross-phases response prediction of conditioned fear. *eLife*. 2022;11, e78717. https://doi.org/10.7554/eLife.78717.
61. Eisenstein EM, Eisenstein D. A Behavioral Homeostasis Theory of Habituation and Sensitization: II. Further Developments and Predictions. *Rev Neurosci*. 2006;17(5): 533–558. https://doi.org/10.1515/REVNEURO.2006.17.5.533.
62. Waldeck D, Pancani L, Holliman A, et al. Adaptability and psychological flexibility: overlapping constructs? *J Context Behav Sci*. 2021;19:72–78. https://doi.org/10.1016/j.jcbs.2021.01.002.
63. Smith BW, Tooley EM, Montague EQ, et al. The role of resilience and purpose in life in habituation to heat and cold pain. *J Pain*. 2009;10(5):493–500. https://doi.org/10.1016/j.jpain.2008.11.007.
64. Misohkowski DP-B, E E, Banker L, et al. Pain or nociception? Subjective experience mediates the effects of acute noxious heat on autonomic responses - corrected and republished. *Artic Pain*. 2019;160(6):1469–1481. https://doi.org/10.1097/j.pain.0000000000001573.
65. Hoeppli ME, Nahman-Averbuch H, Hinkle WA, et al. Dissociation between individual differences in self-reported pain intensity and underlying fMRI brain activation. *Nat Commun*. 2022;13(1):3569. https://doi.org/10.1038/s41467-022-31039-3.
66. Paul K, Tik M, Hahn A, et al. Give me a pain that I am used to: distinct habituation patterns to painful and non-painful stimulation. *Sci Rep*. 2021;11(1), 22929. https://doi.org/10.1038/s41598-021-01881-4.
67. Xu A, Larsen B, Baller EB, et al. Convergent neural representations of experimentally-induced acute pain in healthy volunteers: A large-scale fMRI meta-analysis. *Neurosci Biobehav Rev*. 2020;112:300–323. https://doi.org/10.1016/j.neubiorev.2020.01.004.
68. Piché M, Arsenault M, Rainville P. Dissection of perceptual, motor and autonomic components of brain activity evoked by noxious stimulation. *PAIN*. 2010;149(3): 453–462. https://doi.org/10.1016/j.pain.2010.01.005.
69. McDiarmid TA, Bernardos AC, Rankin CH. Habituation is altered in neuropsychiatric disorders-A comprehensive review with recommendations for experimental design and analysis. *Neurosci Biobehav Rev*. 2017;80:286–305. https://doi.org/10.1016/j.neubiorev.2017.05.028.