

Associative learning and extinction of conditioned threat predictors across sensory modalities

Laura. R. Koenen¹, Robert. J. Pawlik², Adriane Icenhour³, Liubov Petrakova², Katarina Forkmann³, Nina Theysohn⁴, Harald Engler¹ & Sigrid Elsenbruch^{1,2,3}✉

The formation and persistence of negative pain-related expectations by classical conditioning remain incompletely understood. We elucidated behavioural and neural correlates involved in the acquisition and extinction of negative expectations towards different threats across sensory modalities. In two complementary functional magnetic resonance imaging studies in healthy humans, differential conditioning paradigms combined interoceptive visceral pain with somatic pain (study 1) and aversive tone (study 2) as exteroceptive threats. Conditioned responses to interoceptive threat predictors were enhanced in both studies, consistently involving the insula and cingulate cortex. Interoceptive threats had a greater impact on extinction efficacy, resulting in disruption of ongoing extinction (study 1), and selective resurgence of interoceptive CS-US associations after complete extinction (study 2). In the face of multiple threats, we preferentially learn, store, and remember interoceptive danger signals. As key mediators of placebo effects, conditioned responses may be particularly relevant to clinical conditions involving disturbed interoception and chronic visceral pain.

¹Institute of Medical Psychology and Behavioral Immunobiology, Center for Translational Neuro- and Behavioral Sciences, University Hospital Essen, University of Duisburg-Essen, Essen, Germany. ²Department of Medical Psychology and Medical Sociology, Faculty of Medicine, Ruhr University Bochum, Bochum, Germany. ³Translational Pain Research Unit, Department of Neurology, University Hospital Essen, University of Duisburg-Essen, Essen, Germany. ⁴Institute of Diagnostic and Interventional Radiology and Neuroradiology, University Hospital Essen, University of Duisburg-Essen, Essen, Germany. ✉email: sigrid.elsenbruch@ruhr-uni-bochum.de

Expectations shape our experience of reality, and are essential for adaptive behaviour in any complex environment. In the face of danger, negative expectations are formed and dynamically updated by experience, involving associative learning and memory processes. As a key emotional response during the expectation of threat, conditioned fear is essential to trigger adaptive escape or avoidance responses. Learned fear can however also turn maladaptive and contribute to pathology, as underscored by knowledge from fear conditioning accomplished in the context of anxiety, psychological trauma, and stress-related disorders¹. More recently, the scope has been broadened to acute and chronic pain², embedded within the fear-avoidance model³. In keeping with its biological salience, pain is a ubiquitous and fundamentally threatening experience. Interoceptive pain arising from visceral organs appears to be particularly threatening and fear-inducing^{4,5}, resulting in much suffering in highly prevalent disorders of the gut–brain axis like the irritable bowel syndrome (IBS)⁶. Interoceptive, visceral pain is highly modifiable by cognitions and emotions^{6,7}, including negative expectations during the anticipation of pain as key mediators of nocebo effects^{8–12}. Despite broad clinical implications of nocebo effects reaching far beyond chronic pain^{13–15}, the formation and persistence of negative pain-related expectations by classical conditioning remain incompletely understood, especially with respect to neurobiological mechanisms and their possible specificity to threat modality.

Human fear conditioning studies with experimental pain as unconditioned stimuli (US) have implemented exteroceptive, somatic^{16–19} or interoceptive, visceral pain as salient threats^{20–23}, but knowledge about common and distinct threat-specific neural mechanisms remains limited, especially regarding the extinction and retrieval of pain-related fear memories^{24,25}. Combining interoceptive and exteroceptive threats from different sensory modalities, as accomplished herein, constitutes a unique opportunity to elucidate specificity to threat modality in a clinically-relevant context^{4,26}, and is timely given recent conceptual advances regarding interoception^{27,28} and interoceptive psychopathology²⁹. The experience of multiple threats from different sensory modalities closely mimics the clinical reality of patients with diverse symptoms, especially those with complex comorbidities as they often characterise patients with chronic pain. In fact, any normal environment presents multiple salient threats, which may impact learning and memory processes relevant to nocebo effects that remain incompletely understood even in healthy individuals. Existing studies with multiple threats support the role of the insula, a key region of the salience network, in sensory modality-specific effects underlying aversive expectancy^{26,30–32}. The engagement of the salience network, together with regions of the fear and extinction networks, remains to be tested not only for the formation but especially for the extinction of conditioned responses to multiple threats. Conditioned negative expectations may be markedly resistant to extinction, as suggested by studies involving somatic pain stimuli^{33,34}. This may be particularly the case for interoceptive memory traces, as suggested by early classical interoceptive conditioning studies carried out by soviet psychologists³⁵, complemented by modern approaches on fear learning of interoceptive and exteroceptive cues³⁶ and on the partially distinct neural representation of aversive visceral signals³⁷. Impaired extinction efficacy and other phenomena related to memory processes can reportedly facilitate the return of fear and increase the risk of relapse³⁸, with broad implications for the chronicity and treatment of pain and fear-related disorders³⁹.

We herein elucidated the behavioural and neural mechanisms involved in the acquisition and extinction of negative expectations towards different types of interoceptive and exteroceptive

threats across sensory modalities. To this end, we analysed data from two independent differential fear conditioning studies with methodological and conceptual overlap, allowing to assess reproducibility, and offering converging insight into conditioned anticipatory responses to threats from different sensory modalities. In both functional magnetic resonance imaging (fMRI) studies, visceral pain induced by rectal distension was implemented as clinically-relevant interoceptive US together with an exteroceptive US, which was either an equally painful thermal stimulus (study 1), or a non-nociceptive, yet equally aversive tone (study 2). In addition to threat modality-specific predictive cues (conditioned stimuli, CS), unpaired safety cues in both studies allowed us to compare conditioned differential responses to interoceptive versus exteroceptive threat predictors for different phases of conditioning. For the acquisition, we tested the general hypothesis that in the face of multiple threats, predictive learning is shaped by the salience of the US, as suggested by preparedness theory⁴⁰ and the evolutionary significance of the interoceptive modality, as illustrated by one-trial learning phenomena like conditioned nausea and taste aversion⁴¹. Given initial evidence that pain-modality shapes not only the perception and processing of stimuli^{4,5,42}, but also anticipatory responses including conditioned fear^{26,36}, we expected greater differential conditioned responses involving regions of the fear and salience networks to cues predicting interoceptive threat. We further tested whether conditioned responses to interoceptive threat predictors are more resistant to effective extinction, involving regions of the extinction network. The return of conditioned responses induced by reinstatement, i.e. unexpected re-exposure to the US, constitutes a promising translational tool to assess extinction efficacy³⁸ that has rarely been applied in brain imaging studies on pain-related fear²⁰. To this end, our paradigms incorporated different reinstatement procedures following extinction phases, allowing us to test in reinstatement-test phases if the interoceptive CS–US association is more susceptible to reinstatement effects.

In sum, conditioned responses to interoceptive threat predictors were enhanced in both studies after the acquisition, consistently involving the insula and cingulate cortex as key regions of the salience network. Our results supported that unexpected exposure to interoceptive threats had a greater impact on extinction efficacy, resulting in disruption of ongoing extinction (study 1), and selective resurgence of interoceptive CS–US associations after complete extinction (study 2). Together, our findings are an important step towards unravelling how negative expectations are shaped by associative learning and memory processes in the face of multiple threats. A more refined understanding of conditioned nocebo effects in the context of clinically-relevant interoceptive and exteroceptive threats may ultimately contribute to an improved consideration of expectancy effects to the benefit of patients, broaden the rapidly evolving scope of the gut–brain axis in health neuroscience and disease^{43–45}, and fits into a framework of the cognitive neurosciences interfacing between mind and body²⁷.

Results

Participants. Out of a total of $N = 77$ healthy adults who participated, $N = 12$ were excluded due to technical difficulties with MRI data acquisition ($N = 6$), movement artefacts ($N = 4$), or failure to reach visceral pain threshold within predetermined maximal distension pressure ($N = 2$). As a result, we herein report on data from $N = 42$ volunteers for study 1 (all female, age 34.5 ± 2.0 years; BMI 22.7 ± 0.4 kg/m²), and $N = 23$ volunteers for study 2 (10 female, age 26.7 ± 1.0 years; BMI 22.4 ± 0.7 kg/m²). Consistent with stringent and highly-parallelised exclusion

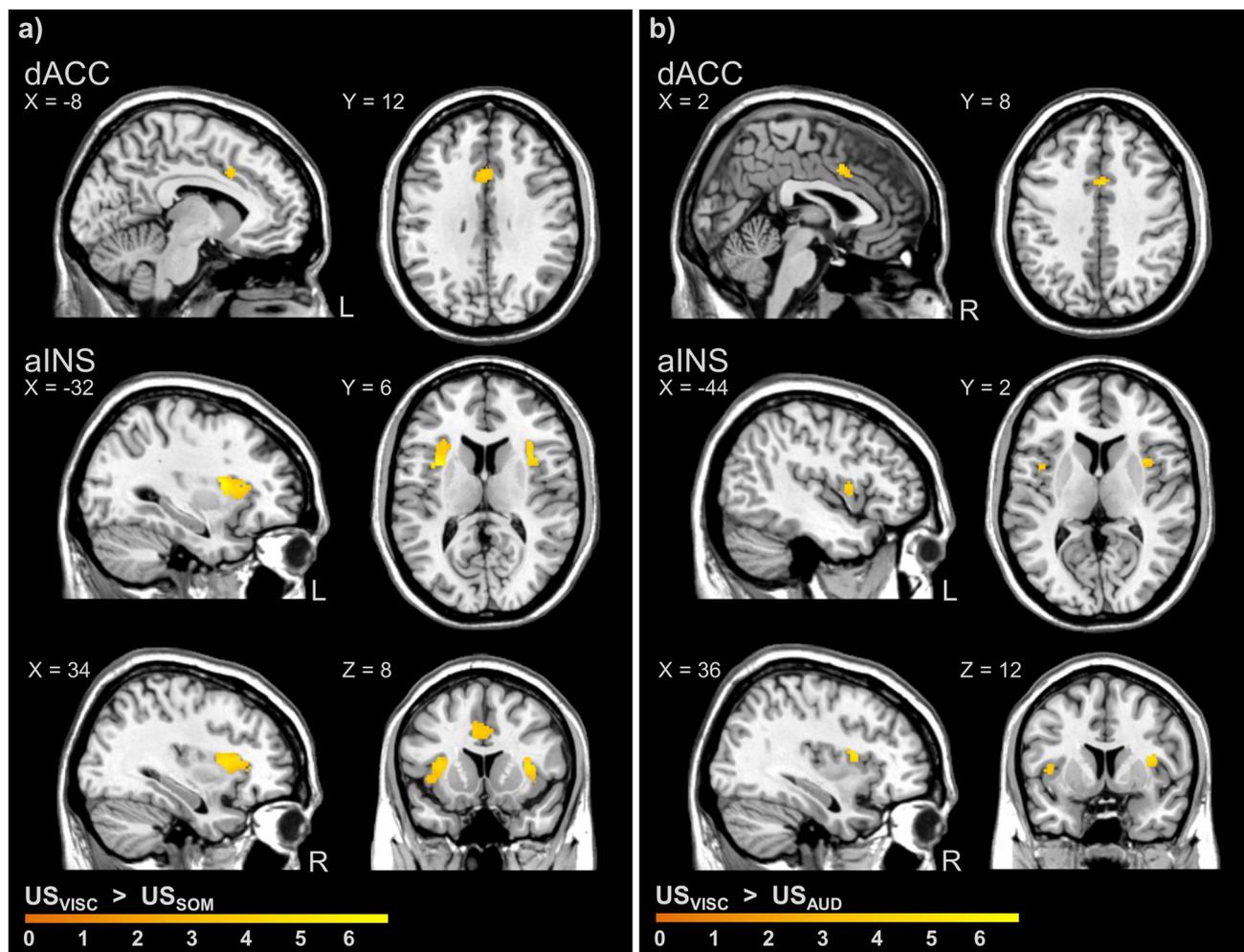


Fig. 1 Differences in neural activation induced by interoceptive versus exteroceptive threats implemented as unconditioned stimuli (US) during acquisition in studies 1 and 2. Interoceptive threat (US_{VISC}) induced significantly greater neural activation in dACC and aINS compared to both exteroceptive threats (**a** study 1, compared to US_{SOM} ; **b** study 2, compared to US_{AUD} ; all $P_{FWE} < 0.05$). Moreover, interoceptive and exteroceptive threats induced shared neural activation in different brain regions across studies (full results in Supplementary Table S3). Neural activations in regions of interest were superimposed on a structural T1-image and thresholded at $P < 0.001$ uncorrected for visualisation purposes; colour bars indicate t -scores. For details, see Table 1. For whole-brain results on differential activation, see Supplementary Fig. S4. aINS anterior insula, AUD auditory, dACC dorsal anterior cingulate cortex, FWE family-wise error, SOM somatic, US unconditioned stimuli, VISC visceral.

criteria, in both samples mean gastrointestinal symptom scores (study 1: 3.73 ± 0.4 ; study 2: 1.65 ± 0.4), and HADS scores for anxiety (study 1: 4.24 ± 0.4 ; study 2: 3.00 ± 0.4) and depression (study 1: 2.61 ± 0.3 ; study 2: 1.04 ± 0.3) were low. Chronic stress scores were well-within the normal range (study 1: 36.93 ± 1.5 ; study 2: 33.57 ± 1.5). Pain thresholds for visceral pain (study 1: 40.0 ± 1.7 mmHg; study 2: 40.2 ± 1.6 mmHg) and for thermal heat pain (study 1: 44.9 ± 0.5 °C) were well-within the range of findings in our previous work^{4,21,46}. Unpleasantness thresholds for auditory stimuli (93.96 ± 1.52 dB SPL (range: 75–108 dB SPL) were within expected ranges and comparable to previous studies from our own group^{47,48}.

Unconditioned stimuli: acquisition phase. Negative valence was consistently greater for interoceptive compared to exteroceptive threats implemented as US in both studies, despite careful matching to US pain intensity (study 1) and US unpleasantness (study 2) prior to acquisition. This was indicated by significantly greater post-acquisition US_{VISC} unpleasantness ratings compared to US_{SOM} in study 1 (US_{VISC} : 66.14 ± 3.8 mm, US_{SOM} : 24.62 ± 5.1 mm; $t(41) = 6.62$; $P < 0.001$; $d = 1.43$), as well as to US_{AUD} in

study 2 (US_{VISC} : 79.30 ± 3.2 mm, US_{AUD} : 62.78 ± 5.3 mm; $t(22) = 3.42$; $P = 0.005$; $d = 0.75$). Note that in study 1, post-acquisition US_{VISC} were also perceived as more intense (US_{VISC} : 77.3 ± 1.5 mm, US_{SOM} : 66.2 ± 2.4 mm; $t(41) = 4.23$; $P < 0.001$; $d = 0.83$), which was highly intercorrelated with US unpleasantness within both modalities (US_{VISC} : $r = 0.71$; $P \leq 0.001$; US_{SOM} : $r = 0.65$; $P \leq 0.001$). At the neural level, US_{VISC} induced enhanced activation when compared to US_{SOM} (Fig. 1a) as well as compared to US_{AUD} (Fig. 1b), involving aINS and dACC in both studies, and additionally the amygdala in study 1 (Table 1). Further, interoceptive US_{VISC} consistently induced lower neural activation compared to both exteroceptive US_{SOM} and US_{AUD} within pINS (Table 1). Of note, these differences between US modalities remained largely unchanged when considering post-acquisition differences in US unpleasantness (study 1, 2) or US intensity ratings (study 1 only) as covariates of no interest (Supplementary Tables S1 and S2). Moreover, conjunction analyses (against global null) revealed shared neural activation in study 1 ($US_{VISC} \cap US_{SOM}$) in the aINS, whereas shared activation was detectable in study 2 only in uncorrected whole-brain analyses but not in FWE-corrected ROI analyses ($US_{VISC} \cap US_{AUD}$, see Supplementary Table S3).

Table 1 Differences in neural activation induced by interoceptive versus exteroceptive threats as unconditioned stimuli (US) during acquisition.

Contrast	Region	MNI-Coordinates				t-value	P
		H	x	y	z		
Study 1 (N = 42)							
US _{VISC} > US _{SOM} ROI analyses	aINS	L	-32	6	12	6.72	<0.001
		R	34	14	8	5.46	<0.001
Whole-brain analyses	dACC	L	-8	12	38	4.65	0.001
	Amygdala	L	-24	-4	-12	3.42	0.019
	PCC	R	4	-40	12	3.86	<0.001
	Inferior frontal gyrus, opercular (dIPFC)	R	38	8	32	3.84	<0.001
	Supramarginal gyrus (S2)	L	-62	-26	28	4.02	<0.001
		R	66	-28	28	3.63	<0.001
	Vermis	-	0	-54	-36	5.67	<0.001
US _{VISC} < US _{SOM} ROI analyses	Cerebellum	R	32	-52	-28	4.01	<0.001
Whole-brain analyses	pINS	R	38	-12	18	6.20	<0.001
	Superior frontal gyrus, medial (dmPFC)	L	-10	48	34	3.72	<0.001
Whole-brain analyses	PHIP	R	20	-6	-22	4.10	<0.001
	Postcentral gyrus (S1)	L	-54	-12	48	4.08	<0.001
		R	44	-22	54	3.47	0.001
	Postcentral gyrus (S2)	L	-62	-6	30	3.45	<0.001
		R	64	-4	24	5.41	<0.001
	SMA	R	8	-20	68	5.40	<0.001
	Fusiform gyrus	R	36	-44	-10	4.41	<0.001
	Rolandic operculum	L	-44	-14	18	4.18	<0.001
	Middle temporal gyrus	L	-62	-8	-14	5.07	<0.001
		R	66	-10	-16	3.42	0.001
Study 2 (N = 23)							
US _{VISC} > US _{AUD} ROI analyses	aINS	L	-44	2	8	4.78	0.004
		R	36	8	12	5.71	0.001
Whole-brain analyses	dACC	R	2	8	40	5.16	0.002
	SMA	R	6	4	46	6.25	<0.001
	Rolandic operculum	L	-46	2	8	5.36	<0.001
US _{VISC} < US _{AUD} ROI analyses	pINS	L	-38	-12	18	5.22	0.004
Whole-brain analyses		R	38	-10	16	5.62	0.002
	Middle temporal gyrus	R	56	-36	6	9.87	<0.001
	Superior temporal gyrus	L	-54	-36	10	7.60	<0.001
	Rolandic operculum	R	44	-10	22	6.50	<0.001
	Lingual area	L	-16	-48	-8	4.31	<0.001
		R	20	-52	-10	5.07	<0.001
	Angular gyrus	L	-48	-54	24	4.57	<0.001
	Precuneus	R	6	-52	44	4.14	<0.001
	Middle occipital gyrus	L	-44	-76	8	4.10	<0.001

Differential neural activation induced by interoceptive threat (US_{VISC}) compared to exteroceptive threat (study 1, US_{SOM}; study 2, US_{AUD}) implemented as unconditioned stimuli (US) during acquisition. Results of second-level paired t-tests are presented. Peak voxel indicate results of ROI analyses (cluster size $k_E \geq 3$; all $P_{FWE} < 0.05$) and whole-brain analyses (in *italic font*; cluster size $k_E \geq 10$; all $P_{uncorrected} < 0.001$). Exact unilateral P-values are provided. For a visualisation, see Fig. 1 and Supplementary Fig. S4. For analysis of shared responses, see Supplementary Table S3. For analyses controlling for US ratings, see Supplementary Tables S1 and S2.

aINS anterior insula, dACC dorsal anterior cingulate cortex, FWE family-wise error, H hemisphere, MNI Montreal Neurological Institute, PCC posterior cingulate gyrus, pINS posterior insula, PHIP parahippocampus, S1 primary somatosensory cortex, S2 secondary somatosensory cortex, SMA supplementary motor area, US unconditioned stimuli.

Conditioned stimuli: acquisition phase. Repeated CS⁺-US pairings during acquisition resulted in the conditioned negative valence of all threat predictors in both studies, as evidenced by significant rmANOVA time effects (Supplementary Tables S4 and S5). Interestingly, this increase in negative valence was consistently enhanced for interoceptive (ΔCS^+_{VISC}) versus exteroceptive threat predictors (ΔCS^+_{SOM} and ΔCS^+_{AUD} , respectively), despite comparable contingency awareness between modalities (Table 2). This indication of enhanced conditioned behavioural responses to interoceptive threat predictors was supported by significant time \times modality rmANOVA interaction effects in both studies (study 1: $F(1,41)=16.71$; $P < 0.001$; $\eta_p^2 = 0.29$, Supplementary Table S4; study 2: $F(1,22) = 8.19$; $P = 0.009$; $\eta_p^2 = 0.27$, Supplementary

Table S5), and between-modality comparisons revealing significantly enhanced conditioned negative valence post-acquisition for ΔCS^+_{VISC} (versus ΔCS^+_{SOM} : $t(41) = 4.19$; $P < 0.001$; $d = 0.59$; Fig. 2a and Supplementary Table S6; versus ΔCS^+_{AUD} : $t(22) = 3.28$; $P = 0.007$; $d = 0.37$; Fig. 2b and Supplementary Table S7; as well as by enhanced SCR for CS⁺_{VISC} in a subset of participants in study 2, see Supplementary Fig. S1). At the neural level, threat predictors induced shared differential activation in highly overlapping brain regions across studies, including aINS, hippocampus, MCC, dACC as revealed by conjunction analysis (in study 1: $\Delta CS^+_{VISC} \cap \Delta CS^+_{SOM}$; in study 2: $\Delta CS^+_{VISC} \cap \Delta CS^+_{AUD}$; full results in Supplementary Table S8). Interestingly, differential neural responses in pINS and MCC were enhanced for ΔCS^+_{VISC}

Table 2 Contingency awareness across learning phases.

Study 1 (N = 42)	CS ⁺ _{VISC}	CS ⁺ _{SOM}	CS ⁻	P*
ACQ	77.3 ± 3.0	70.2 ± 3.8	32.7 ± 4.4	0.052
EXT	24.2 ± 5.0	25.0 ± 5.1	10.5 ± 2.7	0.767
RST-TEST				
US _{VISC} -subgroup (N = 22)	16.6 ± 6.8	10.0 ± 5.2	8.0 ± 3.5	0.680
US _{SOM} -subgroup (N = 20)	15.9 ± 6.1	17.3 ± 6.7	16.3 ± 5.3	0.875

Study 2 (N = 23)	CS ⁺ _{VISC}	CS ⁺ _{AUD}	CS ⁻	P*
ACQ	80.9 ± 6.0	85.1 ± 4.3	15.5 ± 4.3	0.552
EXT	0.0 ± 0.0	2.7 ± 2.3	1.6 ± 1.6	0.237
RST-TEST	2.6 ± 1.8	1.7 ± 1.4	1.3 ± 1.1	0.504

Contingency awareness for threat-predictive conditioned stimuli (CS⁺) and safety cues (CS⁻) assessed with visual analogue scales (0–100 VAS, mm, representing % probability that a CS is followed by the specific US) after acquisition (ACQ), extinction (EXT), and reinstatement-test (RST-TEST) in study 1 (CS⁺_{VISC}, CS⁺_{SOM}) and study 2 (CS⁺_{VISC}, CS⁺_{AUD}). Note that actual CS-US contingencies for all CS⁺ during ACQ were 80% in study 1 and 83% in study 2; all CS were presented without US in EXT and RST-TEST phases. CS⁻ were never paired with US in any phase. Data are given as mean ± standard error of the mean.

AUD auditory, CS conditioned stimuli, SOM somatic, US unconditioned stimuli, VISC visceral.

*Exact P-values for between-modality paired t-tests assessing differences between CS⁺ within each study, shown uncorrected for multiple testing.

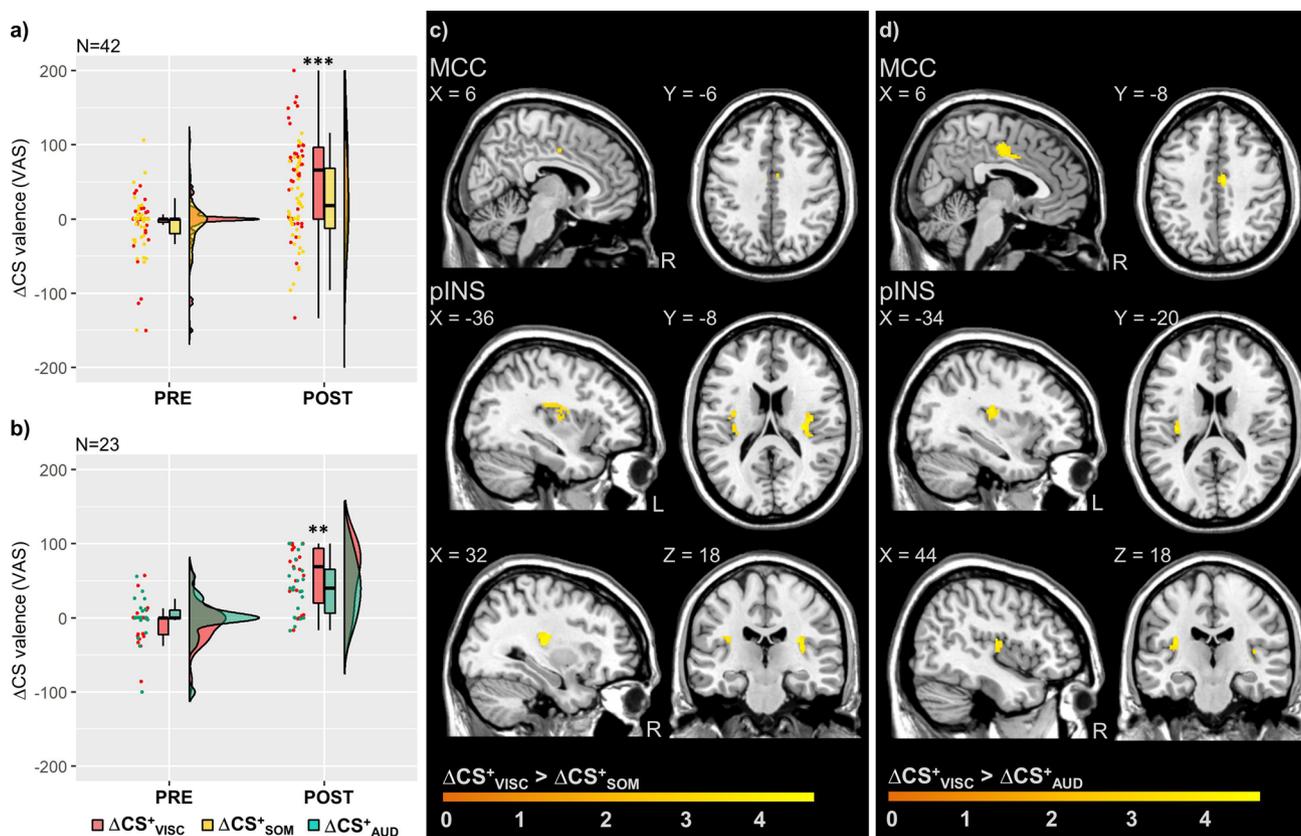


Fig. 2 Behavioural and neural responses to cues (CS) predicting interoceptive versus exteroceptive threats during acquisition in studies 1 and 2. Conditioned predictors of interoceptive threat (ΔCS^+_{VISC}) acquired significantly greater negative valence than conditioned predictors of exteroceptive threats after acquisition **a** study 1, compared to ΔCS^+_{SOM} ; $***P < 0.001$; **b** study 2, compared to ΔCS^+_{AUD} ; $**P < 0.01$; results of Bonferroni-corrected paired t-tests between modalities, full details in Supplementary Tables S4–S7. Individual delta (Δ) scores were computed for differential CS valence of each CS⁺ relative to the CS⁻. Data are presented as individual data points, boxplots, and densities (raincloud plots¹¹⁷). At the neural level, interoceptive threat predictors (ΔCS^+_{VISC}) induced enhanced differential neural responses in MCC and pINS compared to both exteroceptive threat predictors **c** study 1, $\Delta CS^+_{VISC} > \Delta CS^+_{SOM}$; **d** study 2, $\Delta CS^+_{VISC} > \Delta CS^+_{AUD}$; all $P_{FWE} < 0.05$; details in Table 3, but threat predictors also induced shared differential activation in overlapping brain regions across studies (full results in Supplementary Table S8). For whole-brain results, see Supplementary Fig. S5. Neural activations in regions of interest were superimposed on a structural T1-image and thresholded at $P < 0.001$ uncorrected for visualisation purposes; colour bars indicate t-scores. AUD auditory, CS conditioned stimuli, FWE family-wise error, MCC midcingulate cortex, pINS posterior insula, SOM somatic, VAS visual analogue scale, VISC visceral.

Table 3 Differences in neural activation induced by cues (CS) predicting interoceptive versus exteroceptive threats during acquisition.

Contrast	Region	MNI-coordinates				t-value	P
		H	x	y	z		
Study 1 (N = 42)							
$\Delta CS^+_{VISC} > \Delta CS^+_{SOM}$ ROI analyses	pINS	L	-36	-8	18	4.31	0.008
		R	32	-22	12	4.89	0.002
Whole-brain analyses	MCC	R	6	-6	40	3.56	0.023
	Superior frontal gyrus, medial (vmPFC)	R	12	60	0	3.62	<0.001
	Precentral gyrus (S1)	R	14	-30	74	5.41	<0.001
	Rolandic operculum	L	-42	-30	16	3.49	0.001
	Thalamus	L	-16	-26	0	3.79	<0.001
		R	14	-22	0	4.14	<0.001
	Putamen	R	28	-8	14	5.25	<0.001
	Inferior temporal gyrus	L	-58	-50	-14	3.66	<0.001
	Middle occipital lobe	L	-32	-80	40	4.17	<0.001
	Vermis	R	4	-44	-16	4.12	<0.001
	Cerebellum	L	-36	-74	-34	4.29	<0.001
	R	36	-78	-26	4.84	<0.001	
$\Delta CS^+_{VISC} < \Delta CS^+_{SOM}$	-	-	-	-	-	-	
Study 2 (N = 23)							
$\Delta CS^+_{VISC} > \Delta CS^+_{AUD}$ ROI analyses	pINS	L	-34	-20	18	4.86	0.007
		R	44	-14	14	4.39	0.017
		MCC	R	6	-10	42	4.77
Whole-brain analyses	dACC	R	4	6	36	3.80	0.026
	Superior temporal gyrus	L	-42	-28	8	6.05	<0.001
		R	50	-22	8	6.72	<0.001
$\Delta CS^+_{VISC} < \Delta CS^+_{AUD}$	-	-	-	-	-	-	

Differential neural activation induced by predictors of interoceptive (CS^+_{VISC}) versus exteroceptive threat (study 1, CS^+_{SOM} ; study 2, CS^+_{AUD}) relative to safety-predictive CS^- , during acquisition. Results of second-level paired *t*-tests (study 1: $\{CS^+_{VISC} < CS^-\} > \{CS^+_{SOM} < CS^-\}$; study 2: $\{CS^+_{VISC} < CS^-\} > \{CS^+_{AUD} < CS^-\}$; and vice versa) are presented. Peak voxel indicates results of ROI analyses (cluster size $k_E \geq 3$; all $P_{FWE} < 0.05$) and whole-brain analyses (in *italic font*; cluster size $k_E \geq 10$; all $P_{uncorrected} < 0.001$). Exact unilateral *P*-values are provided. For a visualisation, see Fig. 2 and Supplementary Fig. S5. For analysis of shared responses, see Supplementary Table S8. For analyses controlling for US ratings, see Supplementary Tables S9 and S10. AUD auditory, CS conditioned stimuli, dACC dorsal anterior cingulate cortex, FWE family-wise error, H hemisphere, MCC midcingulate cortex, MNI Montreal Neurological Institute, pINS posterior insula, S1 primary somatosensory cortex, SOM somatic, VISC visceral, vmPFC ventromedial prefrontal cortex.

compared to both ΔCS^+_{SOM} (Fig. 2c), as well as compared to ΔCS^+_{AUD} (Fig. 2d), a finding that was strikingly similar between both studies ($P_{FWE} < 0.05$, Table 3). Of note, these findings were not appreciably altered when considering post-acquisition differences in US ratings as covariates of no interest (Supplementary Tables S9 and S10). Moreover, exploratory correlational analyses revealed significant correlations between CS and US valence, as well as between differential CS- and US-induced activations in peak activations in the insula and cingulate cortices (Supplementary analyses).

Conditioned stimuli: extinction phase. Repeated presentations of all CS without any US-induced extinction of conditioned negative valence of threat predictors (for significant rmANOVA time effects observed in both studies, see Supplementary Tables S4 and S5). However, this was shaped by CS type, as supported by significant time \times modality interaction effects (study 1: $F(1,41) = 5.51$; $P = 0.024$; $\eta_p^2 = 0.12$; study 2: $F(1,22) = 6.36$; $P = 0.019$; $\eta_p^2 = 0.22$). In study 1 with an immediate extinction and a small number of extinction trials, post-extinction negative valence remained significantly greater for interoceptive compared to exteroceptive threat predictors (ΔCS^+_{VISC} : 35.38 ± 7.2 mm, ΔCS^+_{SOM} : 9.38 ± 7.0 mm; $t(41) = 3.42$; $P = 0.003$; $d = 0.57$). Contingency ratings indicated no differences between modalities, but an overestimation of true contingencies (CS^+_{VISC} : $24.2 \pm 5.0\%$, CS^+_{SOM} : $25.0 \pm 5.1\%$, with de facto 0% during extinction; Table 2). On the other hand, in study 2 with extinction accomplished 24 h after acquisition and a larger number of

extinction trials, no mid- or post-extinction difference in negative valence between interoceptive and exteroceptive cues was detectable (ΔCS^+_{VISC} : 5.00 ± 4.0 mm, ΔCS^+_{AUD} : 3.00 ± 2.8 mm; $P > 0.99$; $d = 0.12$; for mid-extinction results, see Supplementary Table S7), and contingency ratings were widely accurate, with no differences between modalities (Table 2). At the neural level, we observed shared differential neural activation induced by interoceptive and exteroceptive threat predictors (i.e. $\Delta CS^+_{VISC} \cap \Delta CS^+_{SOM}$ and $\Delta CS^+_{VISC} \cap \Delta CS^+_{AUD}$, respectively; full results in Supplementary Table S11) within several ROIs, including hippocampus in both studies, as well as vmPFC, pINS, and amygdala (study 1, Supplementary Fig. S2a) and aINS, and dACC (study 2, Supplementary Fig. S2b). No differences between modalities in differential CS-induced neural activation were found in any ROI in either study. However, whole-brain analyses revealed differences in the temporal gyrus (Supplementary Table S12).

Conditioned stimuli: reinstatement-test phase. In the reinstatement-test phase, unpaired CS presentations were implemented immediately after single threat reinstatement in study 1 [i.e. unexpected exposure to only US_{VISC} in one subgroup ($N = 22$), and to only US_{SOM} in another subgroup ($N = 20$)], or after multiple threat reinstatement (i.e. unexpected exposure to both US_{VISC} and US_{SOM} in all participants) in study 2. After single threat reinstatement (US_{VISC} ($N = 22$); US_{SOM} ($N = 20$)) in study 1, a significant time \times modality \times group interaction was observed for negative CS valence ($F(1,1,40) = 5.56$; $P = 0.023$; $\eta^2 = 0.12$, see Supplementary Table S4). After multiple threat reinstatement with

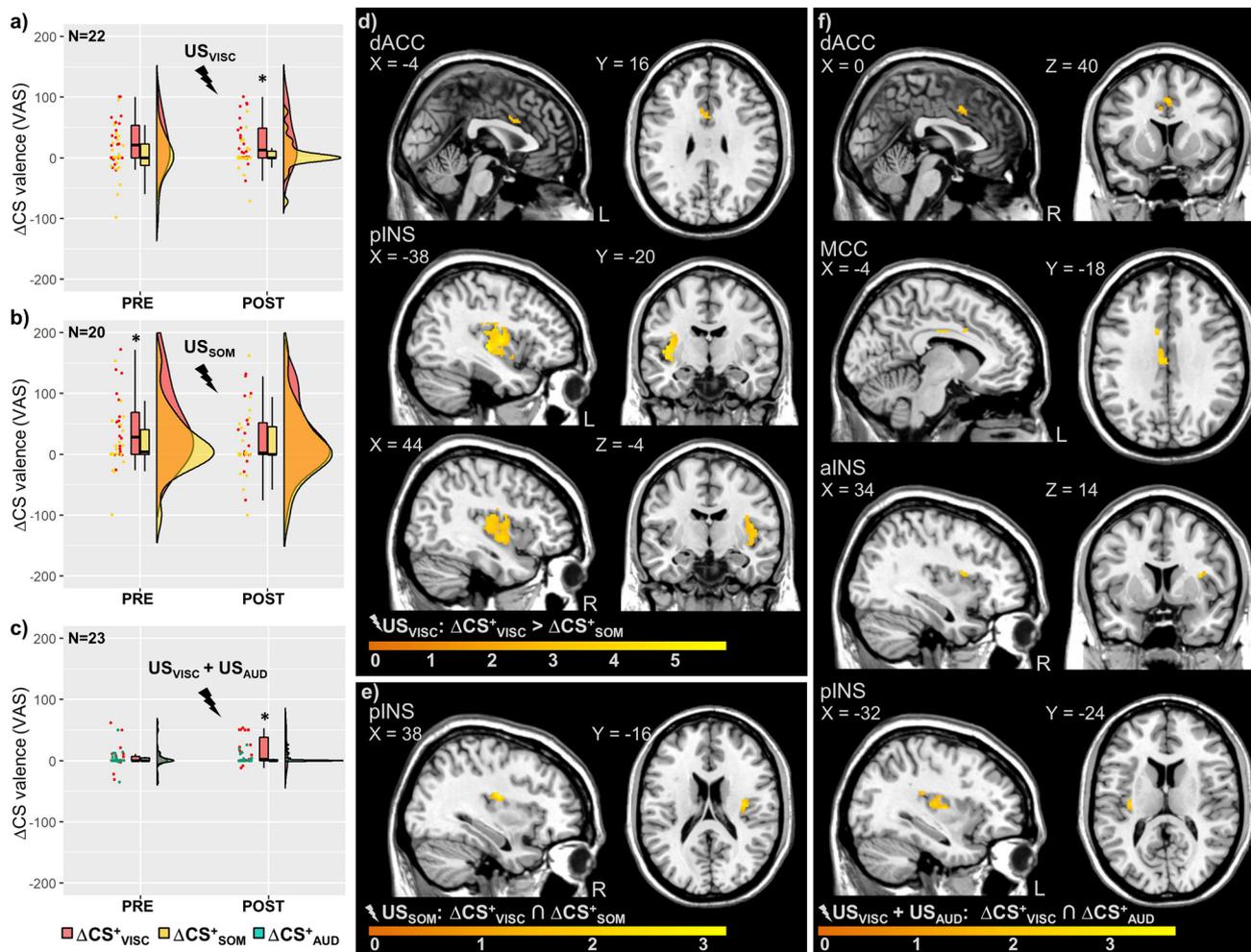


Fig. 3 Behavioural and neural responses to cues (CS) predicting interoceptive versus exteroceptive threats during reinstatement-test in studies 1 and 2. After reinstatement with unexpected US, negative valence was greater for interoceptive compared to exteroceptive threat predictors in reinstatement groups involving US_{VISC} **a** study 1, US_{VISC} -subgroup, ΔCS^+_{VISC} versus ΔCS^+_{SOM} ; **c** study 2, ΔCS^+_{VISC} versus ΔCS^+_{AUD} ; *both $P < 0.05$, results of Bonferroni-corrected paired t -tests between modalities, full results in Supplementary Tables S4–S7, whereas no difference was observed in the US_{SOM} -subgroup **b** study 1, ΔCS^+_{VISC} versus ΔCS^+_{SOM} . Individual delta (Δ) scores were computed for differential CS valence of each CS⁺ relative to the CS⁻. Data are presented as individual data points, boxplots, and densities (raincloud plots¹¹⁷). At the neural level, differential activation induced by interoceptive cues was enhanced in the US_{VISC} -subgroup within dACC and pINS during reinstatement-test **d** study 1, ΔCS^+_{VISC} compared to ΔCS^+_{SOM} , all $P_{FWE} < 0.05$. While no differential activation was observed for ΔCS^+_{VISC} compared to ΔCS^+_{SOM} in the US_{SOM} -subgroup in study 1 or compared to ΔCS^+_{AUD} in study 2, shared differential neural activation was induced by both threat predictors in regions of interest, such as in the insula and cingulate cortex **e** $\Delta CS^+_{VISC} \cap \Delta CS^+_{SOM}$ in study 1, US_{SOM} -subgroup; **f** $\Delta CS^+_{VISC} \cap \Delta CS^+_{AUD}$ in study 2. For full results, see Tables 4–6. For whole-brain results on differential activation, see Supplementary Fig. S7. Neural activations in regions of interest were superimposed on a structural T1-image and thresholded at $P < 0.01$ uncorrected for visualisation purposes; colour bars indicate t -scores. aINS anterior insula, AUD auditory, CS conditioned stimuli, dACC dorsal anterior cingulate cortex, FWE family-wise error, MCC midcingulate cortex, pINS posterior insula, SOM, somatic, US unconditioned stimuli, VAS visual analogue scale, VISC visceral.

both US_{VISC} and US_{SOM} in study 2, the interaction effect was not significant ($P = 0.055$; $\eta_p^2 = 0.16$, full results in Supplementary Table S5). Planned between-modality comparisons for the POST RST-TEST time point (Supplementary Tables S6 and S7) revealed greater differential negative valence of interoceptive compared to exteroceptive threat predictors in those groups involving US_{VISC} exposure during reinstatement (study 1, US_{VISC} -subgroup, ΔCS^+_{VISC} versus ΔCS^+_{SOM} , $t(21) = 2.73$; $P = 0.025$; $d = 0.58$; Fig. 3a; study 2, ΔCS^+_{VISC} versus ΔCS^+_{AUD} , $t(22) = 2.49$; $P = 0.042$; $d = 0.73$; Fig. 3c), whereas no difference was observed in the US_{SOM} -subgroup (study 1, ΔCS^+_{VISC} versus ΔCS^+_{SOM} , $P = 0.355$; $d = 0.21$; Fig. 3b). Exploratory within-group comparisons (PRE-POST) for study 1 showed no significant changes for ΔCS^+_{VISC} and ΔCS^+_{SOM} (both $P > 0.05$) in the US_{VISC} -subgroup, whereas in the US_{SOM} -subgroup ΔCS^+_{VISC} significantly decreased ($P = 0.005$ uncorrected, Supplementary

Table S6). The same comparisons for study 2 (PRE-POST) revealed a significant increase for ΔCS^+_{VISC} ($P = 0.038$ uncorrected), and no change for ΔCS^+_{AUD} ($P > 0.05$) (Supplementary Table S7).

At the neural level, threat predictors induced shared differential neural activation within hippocampus and vmPFC in the US_{VISC} -subgroup (Table 4), and in the pINS in the US_{SOM} -subgroup (study 1, $\Delta CS^+_{VISC} \cap \Delta CS^+_{SOM}$, Fig. 3e, Table 5). After multiple threat reinstatement in study 2, threat predictors induced shared activation in the hippocampus, aINS, pINS, dACC, and MCC ($\Delta CS^+_{VISC} \cap \Delta CS^+_{SOM}$, Fig. 3f, Table 6). Differential activation induced by interoceptive versus exteroceptive threat predictors was enhanced within pINS (L: $x = -38$, $y = -20$, $z = 0$, $t(21) = 5.23$, $P_{FWE} = 0.004$, R: $x = 44$, $y = -4$, $z = -4$, $t(21) = 4.91$, $P_{FWE} = 0.007$) and dACC (L: $x = -4$, $y = 16$, $z = 22$, $t(21) = 4.55$, $P_{FWE} = 0.023$) only in the US_{VISC} -subgroup (study 1, ΔCS^+_{VISC} compared

Table 4 Neural activation induced by cues (CS) predicting interoceptive versus exteroceptive threats during reinstatement-test after visceral single threat reinstatement in study 1.

Contrast	Region	MNI-coordinates				t-value	P
		H	x	y	z		
Study 1: US _{VISC} -subgroup ('single threat reinstatement'; N = 22)							
$\Delta CS^+_{VISC} \cap \Delta CS^+_{SOM}$ ROI analyses ¹	HIP	L	-16	-40	6	2.78	0.009
		R	40	-26	-12	2.96	0.004
Whole-brain analyses ¹	Superior frontal gyrus, orbital (vmPFC)	L	-16	56	-12	3.79	<0.001
	Inferior temporal gyrus	L	-48	-24	-20	2.48	<0.001
	Middle occipital lobe	L	-22	-82	14	2.52	<0.001
	Cerebellum	L	-28	-78	-34	3.67	<0.001
		R ^a	6	-72	-34	4.61	<0.001
ROI analyses ²	vmPFC	L	-16	56	-12	3.79	<0.001
Whole-brain analyses ²	PCC	L	-16	-44	8	3.98	<0.001
	Superior frontal gyrus (dmPFC)	R	20	6	70	3.00	<0.001
	Middle frontal gyrus, orbital (vlPFC)	L ^a	-36	46	-4	4.56	<0.001
		R	38	46	-2	2.93	<0.001
	Inferior frontal gyrus, triangular (vlPFC)	L	-38	20	6	2.53	<0.001
	Superior parietal gyrus	R	34	-54	58	2.86	<0.001
	Inferior parietal gyrus	L	-38	-62	50	2.40	<0.001
$\Delta CS^+_{VISC} > \Delta CS^+_{SOM}$ ROI analyses	pINS	L	-38	-20	0	5.23	0.004
		R	44	-4	-4	4.91	0.007
Whole-brain analyses	dACC	L	-4	16	22	4.55	0.023
	MCC	R	12	-10	50	5.18	<0.001
	PHIP	R	16	-38	-10	4.10	<0.001
	Thalamus	R	20	-24	-2	3.97	<0.001
	Lingual	L	-14	-42	-10	5.32	<0.001
		R	14	-52	6	4.96	<0.001
	Heschl gyrus	L	-52	-14	8	5.44	<0.001
	Precentral gyrus	L	-36	-28	66	4.23	<0.001
		R	52	4	18	4.42	<0.001
	Postcentral gyrus	L	-50	-20	54	4.80	<0.001
		R	60	-16	30	4.16	<0.001
	Rolandic operculum	R	46	-10	14	4.02	<0.001
	Precuneus	L	-16	-58	62	5.03	<0.001
		R	12	-62	68	4.08	<0.001
	Cuneus	R	22	-66	30	3.76	0.001
	Superior parietal gyrus	R	22	-64	64	3.77	0.001
	Calcarine fissure	L	-10	-70	20	4.04	<0.001
Cerebellum	R	10	-64	-48	4.17	<0.001	
$\Delta CS^+_{VISC} < \Delta CS^+_{SOM}$	-	-	-	-	-	-	

Shared and differential neural activation induced by predictors of interoceptive (CS^+_{VISC}) and exteroceptive threat (CS^+_{SOM}) relative to safety-predictive CS^- , during reinstatement-test (RST-TEST) after single threat reinstatement with US_{VISC}. For shared activation, results of conjunction analyses against global null are presented ($\{CS^+_{VISC} > CS^-\} \cap \{CS^+_{SOM} > CS^-\}$; $\{CS^+_{VISC} < CS^-\} \cap \{CS^+_{SOM} < CS^-\}$). For differential activation, results of second-level paired t-tests are presented ($\{CS^+_{VISC} < CS^-\} > \{CS^+_{SOM} < CS^-\}$; and vice versa). Peak voxel indicates results of ROI analyses (cluster size $k_E \geq 3$; all $P_{FWE} < 0.05$) and whole-brain analyses (in *italic font*; cluster size $k_E \geq 10$; all $P_{uncorrected} < 0.001$). Exact unilateral P-values are provided. For a visualisation, see Fig. 3 and Supplementary Fig. S7.

CS conditioned stimuli, FWE family-wise error, dACC dorsal anterior cingulate cortex, dmPFC dorsomedial prefrontal cortex, H hemisphere, HIP hippocampus, MCC midcingulate cortex, MNI Montreal Neurological Institute, PCC posterior cingulate cortex, PHIP parahippocampus, pINS posterior insula, ROI regions of interest, SOM somatic, VISC visceral, vlPFC ventrolateral prefrontal cortex, vmPFC ventromedial prefrontal cortex.

^aResults also significant against conjunction null).

to ΔCS^+_{SOM} , Fig. 3d, Table 4). In contrast, other groups revealed no differences between modalities in differential neural activation induced by interoceptive compared to exteroceptive threat predictors at all (Tables 5 and 6).

Discussion

Adaptive human behaviour in complex environments with multiple threats is guided by evolutionary-driven survival strategies that are preserved across species. In the face of imminent threat, learning from experience is particularly fundamental to the ability to identify and remember predictors of danger to facilitate avoidance or escape. As a translational model at the interface of psychology and the neurosciences, Pavlovian conditioning has proven valuable to elucidating behavioural and neural mechanisms underlying conditioned fear during the expectation of threat^{49–51}, with widely

appreciated clinical implications for anxiety and stress-related disorders¹. Herein, we broadened the scope to unravel learning and memory processes underlying negative expectations in the face of multiple threats from different sensory modalities, with a particular focus on the pain. Pain is a ubiquitous and highly salient threat and a crucial part of the organism's survival system that evokes strong adaptive responses, including cognitive and emotional processes orchestrated within the brain. These guide behaviour not only in response to actual pain experience⁵², but more importantly also during pain expectation^{8,53,54}. Interoceptive, visceral pain appears to be particularly threatening^{4,5}, engages partly distinct neural representations³⁷, and may have a specific functional role in shaping brain dynamics²⁷. Given the evolutionary significance of aversive signals originating from within our bodies, interoceptive conditioning could evoke greater and more persisting conditioned

Table 5 Neural activation induced by cues (CS) predicting interoceptive versus exteroceptive threats during reinstatement-test after somatic single threat reinstatement in study 1.

Contrast	Region	MNI-coordinates				t-value	P
		H	x	y	z		
Study 1: US _{SOM} -subgroup ('single threat reinstatement'; N = 20)							
$\Delta CS^+_{VISC} \cap \Delta CS^+_{SOM}$							
ROI analyses ¹	pINS	R	38	-16	22	3.16	0.004
Whole-brain analyses ¹	SMA	L	-14	-2	48	3.94	<0.001
	Superior frontal gyrus	R	22	-12	68	3.58	<0.001
	Superior frontal gyrus (dmPFC)	R	22	62	6	2.58	<0.001
	Middle frontal gyrus (vIPFC)	L	-32	26	50	2.56	<0.001
		R	44	48	16	2.57	<0.001
	Inferior frontal gyrus, triangular (vIPFC)	R	54	34	6	3.20	<0.001
	Inferior frontal gyrus, opercular (vIPFC)	R	48	10	12	2.84	<0.001
	Rectus (vmPFC)	L	-4	24	-22	3.19	<0.001
		R	12	40	-22	2.57	<0.001
	Precentral gyrus	L	-16	-18	70	2.82	<0.001
	Superior temporal gyrus	L	-62	-16	10	2.18	<0.001
		R	68	-32	16	2.86	<0.001
	Temporal pole: superior temporal gyrus	L	-40	20	-24	2.60	<0.001
	Middle temporal gyrus	R	68	-28	-2	3.03	<0.001
		R	36	18	-30	3.86	<0.001
	Rolandic operculum	L	-44	-16	18	2.55	<0.001
		R	42	-16	22	3.52	<0.001
	Inferior parietal gyrus	R	-48	-38	42	2.21	<0.001
	Superior occipital lobe	L	-18	-68	28	2.55	<0.001
Whole-brain analyses ²	Superior frontal gyrus, orbital (vmPFC)	R	16	58	-10	3.00	<0.001
$\Delta CS^+_{VISC} > \Delta CS^+_{SOM}$	-	-	-	-	-	-	-
$\Delta CS^+_{VISC} < \Delta CS^+_{SOM}$	-	-	-	-	-	-	-

Shared and differential neural activation induced by predictors of interoceptive (CS^+_{VISC}) and exteroceptive threat (CS^+_{SOM}) relative to safety-predictive CS^- , during reinstatement-test (RST-TEST) after single threat reinstatement with US_{SOM}. For shared activation, results of conjunction analyses against global null are presented ($\{CS^+_{VISC} > CS^-\} \cap \{CS^+_{SOM} > CS^-\}$; $\{CS^+_{VISC} < CS^-\} \cap \{CS^+_{SOM} < CS^-\}$). For differential activation, results of second-level paired t-tests are presented ($\{CS^+_{VISC} < CS^-\} > \{CS^+_{SOM} < CS^-\}$ and vice versa). Peak voxel indicates results of ROI analyses (cluster size $k_E \geq 3$; all $P_{FWE} < 0.05$) and whole-brain analyses (in *italic font*; cluster size $k_E \geq 10$; all $P_{uncorrected} < 0.001$). Exact unilateral P-values are provided. For a visualisation, see Fig. 3 and Supplementary Fig. S7. CS conditioned stimuli, FWE family-wise error, dmPFC dorsomedial prefrontal cortex, H hemisphere, MNI Montreal Neurological Institute, pINS posterior insula, ROI regions of interest, SMA supplementary motor area, SOM somatic, US unconditioned stimuli, VISC visceral, vIPFC ventrolateral prefrontal cortex, vmPFC ventromedial prefrontal cortex.

responses relevant to nocebo mechanisms underlying hypervigilance and hyperalgesia.

In two independent fMRI studies, we implemented specific, yet complementary differential conditioning paradigms to elucidate the acquisition and extinction of conditioned responses to predictors of interoceptive and exteroceptive threats. Experimental visceral pain as clinically-relevant interoceptive US was significantly more unpleasant when compared to the exteroceptive US from two sensory modalities, i.e. exteroceptive somatic pain in study 1 and aversive tone in study 2, despite careful matching to intensity and unpleasantness, respectively, supporting possible differences in habituation processes⁴. In addition to shared neural activation induced by US of different modalities, interoceptive US interestingly evoked greater neural activation within the anterior insula and dorsal anterior cingulate cortex as key regions of the salience network, with well-established roles in the central integration of interoceptive sensory signals with emotional and cognitive facets⁵⁵⁻⁵⁷. These findings, which appeared robust even when considering differences in US ratings as nuisance variables, reproduce and complement earlier efforts to elucidate the specificity of interoceptive visceral pain in shaping aversive anticipation and central pain processing, not only in direct comparison to an exteroceptive painful threat^{4,26,58}, but also to a non-nociceptive, yet a priori equally aversive auditory threat. In line with a notable recent publication detailing a multivariate brain measure, the Neurologic Pain Signature (NPS), for visceral and somatic stimulation across different independent datasets³⁷, our results from two independent studies support the unique salience of interoceptive pain as a US, above and beyond specific yet highly

intertwined perceptual characteristics of intensity and unpleasantness, and underscore the suitability of this experimental model to elucidating the role of threat modality in associative learning and extinction processes in a clinically-relevant context.

To assess whether US modality distinctly shapes the formation of learned negative expectations, we accomplished analyses of differential conditioned responses to modality-specific threat predictors (CS). Results for the acquisition phases of both studies showed enhanced differential behavioural and neural responses to interoceptive threat predictors, suggesting preferential learning for the visceral modality. This was supported at the behavioural level by greater increases in the negative valence of interoceptive versus exteroceptive predictive cues, which were observed despite comparable contingency awareness, and greater visceral cue-induced SCR responses suggested by exploratory analyses of a subset of data. Within the brain, we documented shared differential activation to all threat predictors compared to safety cues in highly overlapping brain regions across studies, in line with a recent meta-analysis documenting a consistent and robust pattern of an 'extended fear network' across diverse fear conditioning paradigms⁵⁰, as well as a meta-analysis supporting that pain-related and non-pain-related conditioned fear recruits overlapping but distinguishable neural networks²⁴. Importantly, we also consistently demonstrated differences between interoceptive versus exteroceptive threat predictors in both studies. Specifically, conditioned interoceptive threat predictors induced greater differential neural activation in the posterior insula and midcingulate cortex. A recent meta-analysis focusing on pain anticipation supported an interplay of insular and cingulate regions in the representation of the affective qualities of sensory

Table 6 Neural activation induced by cues (CS) predicting interoceptive versus exteroceptive threats during reinstatement-test after multiple threat reinstatement in study 2.

Contrast	Region	MNI-coordinates				t-value	P	
		H	x	y	z			
Study 2: RST with US _{VISC} and US _{SOM} ('multiple threat reinstatement'; N = 23)								
$\Delta CS^+_{VISC} \cap \Delta CS^+_{AUD}$ ROI analyses ¹	aINS	R	34	6	14	2.94	0.011	
	pINS	L	-32	-24	8	3.52	<0.001	
	dACC	L	0	10	40	2.47	0.009	
	MCC	L	-4	-18	32	2.70	0.032	
	HIP	L	-30	-32	-2	3.72	<0.001	
Whole-brain analyses ¹	Inferior frontal gyrus, triangular (dlPFC)	L	-44	12	24	3.03	<0.001	
	ACC	L	2	36	28	3.21	<0.001	
	PCC	R	12	-46	20	3.05	<0.001	
	Thalamus	L	-20	-28	10	5.20	<0.001	
		R	14	-24	14	4.62	<0.001	
	Middle temporal gyrus	L	-60	-48	4	3.05	<0.001	
	Precuneus	L	-2	-58	30	3.47	<0.001	
	Paracentral lobe	R	10	-34	50	2.13	<0.001	
	Gyrus fusiformis	L	-30	-66	-6	2.38	<0.001	
		R	38	-56	-14	2.63	<0.001	
	Vermis	R	4	-36	-16	2.33	<0.001	
	Cerebellum	R	26	-52	-22	2.95	<0.001	
	Whole-brain analyses ²	Middle temporal gyrus	R	52	-58	12	3.38	<0.001
		Supramarginal gyrus (S2)	R	48	-44	22	2.41	<0.001
		Calcarine fissure	R ^a	28	-58	12	2.61	<0.001
$\Delta CS^+_{VISC} > \Delta CS^+_{AUD}$	-	-	-	-	-	-		
$\Delta CS^+_{VISC} < \Delta CS^+_{AUD}$	-	-	-	-	-	-		

Shared and differential neural activation induced by predictors of interoceptive (CS^+_{VISC}) and exteroceptive threat (CS^+_{AUD}) relative to safety-predictive CS^- , during reinstatement-test (RST-TEST) after multiple threat reinstatement with US_{VISC} and US_{AUD}. For shared activation, results of conjunction analyses against global null are presented (¹ $\{CS^+_{VISC} > CS^-\} \cap \{CS^+_{AUD} > CS^-\}$; ² $\{CS^+_{VISC} < CS^-\} \cap \{CS^+_{AUD} < CS^-\}$). For differential activation, results of second-level paired t-tests are presented ($\{CS^+_{VISC} < CS^-\} > \{CS^+_{AUD} < CS^-\}$; and vice versa). Peak voxel indicates results of ROI analyses (cluster size $k_E \geq 3$; all $P_{FWE} < 0.05$) and whole-brain analyses (in *italic font*; cluster size $k_E \geq 10$; all $P_{uncorrected} < 0.001$). Exact unilateral P-values are provided. For a visualisation, see Fig. 3 and Supplementary Fig. S7.

aINS anterior insula, CS conditioned stimuli, FWE family-wise error, dACC dorsal anterior cingulate cortex, dlPFC dorsolateral prefrontal cortex, H hemisphere, HIP hippocampus, MCC midcingulate cortex, MNI Montreal Neurological Institute, pINS posterior insula, ROI regions of interest, S2 secondary somatosensory cortex, VISC visceral.

^aResults also significant against conjunction null.

events, particularly applying to interoceptive signals⁵⁹, in line with our own recent findings documenting the relevance of posterior insula in visceral compared to somatic pain expectation²⁶. Given the well-established role of the posterior insula in restoring and maintaining homeostasis in the face of imminent danger⁶⁰, its distinct involvement may serve adaptive modulatory functions during the expectation of interoceptive threat. Our findings extend knowledge from other brain imaging studies on the modality-specific aversive expectancy that have compared predictors of somatic pain with aversive pictures³¹ or disgusting odours³⁰, and complement our own data on nocebo effects and underlying mechanisms in visceral pain^{8,9,11,21}. These observations suggest a specific relevance of insular together with cingulate regions in the preferential acquisition of the presumably more salient interoceptive CS-US association. In keeping with the notion that expectations dynamically influence perception and learning^{61,62}, the reciprocal impact of interoceptive threats and their predictors is further substantiated by our exploratory correlational results. These not only indicate that affective qualities of interoceptive versus exteroceptive threats shape conditioned negative expectations, but also suggest a tight link between differential neural responses during the expectation and experience of aversive interoceptive signals. Together, our findings regarding the formation of negative interoceptive expectancies by aversive conditioning support our hypothesis that in the face of multiple danger signals indicating bodily harm, visceral pain evokes preferential interoceptive fear learning. These findings could be viewed as a modern replication of classical

interoceptive conditioning studies carried out by soviet psychologists³⁵, complemented herein by brain imaging techniques. They are in keeping with preparedness theory⁴⁰, and support its applicability to pain-related learning in a broader context of the affective neurosciences, with intriguing putative clinical relevance. The role of fear and hypervigilance is increasingly appreciated in the pathophysiology and treatment of multiple complex and overlapping clinical conditions, including anxiety and chronic pain⁵⁴. Modality-specific conditioning could therefore contribute to unravelling nocebo mechanisms relevant to vulnerability to chronicity and treatment failure, especially when nocebo effects persist rather than extinguish.

Persisting or resurging fear constitutes a core target of cognitive-behavioural treatment approaches like exposure therapy, which is essentially built on the successful and robust extinction of conditioned responses including learned fear. When the threat is no longer present, extinction of conditioned responses to former threat predictors is adaptive, allowing behavioural flexibility in rapidly changing, complex environments. At the same time, the initially acquired memory trace is preserved and can be dynamically reactivated⁶³, which can contribute to impaired extinction efficacy and to relapse in clinical contexts⁶⁴. This may be particularly relevant for highly salient and fear-evoking threats that are crucial to avoid, like interoceptive pain, as essentially already suggested by the Soviet pioneers of classical conditioning³⁵. To elucidate threat modality-specific extinction processes and their underlying neural mechanisms, we tested whether the visceral

CS–US association is more resistant to extinction and more susceptible to memory reactivation or ‘relapse’, induced by unexpected US exposure (i.e. reinstatement). In an effort to model different aspects of extinction learning, including the clinical reality of patients with waxing and waning symptoms, we herein implemented different experimental extinction and reinstatement protocols. Interestingly, when omission of the US occurred directly after acquisition on the same day in study 1, behavioural results indicated persisting conditioned fear in response to visceral but not somatic pain predictors, despite comparable contingency awareness. While these results may indicate a greater resistance to extinction specifically for the interoceptive CS–US association, a cautious interpretation is warranted given the small number of extinction trials and an overall overestimation of reported CS–US contingency awareness. In study 2, with an extinction phase accomplished on a subsequent study day and more extinction trials, conditioned behavioural responses were no longer evident to either predictive cue, not even in a supplementary analysis of a smaller number of extinction trials, and hence do not support a modality-specific resistance to extinction. However, given that our earlier conditioning work repeatedly documented rapid and full extinction in 1-day paradigms with visceral threats only^{20,65}, together the present findings could hence also indicate that full extinction of conditioned emotional responses to multiple threats requires more unreinforced trials, especially for threats of higher salience and immediate extinction learning. In light of increasing knowledge regarding the role of consolidation and reconsolidation in the context of conditioned fear^{66,67}, our divergent findings in studies 1 and 2 call for more mechanistic studies on the temporal dynamics and boundary conditions of pain-related extinction learning in multi-day and multi-threat paradigms, ideally including objective, physiological measures derived from electrodermal activity or pupillometry recordings. Regarding brain imaging results for the extinction phase, no differences were observed in neural responses to threat predictors of different modalities in either study. Instead, shared neural activation induced by both threat predictors during extinction was evident in both studies, involving key areas of the extinction network, particularly the hippocampus, supporting its general role in extinction learning⁶⁸ irrespective of threat modality, or of the number and timing of extinction trials.

Although extinction efficacy is clearly relevant to chronicity and treatment failure in patients^{69,70}, underlying mechanisms remain incompletely understood even in healthy individuals. Reinstatement constitutes a promising translational tool³⁸, which has not been applied in human brain imaging studies in the context of multiple threats. While we implemented different reinstatement procedures in the two studies, all involved unexpected US exposure followed by unpaired cue presentations. This allowed us to analyse differential conditioned responses during a reinstatement-test phase for different (former) threat predictors as an indicator of extinction efficacy. Our behavioural findings provide at least partial support for the notion that extinction efficacy may be reduced for the interoceptive CS–US association, i.e. that reinstatement with the visceral US had a greater impact on differential responses to CS. After reinstatement involving unexpected exposure to the visceral US, either as a ‘single threat’ (US_{VISC}-subgroup of study 1) or as a ‘multiple threat’ (all participants in study 2), we observed greater negative valence of former interoceptive versus exteroceptive threat predictors. On the other hand, reinstatement with the somatic US alone (US_{SOM}-subgroup of study 1) did not induce differences between CS modalities. Our hypothesis is most clearly supported by the results of study 2, in which reinstatement induced a selective resurgence of the interoceptive CS–US association, consistent with a return of interoceptive fear. Interpretation of findings in

study 1 is complicated by the fact that extinction was immediate and shorter, as explained above, and evidently did not lead to a complete resolution of conditioned responses. Herein, reinstatement can rather be conceptualised as a disruption of the ongoing extinction process, which would make a return (i.e. a *de novo* increase) of conditioned responses difficult to detect. While we did not plan for this, and results may be hampered by limited statistical power given relatively small sample sizes of reinstatement groups, applicability to real-life scenarios is intriguing: Unexpected and unsignaled threats like painful episodes can obviously occur at any time point during an ongoing extinction process. Understanding how easily such a process can be disrupted would hence inform our understanding of adaptive extinction learning with relevance to factors that may interfere with successful exposure-based treatment. Given these considerations, one interpretation of data from study 1 is that unexpected exposure to interoceptive threat disrupted or in fact halted the ongoing extinction of conditioned responses to interoceptive threat predictors, whereas the extinction process for the same predictors effectively continues after unexpected exposure to exteroceptive threat, as supported by within-modality comparisons. Interestingly, residual effects of cue aversiveness have been shown to predict a reinstatement effect in healthy volunteers⁷¹, consistent with evidence from a clinical setting⁷². Hence, residual fear responses after extinction may serve as a predictor of treatment outcome. In patients with persistent fear and/or pain, achieving a robust and sustained extinction of conditioned responses constitutes an important treatment goal, and maybe particularly challenging for interoceptive memory traces based on our data in healthy individuals. In other words, we may be primed to preferentially learn, store, and remember cues that signal internal harm.

Within the brain, after reinstatement with visceral US alone or in combination with the somatic US, shared differential neural responses to both threat-predictive cues were consistently observed in the hippocampus as a core region of the extinction network. This finding is well in line with earlier studies implementing only one threat modality^{73–75}, including visceral pain in healthy individuals⁷⁶ and in patients with chronic visceral pain^{77,78}. In addition to the hippocampus, posterior insula and cingulate regions were differentially activated following reinstatement with visceral threats alone, a neural activation pattern that closely resembled neural responses detected during the acquisition, yet not observed during extinction. These enhanced differential responses within the insula and cingulate cortex during reinstatement-test may reflect the reactivation of the excitatory memory trace, presumably triggering preparatory responses in expectation of the reoccurrence of threat. Notably, single threat visceral reinstatement (study 1), resulted in enhanced differential neural responses for interoceptive compared to the exteroceptive threat predictors, whereas multiple threat reinstatement (study 2) led to shared differential activation of these regions to both interoceptive and exteroceptive threat cues. The latter finding may be explained by generalisation effects induced by the unexpected exposure to multiple threats in close temporal proximity. This may promote a generalisation of threat value from the more salient interoceptive to the less salient exteroceptive threat, ultimately resulting in the reactivation of neural responses to all former threat predictors regardless of their salience. While this is speculative, the inability to adequately differentiate between stimuli of different threat value, particularly when confronted with recent adversity, has previously been discussed as one mechanism contributing to enhanced relapse risk in clinical populations undergoing extinction-based treatment⁷⁴. Together, these observations underline the critical importance of factors promoting either discrimination or generalisation of

conditioned responses in the prediction and prevention of fear relapse^{79,80}, extending the concept that as a result of conditioning, impaired discrimination of conditioned⁸¹ and unconditioned responses⁸² could play a role in the development of chronic pain.

Conditioned interoceptive fear should generally be considered adaptive, and an essential component of evolutionary-driven survival behaviour. However, when contextualised within a broader nocebo framework in which conditioned negative expectations drive maladaptive avoidance, hypervigilance and hyperalgesia^{8,83,84}, putative clinical implications and future directions are noteworthy. Our findings support that interoceptive threat predictors may more readily evoke conditioned fear, which could drive the transition from acute to chronic pain as well as symptom chronicity, especially in vulnerable individuals. Furthermore, the risk for impaired extinction efficacy and relapse phenomena may be more pronounced in the context of aversive interoceptive signals, especially in combination with stress^{83,85}, which demonstrably amplifies visceral nocebo effects⁸⁶, and may contribute to a negative recall bias about aversive visceral experiences⁸⁷. Given the evidence supporting altered extinction learning in patients with chronic pain, including IBS^{77,78}, translational research in clinical populations is urgently needed. While our results are limited by the lack of complete SCR data as a biological marker of learning as well as by a more definite exclusion of pain-modality-specific vascular artefacts induced by gasping or other respiratory or movement-related effects that could be more closely inspected whether pulse oximetry or respiration had been measured, they do provide a more refined understanding of conditioned nocebo effects in the context of clinically-relevant interoceptive and exteroceptive threats. Merging our clinically-driven perspective with the rapidly expanding general literature on interoception and predictive processing provides opportunities for the development or refinement of computational models based on the precision of interoceptive versus exteroceptive signals, advancing not only the definition of salience itself but also clarification its mechanistic basis^{88–93}. The translation of this knowledge may ultimately help understand and minimise negative expectancy effects in patients with chronic pain^{94,95}, especially in disorders of gut–brain interactions, adds a brain perspective to the eloquent claim that the gut is ‘smart’ due to its capability to learn and remember⁹⁶, and supports further efforts towards extinction-based treatment approaches for these highly prevalent conditions^{10,97}.

Methods

Participants. For the purposes of this report, we analysed unpublished data from healthy volunteers who were recruited to serve as controls in two conceptually connected fMRI conditioning studies conducted within a collaborative research unit (SFB 1280 ‘Extinction Learning’, funded by the German Research Foundation). We utilised data from healthy volunteers recruited as part of a patient study (study 1), and included data from the placebo arm of a pharmacological study (study 2; German Clinical Trials Register, registration ID: DRKS00016706). Recruitment and screening of all healthy volunteers in both studies followed highly-standardised and established procedures in our line of visceral pain research^{4,86}. An initial structured telephone screening was followed by a personal interview and a medical examination. Interview and examinations were accomplished in a medically-equipped room within a clinical research unit at the University Hospital Essen, Germany. Exclusion criteria common to both studies were <18 or >45 years of age, body mass index (BMI) <18 or >30, and any known medical condition or regular medication use (except thyroid medication and hormonal contraceptives). The usual exclusion criteria for magnetic resonance imaging (MRI) applied, and structural brain abnormalities were ruled out upon structural MRI. Perianal tissue damage (e.g. haemorrhoids, fissures), which may interfere with rectal balloon distensions were excluded by digital rectal examination. Pregnancy was excluded with a commercially available urinary pregnancy test (Biorepair GmbH, Sinsheim, Germany) on the day of the experiment. Prior participation in any previous or other ongoing studies involving pain-related conditioning was also exclusionary. Standardised questionnaires were used to screen for recent gastrointestinal complaints⁹⁸, symptoms of depression or anxiety (Hospital Anxiety and Depression Scale, HADS)⁹⁹, as well as to confirm right-handedness^{100,101}. As part of a comprehensive psychosocial questionnaire battery, chronic perceived stress was also assessed (Trier Inventory of

Chronic Stress, TICS)¹⁰². All the participants reported normal hearing and normal or corrected-to-normal vision. The work was conducted in accordance with the Declaration of Helsinki, and studies were approved by the ethics committee of the University Hospital Essen (protocol numbers 10–4493 and 16–7237), and followed the relevant ethical guidelines and regulations. All volunteers provided written informed consent and were paid for their participation.

Overview of study designs and procedures. To elucidate the formation and extinction of conditioned responses to threat-predictive cues (conditioned stimuli, CS) in the face of multiple different biologically-salient threats, we implemented two differential delay conditioning studies with visual CS⁺ predicting interoceptive threat (visceral pain: US_{VISC}) or exteroceptive threat (study 1: somatic thermal pain, US_{SOM}; study 2: aversive auditory stimulus, US_{AUD}, study 2), and unpaired CS⁻. All experimental procedures were conducted in the MRI-suite of the Institute of Diagnostic and Interventional Radiology and Neuroradiology at the University Hospital Essen, Germany. In both studies, perceptual thresholds for each US modality were initially assessed, and individual US stimulus intensities for implementation during conditioning were identified with a calibration and matching procedure (for details, see below). During conditioning, both studies implemented the same sequence of experimental learning phases, namely acquisition, extinction, and reinstatement-test phases (Fig. 4, details below). In study 1, all phases were accomplished consecutively on a single study day, whereas in study 2, extinction and reinstatement-test phases were implemented 24 h after acquisition. During all learning phases, fMRI was applied to assess shared and differential neural activation induced by US and CS. US- and CS-related behavioural measures were acquired for each phase using digitised visual analogue scales (VAS). Moreover, electrodermal activity was continuously recorded aiming for analysis of skin conductance responses (SCR) as a psychophysiological measure of learning using an MRI-compatible system (Biopac Systems, Inc., Goleta, CA, USA; MP100 in study 1, MP160 in study 2), but technical difficulties resulted in incomplete data. Results of exploratory analyses of skin conductance responses to CS for a subset of participants in studies 1 and 2 are provided as Supplementary Material (Supplementary Fig. S1). Note that participants were informed that the study goals were to investigate neural mechanisms underlying visceral pain-related learning and memory processes. Importantly, no detailed information was provided about experimental phases, or about the contingencies between CS and US.

Unconditioned stimuli (US). For interoceptive US (US_{VISC}), applied in both studies, pressure-controlled rectal distensions were carried out with a barostat system (modified ISOBAR 3 device, G & J Electronics, Toronto, ON, Canada). Graded distensions of the rectum with an inflatable balloon constitute a well-established experimental model to assess viscerosensation and visceral pain, especially in the context of IBS¹⁰³. The distension model allows the controlled and finely-tuned application of distensions inducing mild, intermediate or strong sensations of urgency, discomfort, and pain that closely resemble aversive visceral sensations experienced by patients, which are also commonly but less frequently experienced by healthy persons. Building on our long-standing experimental expertise with different sensory modalities^{4,47,48,86}, for the exteroceptive US, cutaneous thermal stimuli (US_{SOM}) were applied on the left ventral forearm with a thermode (PATHWAY model CHEPS; Medoc Ltd. Advanced Medical Systems, Ramat Yishai, Israel) in study 2, an aversive tone (US_{AUD}) with a saw-tooth waveform profile and a frequency of 1 kHz created using Audacity 1.3.10-beta (<http://www.audacity.sourceforge.net/>) was presented by an MRI-compatible sound system (Amplifier mkII+S/N 2016-2-2-03, MR confon GmbH, Magdeburg, Germany) bi-aurally through headphones. Individual perceptual thresholds for US are reported herein only for the purpose of descriptive characterisation of the study samples, as they primarily served as anchors for US calibration and matching.

In a continuation of our previous work on specificity to pain modality^{4,26,86}, US_{VISC} and US_{SOM} were matched to perceived pain intensity in study 1. In study 2, US_{VISC} were compared to a non-nociceptive, yet equally aversive US_{AUD} by matching to perceived unpleasantness. In both studies, visceral stimuli served as an anchor for calibration and matching of exteroceptive US, aiming to identify individual US stimulation intensities within a predefined perceptual range of 60–80 mm (assessed on 0–100 mm VAS, ends labelled not painful and extremely painful in study 1, and not unpleasant and very unpleasant in study 2). To this end, a distension pressure 5 mmHg below the individual rectal pain threshold was initially chosen and rated on VAS until pressure within the predefined range was identified. This was successfully accomplished in both studies (VAS ratings for US_{VISC}: 70.1 ± 0.9 mm in study 1; 65.8 ± 3.6 mm in study 2). For matching, visceral stimuli were presented with thermal (study 1) or auditory stimuli (study 2), respectively, and participants were prompted to compare the stimuli on a response device with Likert-type response options indicating more, less, or equally painful (study 1) or unpleasant (study 2) stimuli. If the rating showed a deviation, the intensity of exteroceptive stimuli was successively adjusted until ratings indicated equal perception at least twice consecutively. Of note, stimulus durations for interoceptive and exteroceptive US were adjusted for each individual, aiming at matched durations of ascending and plateau phases of US stimulation (20 s study 1; 14 s study 2). For additional details, see Supplementary Methods.

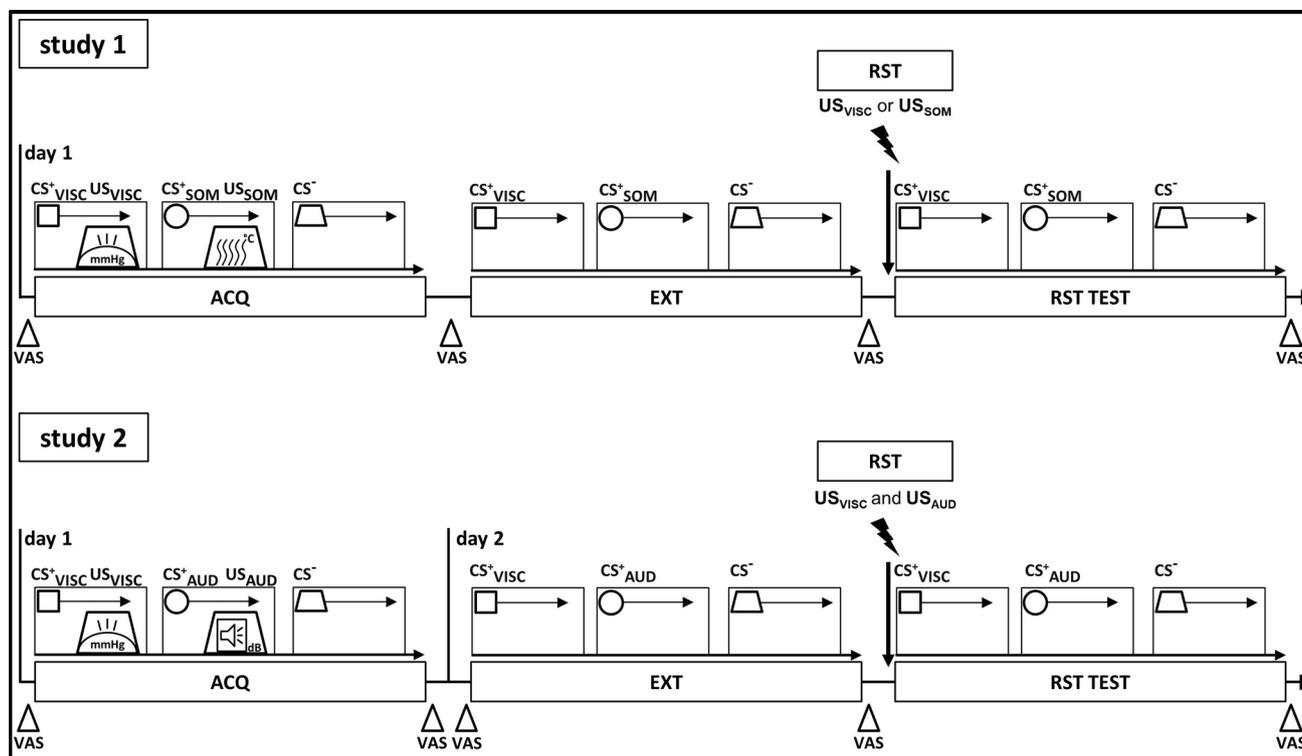


Fig. 4 Schematic overview of study designs. All participants in studies 1 and 2 underwent acquisition (ACQ), extinction (EXT), and reinstatement-test (RST-TEST) phases. As for unconditioned stimuli (US), visceral pain (US_{VISC}) and either equally painful somatic pain (study 1, US_{SOM}), or equally-unpleasant auditory stimuli (study 2, US_{AUD}) were implemented during acquisition (ACQ) and reinstatement (RST). As conditioned stimuli (CS), distinct visual geometrical symbols were paired with US (CS^+_{VISC} ; CS^+_{SOM} ; CS^+_{AUD}) or were presented without US (CS^-) during acquisition (differential delay conditioning). All CS were presented without US during EXT and RST-TEST. RST procedures involved unannounced US from one modality ('single threat reinstatement' in study 1: US_{VISC} in one subgroup; US_{SOM} in another subgroup) or from both modalities (multiple threat reinstatement in study 2: US_{VISC} and US_{AUD} in all participants). During all phases, functional magnetic resonance imaging (fMRI) was accomplished to assess shared and differential CS- and US-induced neural activation in regions of interest. Before and after each phase, behavioural measures were acquired with visual analogue scales (VAS).

Note that after matching, in study 1 a short adaptation phase was accomplished in the MR scanner to accommodate for possible habituation effects previously observed for thermal pain stimuli⁴. This involved a short series of unannounced US presentations (i.e. five visceral and five heat pain in pseudorandomized order), followed by another matching procedure when necessary. Supplementary control analysis of movement data (linear and degree movement) was accomplished for the unannounced US delivered in the habituation phase conducted prior to the acquisition, in order to explore possible pain-modality-specific movements (e.g. due to gasping) potentially confounding differential neural activation in subsequent experimental phases (Supplementary Fig. S3).

Based on these careful matching procedures, the following stimulus intensities for implementation during acquisition and reinstatement were identified: For US_{VISC} distension pressures, 34.7 ± 1.6 mmHg in study 1, 35.9 ± 1.8 mmHg in study 2; for US_{SOM} thermode temperature, 45.1 ± 0.3 °C; for US_{AUD} loudness, 94.74 ± 1.18 dB SPL (range: 89–108 dB SPL), all in line with our earlier results involving the application of the same pain^{4,86} or auditory stimuli^{47,48}.

Experimental phases. During acquisition, both studies involved three distinct conditioned stimuli (CS), which were contingently paired with the interoceptive US (CS^+_{VISC}) and one exteroceptive US (CS^+_{SOM} or CS^+_{AUD} , respectively), or remained unpaired (CS^-). Visual geometric symbols served as CS, and allocation of a specific CS symbol to a specific US (US_{VISC} , US_{SOM} , or US_{AUD}) or designation as CS^- was counterbalanced across participants. Within each study, participants were pseudo-randomly assigned to a different order sequence of CS-US pairings to avoid potential sequence effects. The programming of pairings aimed for an essentially pseudorandomized order, but avoided more than two successive pairings of one modality, and ensured that sequences alternately started with an interoceptive or exteroceptive CS-US pairing. All acquisition sequences were also balanced for the number of CS^+ and CS^- presentations. The number of CS presentations and reinforcement schedules were similar in the two studies (study 1: 10 presentations per CS, 8 CS-US pairings, 80% reinforcement; study 2: 12 presentations per CS, 10 CS-US pairings, 83% reinforcement). All CS^+ were presented 6–12 s before US, with CS and US co-terminating. Inter-stimulus intervals consisted of a black screen with a white frame (durations: 5–8 s).

During extinction and reinstatement-test phases, CS was presented without any US. Given earlier evidence of rapid extinction for $CS-US_{VISC}$ associations^{20,65} and to ensure tolerability of total scanning time for participants in study 1 (all phases accomplished consecutively), the number of extinction and reinstatement-test trials, respectively, was lower (5 presentations per CS) than in study 2 (12 presentations per CS). Subsequent to the extinction phase, reinstatement procedures involving the unannounced and unexpected re-exposure to the US were accomplished, followed by a reinstatement-test phase, consisting of the same number of CS presentations as during extinction in pseudorandomized order. Given a lack of human reinstatement studies involving the multiple US, and unresolved methodological challenges in the field³⁸, we implemented different reinstatement procedures in an effort to provide procedure-specific, yet complementary data. In study 1, participants were pseudo-randomly assigned to subgroups undergoing a reinstatement procedure with either interoceptive US alone (4 US_{VISC} , $N = 22$) or the exteroceptive US alone (4 US_{SOM} , $N = 20$). By doing so, we aimed to test for reinstatement effects after unexpected re-exposure to threat from one modality (single threat reinstatement) within each reinstatement subgroup. In study 2, all participants ($N = 23$) underwent a reinstatement procedure with both interoceptive and exteroceptive US (3 US_{VISC} , 3 US_{AUD}), aiming to test for reinstatement effects after unexpected re-exposure to threats from multiple modalities (multiple threat reinstatement). Note that the number of unexpected US presentations was chosen based on work in the field³⁸ and our own earlier studies^{20,77,104}. All US intensities and durations implemented as part of reinstatement procedures were identical to those applied during acquisition.

Behavioural measures. For the purposes of concise and parallelised US- and CS-related behavioural and neural analyses across different sensory modalities in two independent studies, we focussed our behavioural data analysis on unpleasantness ratings as a clinically-relevant indicator of emotional valence. Emotional valence is relevant to all types of threat, shapes the perception of aversive stimuli, including pain¹⁰⁵, and drives threat-related behaviours like approach and avoidance¹⁰⁶. It is highly relevant to the specificity of visceral pain^{26,42}, and sensitive to modulation by placebo/nocebo mechanisms^{8,10,107,108}. Prior pain-related conditioning studies from our own group (reviewed in refs. 8,109) and in the broader fear conditioning

literature support the notion that conditioned changes in cue valence constitute a sensitive and relevant behavioural measure capturing the formation, as well as the extinction and return of fear responses in healthy adults¹¹⁰ and clinical populations¹⁰⁶.

All ratings were accomplished on digitised VAS in the scanner using an MRI-compatible hand-held fibre optic response system (LUMItouch™, Photon Control Inc., Burnaby, BC, Canada) before and after learning phases (for specific assessment time points, see Fig. 4; note that in study 2, an additional VAS rating was accomplished mid-extinction (after 6 trials), which we report on in the Supplementary Table S7). In study 1, VAS anchors were labelled 'very pleasant' (−100 mm) and 'very unpleasant' (+100 mm), with the word 'neutral' (0 mm) marked in the middle of the digitised VAS, as accomplished in our previous conditioning work involving painful US^{20,21,65,77,86,111}. In study 2, VAS anchors were labelled 'not at all unpleasant' (0 mm) and 'very unpleasant' (100 mm), consistent with our prior work across sensory modalities^{47,48}. For the purposes of this report and in light of differing scales, we exclusively analysed differential CS valence, computed as individual delta (Δ) scores for each CS⁺ relative to the CS[−] for each learning phase and within each study group. This allows phase-specific comparisons of Δ CS⁺_{VIS} vs. Δ CS⁺_{SOM} in study 1 and Δ CS⁺_{VIS} vs. Δ CS⁺_{AUD} in study 2, in keeping with contrasts computed for brain imaging analyses (see below). Note that we additionally acquired perceived intensity of US_{VISC} and US_{SOM} in study 1; for findings dedicated to elucidating the contributions of intensity versus unpleasantness in the context of visceral pain specificity, see our earlier work^{4,26} and Supplementary analyses herein using these ratings as a covariate of no interest for fMRI data analyses (Supplementary Tables S1–2, 9–10–S9).

To elucidate cognitive awareness of the specific CS–US pairings for each phase, we report contingency awareness as a secondary behavioural measure. To this end, at the conclusion of each experimental phase, for each CS a VAS with ends labelled 'never' (0 mm) and 'always' (100 mm) assessed the perceived probability (0–100%) of a US following a specific CS, as previously described^{20,77,104}. Note that we herein report contingency awareness with a focus on differences between modalities. Statistical analyses of the accuracy of CS–US associations require more complex computations (for an approach, see ref. ¹¹²), which is beyond the scope herein.

Statistical analyses and reproducibility of behavioural data. Statistical analyses of behavioural data were accomplished separately for each study using IBM SPSS Statistics for Windows, version 20 (IBM Corp., Armonk, N.Y., USA). For Δ CS valence, 2 × 2 repeated-measures analyses of variance (rmANOVA) with the factors time (pre, post) and modality (interoceptive, exteroceptive) were computed for each experimental phase, applying the Greenhouse-Geisser correction when the assumption of sphericity was violated. Given our hypotheses and to ensure readability, we provide statistical details on time × modality interaction effects in the main manuscript; full rmANOVA results including all main and interaction effects are given in Supplementary Tables (Supplementary Tables S4 and S5). Two-tailed paired *t*-tests were computed as planned comparisons for two purposes: (1) To test for hypothesis-driven differences between modalities in Δ CS valence, US valence, and contingency awareness at specific time points within studies and experimental groups (all results reported in the main manuscript; further details in Supplementary Tables S6 and S7); and (2) to explore differences in Δ CS valence within modalities across time points (PRE-POST; full results reported in Supplementary Tables S6 and S7). Only Bonferroni-corrected *P*-values are reported within the main manuscript; full uncorrected results of all paired *t*-tests are provided in Supplementary Tables S6 and S7. For rmANOVA, effect sizes are reported as partial eta squared (η_p^2); for *t*-tests, effect sizes are provided based on Cohen's *d* for correlated designs¹¹³. Correlational analyses were accomplished using Pearson's *r*. Results are reported as mean ± standard error of the mean (SEM).

All data are expected to be reproducible given the same settings and procedures as described herein.

Brain imaging data acquisition and analyses. All MR images were acquired using a whole-body 3 Tesla scanner (Skyra, Siemens Healthcare, Erlangen, Germany) equipped with a 32-channel head coil. For functional imaging, single-shot echo-planar imaging (EPI) sequences with similar settings were used (identical for both studies: TE 28.0 ms, flip angle 90°, GRAPPA *r* = 2 with 38 transversal slices angulated in the direction of the corpus callosum, slice thickness of 3 mm, slice gap 0.6 mm, voxel size 2.3 × 2.3 × 3.0 mm; study 1: TR 2300 ms, FOV 220 × 220 mm², matrix 94 × 94 mm²; study 2: TR 2400 ms, FOV 240 × 240 mm², matrix 104 × 104 mm²). Structural images were acquired prior to functional imaging using the same T1-weighted 3D-magnetisation prepared rapid gradient echo (MPRAGE) sequence in both studies [repetition time (TR) 1900 ms, echo time (TE) 2.13 ms, flip angle 9°, field of view (FOV) 239 × 239 mm², 192 slices, slice thickness 0.9 mm, voxel size 0.9 × 0.9 × 0.9 mm³, matrix 256 × 256 mm², Generalised Partially Parallel Acquisitions (GRAPPA) *r* = 2].

Functional images were analysed with SPM software (SPM12, Wellcome Trust Centre for Neuroimaging, UCL, London, UK) implemented in Matlab (R2016b, Mathworks Inc., Sherborn, MA, USA). A standard realignment procedure as implemented in SPM12 was performed for the estimation of six parameters for translation (linear: *x*, *y*, *z* (mm)) and for rotation (degree: pitch, roll, yaw (°)) to describe the rigid body transformation between each image and a reference image.

Subsequently, functional images were co-registered to individual T1-weighted structural images used as reference images, with the origin set to the anterior commissure. Functional images were normalised to Montreal Neurological Institute (MNI) space using a standardised International Consortium for Brain Mapping (ICBM) template for European brains as implemented in SPM12, and smoothed using an isotropic Gaussian kernel of 8 mm. To correct for low-frequency drifts, a temporal high-pass filter with a cut-off set at 128 s was implemented. Serial autocorrelations were taken into consideration by means of an autoregressive model first-order correction.

First-level analyses were performed using a general linear model applied to the EPI images. The time series of each voxel was fitted with a corresponding task regressor that modelled a box car convolved with a canonical hemodynamic response function (HRF). As regressors, CS type (CS⁺_{VISC}; CS⁺_{SOM}/CS⁺_{AUD}; CS[−]) and US modality (US_{VISC}; US_{SOM}/US_{AUD}, only in analyses of acquisition phases) were included. For analyses of CS-induced activations, durations were used exactly as implemented in the experiments (jittered between 6 and 12 s before US presentation), for analyses of US-induced activations, ascending and plateau phases of US stimulation were included in analyses (20 s in study 1, 14 s in study 2). Six realignment parameters for translation and rotation were additionally implemented as multiple regressors for motion correction. After model estimation, the following first-level contrasts and respective reverse contrasts were computed for analyses of differential CS-related and US-related neural responses separately for each study group: CS⁺_{VISC} > CS[−], CS⁺_{SOM} > CS[−], US_{VISC} > US_{SOM} for study 1; CS⁺_{VISC} > CS[−], CS⁺_{AUD} > CS[−], US_{VISC} > US_{AUD} for study 2. CS contrasts were computed for each phase, US contrasts only for the acquisition phase.

On the second level, for analyses of US-induced neural activation, one-sample *t*-tests based on these differential first-level contrasts and paired *t*-tests were calculated. For analyses of CS-induced differential neural activation, paired *t*-tests were computed for each experimental phase to compare Δ CS⁺_{VISC} versus Δ CS⁺_{SOM} in study 1 ($\{CS^+_{VISC} < CS^-\} > \{CS^+_{SOM} < CS^-\}$; $\{CS^+_{VISC} > CS^-\} > \{CS^+_{SOM} > CS^-\}$) and Δ CS⁺_{VISC} versus Δ CS⁺_{AUD} in study 2 ($\{CS^+_{VISC} < CS^-\} > \{CS^+_{AUD} < CS^-\}$; $\{CS^+_{VISC} > CS^-\} > \{CS^+_{AUD} > CS^-\}$). Additional exploratory analyses included selected covariates of no interest, as indicated in the results. Further, extending our earlier findings revealing not only distinct but also shared neural activations for US⁴ as well as CS²⁶ across modalities, conjunction analyses using first-level contrasts were carried out to identify joint activations (i.e. CS⁺_{VISC} > CS[−] ∩ CS⁺_{SOM} > CS[−], US_{VISC} ∩ US_{SOM} for study 1; CS⁺_{VISC} > CS[−] ∩ CS⁺_{AUD} > CS[−], US_{VISC} ∩ US_{AUD} for study 2). Conjunction analyses were computed (a) using the minimum statistic to the conjunction null to test for shared activation within all tested subjects, and (b) using the minimum statistics to the global null to test for shared activation within some subjects^{114,115}. For correlational analyses exploring associations between differential CS and differential US activation in specific ROI (provided in Supplementary analyses), parameter estimates were extracted for peak-voxels in significant regions of interest (ROIs) as identified by one-sample *t*-tests.

All analyses focused on a priori defined ROIs of the salience and extinction networks^{26,51,55–57,66,68}, including the insula (anterior, aINS; posterior, pINS), subregions of the cingulate cortex (midcingulate cortex, MCC; dorsal anterior cortex, dACC), amygdala, hippocampus, and ventromedial prefrontal cortex (vmPFC). All ROI analyses were carried out using unilateral anatomical templates constructed from the WFU Pick Atlas (Version 2.5.2), as implemented in SPM12. Segmentation of the insula (aINS, pINS) and cingulate cortex (dACC, MCC) was accomplished with masks based on the previous literature¹¹⁶ within the borders of the Wake Forest University (WFU) Pick Atlas. For all reported ROI analyses, family-wise-error (FWE) correction for multiple testing was used with statistical significance set at $P_{FWE} < 0.05$, and coordinates refer to the MNI space. Supplementary whole-brain analyses (uncorrected $P < 0.001$) were additionally carried out (Tables 1, 3–6; Supplementary Tables S1–3, S8–12; Supplementary Figs. S4–S7 for visualisation). Note that the results presented within the main manuscript all focus on ROI analyses unless explicitly specified otherwise.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All fMRI data analysed for the current study are available in the neurovault repository (<https://neurovault.org/collections/GPPGVZAT/>). Behavioural and SCR data are provided in the main manuscript or its Supplementary Information; additional data and information upon request.

Code availability

No custom code or mathematical algorithms were used in the study. All software used for statistical analyses has been declared in the manuscript.

Received: 7 May 2020; Accepted: 18 March 2021;

Published online: 11 May 2021

References

- Pittig, A., Treanor, M., LeBeau, R. T. & Craske, M. G. The role of associative fear and avoidance learning in anxiety disorders: gaps and directions for future research. *Neurosci. Biobehav. Rev.* **88**, 117–140 (2018).
- Vlaeyen, J. W. Learning to predict and control harmful events: chronic pain and conditioning. *Pain* **156**(Suppl 1), S86–S93 (2015).
- Vlaeyen, J. W., Crombez, G. & Linton, S. J. The fear-avoidance model of pain. *Pain* **157**, 1588–1589 (2016).
- Koenen, L. R. et al. Greater fear of visceral pain contributes to differences between visceral and somatic pain in healthy women. *Pain* **158**, 1599–1608 (2017).
- Verne, N. G. et al. Central representation of visceral and cutaneous hypersensitivity in the irritable bowel syndrome. *Pain* **103**, 99–110 (2003).
- Van Oudenhove, L. et al. Biopsychosocial aspects of functional gastrointestinal disorders. *Gastroenterology* <https://doi.org/10.1053/j.gastro.2016.02.027> (2016).
- Boeckxstaens, G. et al. Fundamentals of neurogastroenterology: physiology/motility - sensation. *Gastroenterology* <https://doi.org/10.1053/j.gastro.2016.02.030> (2016).
- Elsenbruch, S. & Labrenz, F. Nocebo effects and experimental models in visceral pain. *Int. Rev. Neurobiol.* **138**, 285–306 (2018).
- Schmid, J. et al. Neural underpinnings of nocebo hyperalgesia in visceral pain: A fMRI study in healthy volunteers. *Neuroimage* **120**, 114–122 (2015).
- Elsenbruch, S. & Enck, P. Placebo effects and their determinants in gastrointestinal disorders. *Nat. Rev. Gastroenterol. Hepatol.* **12**, 472–485 (2015).
- Schmid, J. et al. Neural mechanisms mediating positive and negative treatment expectations in visceral pain: a functional magnetic resonance imaging study on placebo and nocebo effects in healthy volunteers. *Pain* **154**, 2372–2380 (2013).
- Roderigo, T. et al. Effects of acute psychological stress on placebo and nocebo responses in a clinically relevant model of visceroreception. *Pain* **158**, 1489–1498 (2017).
- Evers, A. W. M. et al. Implications of placebo and nocebo effects for clinical practice: expert consensus. *Psychother. Psychosom.* **87**, 204–210 (2018).
- Schedlowski, M., Enck, P., Rief, W. & Bingel, U. Neuro-bio-behavioral mechanisms of placebo and nocebo responses: implications for clinical trials and clinical practice. *Pharm. Rev.* **67**, 697–730 (2015).
- Colloca, L. & Barsky, A. J. Placebo and nocebo effects. *N. Engl. J. Med.* **382**, 554–561 (2020).
- den Hollander, M., Meulders, A., Jakobs, M. & Vlaeyen, J. W. The effect of threat information on acquisition, extinction, and reinstatement of experimentally conditioned fear of movement-related pain. *Pain. Med.* **16**, 2302–2315 (2015).
- Jensen, K., Kirsch, I., Odmalm, S., Kaptchuk, T. J. & Ingvar, M. Classical conditioning of analgesic and hyperalgesic pain responses without conscious awareness. *Proc. Natl Acad. Sci. USA* **112**, 7863–7867 (2015).
- Jensen, K. B. et al. Nonconscious activation of placebo and nocebo pain responses. *Proc. Natl Acad. Sci. USA* **109**, 15959–15964 (2012).
- Meulders, A., Jans, A. & Vlaeyen, J. W. Differences in pain-related fear acquisition and generalization: an experimental study comparing patients with fibromyalgia and healthy controls. *Pain* **156**, 108–122 (2015).
- Gramsch, C. et al. Learning pain-related fear: neural mechanisms mediating rapid differential conditioning, extinction and reinstatement processes in human visceral pain. *Neurobiol. Learn Mem.* **116**, 36–45 (2014).
- Icenhour, A. et al. Learning by experience? Visceral pain-related neural and behavioral responses in a classical conditioning paradigm. *Neurogastroenterol. Motil.* **29** <https://doi.org/10.1111/nmo.13026> (2017).
- Ceunen, E. et al. Learned fear of gastrointestinal sensations in healthy adults. *Clin. Gastroenterol. Hepatol.* **14**, 1552–1558 (2016).
- Yaguez, L. et al. Brain response to visceral aversive conditioning: a functional magnetic resonance imaging study. *Gastroenterology* **128**, 1819–1829 (2005).
- Biggs, E. E. et al. The neural correlates of pain-related fear: a meta-analysis comparing fear conditioning studies using painful and non-painful stimuli. *Neurosci. Biobehav. Rev.* **119**, 52–65 (2020).
- Meulders, A. Fear in the context of pain: Lessons learned from 100 years of fear conditioning research. *Behav. Res. Ther.* **131**, 103635 (2020).
- Koenen, L. R. et al. From anticipation to the experience of pain: the importance of visceral versus somatic pain modality in neural and behavioral responses to pain-predictive cues. *Psychosom. Med.* **80**, 826–835 (2018).
- Azzalini, D., Rebollo, I. & Tallon-Baudry, C. Visceral signals shape brain dynamics and cognition. *Trends Cogn. Sci.* **23**, 488–509 (2019).
- Ceunen, E., Vlaeyen, J. W. & Van Diest, I. On the origin of interoception. *Front Psychol.* **7**, 743 (2016).
- Paulus, M. P., Feinstein, J. S. & Khalsa, S. S. An active inference approach to interoceptive psychopathology. *Annu. Rev. Clin. Psychol.* **15**, 97–122 (2019).
- Sharvit, G., Corradi-Dell'Acqua, C. & Vuilleumier, P. Modality-specific effects of aversive expectancy in the anterior insula and medial prefrontal cortex. *Pain* **159**, 1529–1542 (2018).
- Fazeli, S. & Buchel, C. Pain-related expectation and prediction error signals in the anterior insula are not related to aversiveness. *J. Neurosci.* **38**, 6461–6474 (2018).
- Sharvit, G., Vuilleumier, P. & Corradi-Dell'Acqua, C. Sensory-specific predictive models in the human anterior insula. *f1000Res* **8**, 164 (2019).
- Au Yeung, S. T., Colagiuri, B., Lovibond, P. F. & Colloca, L. Partial reinforcement, extinction, and placebo analgesia. *Pain* **155**, 1110–1117 (2014).
- Colagiuri, B., Quinn, V. F. & Colloca, L. Nocebo hyperalgesia, partial reinforcement, and extinction. *J. Pain.* **16**, 995–1004 (2015).
- Razran, G. The observable unconscious and the inferable conscious in current Soviet psychophysiology: interoceptive conditioning, semantic conditioning, and the orienting reflex. *Psychol. Rev.* **68**, 1–147 (1961).
- Pappens, M. et al. Learning to fear obstructed breathing: comparing interoceptive and exteroceptive cues. *Biol. Psychol.* **92**, 36–42 (2013).
- Van Oudenhove, L. et al. Common and distinct neural representations of aversive somatic and visceral stimulation in healthy individuals. *Nat. Commun.* **11**, 5939 (2020).
- Haaker, J., Golkar, A., Hermans, D. & Lonsdorf, T. B. A review on human reinstatement studies: an overview and methodological challenges. *Learn Mem.* **21**, 424–440 (2014).
- Vlaeyen, J. W. S. & Crombez, G. Behavioral conceptualization and treatment of chronic pain. *Annu. Rev. Clin. Psychol.* **16**, 187–212 (2019).
- Oehman, A. & Mineka, S. Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychol. Rev.* **108**, 483–522 (2001).
- Stockhorst, U., Enck, P. & Klosterhalfen, S. Role of classical conditioning in learning gastrointestinal symptoms. *World J. Gastroenterol.* **13**, 3430–3437 (2007).
- Dunckley, P. et al. Cortical processing of visceral and somatic stimulation: differentiating pain intensity from unpleasantness. *Neuroscience* **133**, 533–542 (2005).
- Mayer, E. A. Gut feelings: the emerging biology of gut-brain communication. *Nat. Rev. Neurosci.* **12**, 453–466 (2011).
- Weltens, N., Iven, J., Van Oudenhove, L. & Kano, M. The gut-brain axis in health neuroscience: implications for functional gastrointestinal disorders and appetite regulation. *Ann. N. Y. Acad. Sci.* **1428**, 129–150 (2018).
- Labanski, A., Langhorst, J., Engler, H. & Elsenbruch, S. Stress and the brain-gut axis in functional and chronic-inflammatory gastrointestinal diseases: a transdisciplinary challenge. *Psychoneuroendocrinology* **111**, 104501 (2020).
- Icenhour, A. et al. Are there sex differences in visceral sensitivity in young healthy men and women? *Neurogastroenterol. Motil.* **31**, e13664 (2019).
- Forkmann, K. et al. Pain-specific modulation of hippocampal activity and functional connectivity during visual encoding. *J. Neurosci.* **33**, 2571–2581 (2013).
- Forkmann, K., Wiech, K., Sommer, T. & Bingel, U. Reinstatement of pain-related brain activation during the recognition of neutral images previously paired with nociceptive stimuli. *Pain* **156**, 1501–1510 (2015).
- Buchel, C. & Dolan, R. J. Classical fear conditioning in functional neuroimaging. *Curr. Opin. Neurobiol.* **10**, 219–223 (2000).
- Fullana, M. A. et al. Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies. *Mol. Psychiatry* **21**, 500–508 (2016).
- Sehlmeyer, C. et al. Human fear conditioning and extinction in neuroimaging: a systematic review. *PLoS ONE* **4**, e5865 (2009).
- Wiech, K. Deconstructing the sensation of pain: the influence of cognitive processes on pain perception. *Science* **354**, 584–587 (2016).
- Atlas, L. Y. & Wager, T. D. How expectations shape pain. *Neurosci. Lett.* **520**, 140–148 (2012).
- Bushnell, M. C., Ceko, M. & Low, L. A. Cognitive and emotional control of pain and its disruption in chronic pain. *Nat. Rev. Neurosci.* **14**, 502–511 (2013).
- Menon, V. & Uddin, L. Q. Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* **214**, 655–667 (2010).
- Seeley, W. W. et al. Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* **27**, 2349–2356 (2007).
- Wiech, K. et al. Anterior insula integrates information about salience into perceptual decisions about pain. *J. Neurosci.* **30**, 16324–16331 (2010).
- Claassen, J. et al. Cerebellum is more concerned about visceral than somatic pain. *J. Neurol. Neurosurg. Psychiatry* **91**, 218–219 (2020).
- Palermo, S., Benedetti, F., Costa, T. & Amanzio, M. Pain anticipation: an activation likelihood estimation meta-analysis of brain imaging studies. *Hum. Brain Mapp.* **36**, 1648–1661 (2015).
- Craig, A. D. Interoception: the sense of the physiological condition of the body. *Curr. Opin. Neurobiol.* **13**, 500–505 (2003).

61. Jepma, M., Koban, L., van Doorn, J., Jones, M. & Wager, T. D. Behavioural and neural evidence for self-reinforcing expectancy effects on pain. *Nat. Hum. Behav.* **2**, 838–855 (2018).
62. Koban, L., Jepma, M., Lopez-Sola, M. & Wager, T. D. Different brain networks mediate the effects of social and conditioned expectations on pain. *Nat. Commun.* **10**, 4096 (2019).
63. Bouton, M. E. & Moody, E. W. Memory processes in classical conditioning. *Neurosci. Biobehav. Rev.* **28**, 663–674 (2004).
64. Vervliet, B., Craske, M. G. & Hermans, D. Fear extinction and relapse: state of the art. *Annu. Rev. Clin. Psychol.* **9**, 215–248 (2013).
65. Icenhour, A. et al. Neural circuitry underlying effects of context on human pain-related fear extinction in a renewal paradigm. *Hum. Brain Mapp.* **36**, 3179–3193 (2015).
66. Quirk, G. J. & Mueller, D. Neural mechanisms of extinction learning and retrieval. *Neuropsychopharmacology* **33**, 56–72 (2008).
67. Schafe, G. E. & LeDoux, J. E. Memory consolidation of auditory pavlovian fear conditioning requires protein synthesis and protein kinase A in the amygdala. *J. Neurosci.* **20**, RC96–RC96 (2000).
68. Fullana, M. A. et al. Fear extinction in the human brain: ameta-analysis of fMRI studies in healthy participants. *Neurosci. Biobehav. Rev.* **88**, 16–25 (2018).
69. Craske, M. G., Hermans, D. & Vervliet, B. State-of-the-art and future directions for extinction as a translational model for fear and anxiety. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **373**, <https://doi.org/10.1098/rstb.2017.0025> (2018).
70. Pittig, A., van den Berg, L. & Vervliet, B. The key role of extinction learning in anxiety disorders: behavioral strategies to enhance exposure-based treatments. *Curr. Opin. Psychiatry* **29**, 39–47 (2016).
71. Zbozinek, T. D., Hermans, D., Prenoveau, J. M., Liao, B. & Craske, M. G. Post-extinction conditional stimulus valence predicts reinstatement fear: relevance for long-term outcomes of exposure therapy. *Cogn. Emot.* **29**, 654–667 (2015).
72. Vasey, M. W., Harbaugh, C. N., Buffington, A. G., Jones, C. R. & Fazio, R. H. Predicting return of fear following exposure therapy with an implicit measure of attitudes. *Behav. Res. Ther.* **50**, 767–774 (2012).
73. Scharfenort, R. & Lonsdorf, T. B. Neural correlates of and processes underlying generalized and differential return of fear. *Soc. Cogn. Affect Neurosci.* **11**, 612–620 (2016).
74. Scharfenort, R., Menz, M. & Lonsdorf, T. B. Adversity-induced relapse of fear: neural mechanisms and implications for relapse prevention from a study on experimentally induced return-of-fear following fear conditioning and extinction. *Transl. Psychiatry* **6**, e858 (2016).
75. Lonsdorf, T. B., Haaker, J. & Kalisch, R. Long-term expression of human contextual fear and extinction memories involves amygdala, hippocampus and ventromedial prefrontal cortex: a reinstatement study in two independent samples. *Soc. Cogn. Affect Neurosci.* **9**, 1973–1983 (2014).
76. Benson, S. et al. Towards understanding sex differences in visceral pain: enhanced reactivation of classically-conditioned fear in healthy women. *Neurobiol. Learn Mem.* **109**, 113–121 (2014).
77. Icenhour, A. et al. Neural circuitry of abdominal pain-related fear learning and reinstatement in irritable bowel syndrome. *Neurogastroenterol. Motil.* **27**, 114–127 (2015).
78. Labus, J. S. et al. Impaired emotional learning and involvement of the corticotropin-releasing factor signaling system in patients with irritable bowel syndrome. *Gastroenterology* **145**, 1253–1261 (2013). e1251–1253.
79. Dunsmoor, J. E. & Murphy, G. L. Categories, concepts, and conditioning: how humans generalize fear. *Trends Cogn. Sci.* **19**, 73–77 (2015).
80. Dunsmoor, J. E. & Paz, R. Fear generalization and anxiety: behavioral and neural mechanisms. *Biol. Psychiatry* **78**, 336–343 (2015).
81. Harvie, D. S., Moseley, G. L., Hillier, S. L. & Meulders, A. Classical conditioning differences associated with chronic pain: a systematic review. *J. Pain.* **18**, 889–898 (2017).
82. Zaman, J. et al. The influence of pain-related expectations on intensity perception of nonpainful somatosensory stimuli. *Psychosom. Med.* **80**, 836–844 (2018).
83. Elsenbruch, S. & Wolf, O. T. Could stress contribute to pain-related fear in chronic pain? *Front. Behav. Neurosci.* **9**, 340 (2015).
84. Labrenz, F., Knuf-Rtveliashvili, S. & Elsenbruch, S. Sex differences linking pain-related fear and interoceptive hypervigilance: attentional biases to conditioned threat and safety signals in a visceral pain model. *Front. Psychiatry* **11**, <https://doi.org/10.3389/fpsy.2020.00197> (2020).
85. Timmers, I. et al. The interaction between stress and chronic pain through the lens of threat learning. *Neurosci. Biobehav. Rev.* **107**, 641–655 (2019).
86. Benson, S. et al. Cortisol affects pain sensitivity and pain-related emotional learning in experimental visceral but not somatic pain: a randomized controlled study in healthy men and women. *Pain* **160**, 1719–1728 (2019).
87. Icenhour, A., Labrenz, F., Roderigo, T., Benson, S. & Elsenbruch, S. The role of chronic stress in normal visceroperception: insights from an experimental visceral pain study in healthy volunteers. *Front. Psychiatry* **11**, 107 (2020).
88. Allen, M., Legrand, N., Costa Correa, C. M. & Fardo, F. Thinking through prior bodies: autonomic uncertainty and interoceptive self-inference. *Behav. Brain Sci.* **43**, e91 (2020).
89. Seth, A. K. Interoceptive inference, emotion, and the embodied self. *Trends Cogn. Sci.* **17**, 565–573 (2013).
90. Grahl, A., Onat, S. & Buchel, C. The periaqueductal gray and Bayesian integration in placebo analgesia. *Elife* **7**, <https://doi.org/10.7554/eLife.32930> (2018).
91. Onat, S. & Buchel, C. The neuronal basis of fear generalization in humans. *Nat. Neurosci.* **18**, 1811–1818 (2015).
92. Allen, M., Levy, A., Parr, T. & Friston, K. J. In the body's eye: the computational anatomy of interoceptive inference. Preprint at <https://www.biorxiv.org/content/10.1101/603928v1> (2019).
93. Allen, M. & Tsakiris, M. The body as first prior: Interoceptive predictive processing and the primacy. In *The Interoceptive Mind: From Homeostasis to Awareness: Volume 27*, 27–45 (Oxford University Press, 2018).
94. Quinn, V. F. & Colagiuri, B. Using learning strategies to inhibit the nocebo effect. *Int. Rev. Neurobiol.* **138**, 307–327 (2018).
95. Urits, I. et al. An update on cognitive therapy for the management of chronic pain: a comprehensive review. *Curr. Pain. Headache Rep.* **23**, 57 (2019).
96. Schemann, M., Frieling, T. & Enck, P. To learn, to remember, to forget—How smart is the gut? *Acta Physiol.* **228**, e13296 (2020).
97. Keefer, L. Behavioural medicine and gastrointestinal disorders: the promise of positive psychology. *Nat. Rev. Gastroenterol. Hepatol.* **15**, 378–386 (2018).
98. Lacourt, T. E. et al. Biological and psychological predictors of visceral pain sensitivity in healthy premenopausal women. *Eur. J. Pain.* **18**, 567–574 (2014).
99. Herrmann-Lingen, C., Buss, U. & Snaith, R. P. *Hospital anxiety and depression scale: HADS-D; deutsche Version; ein Fragebogen zur Erfassung von Angst und Depressivität in der somatischen Medizin; Testdokumentation und Handanweisung* (2007).
100. Reiss, M. & Reiss, G. Studies on motorial asymmetries. *Fortschr. der Neurologie Psychiatr.* **68**, 70–79 (2000).
101. Oldfield, R. C. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* **9**, 97–113 (1971).
102. Schulz, P. & Schlotz, W. The Trier Inventory for Chronic Stress (TICS): scale construction, statistical testing and validation of the scale work overload. *Diagnostica* **45**, 8–19 (1999).
103. Keszthelyi, D., Troost, F. J. & Masclee, A. A. Irritable bowel syndrome: methods, mechanisms, and pathophysiology. Methods to assess visceral hypersensitivity in irritable bowel syndrome. *Am. J. Physiol. Gastrointest. Liver Physiol.* **303**, G141–G154 (2012).
104. Kattoor, J. et al. Fear conditioning in an abdominal pain model: neural responses during associative learning and extinction in healthy subjects. *PLoS ONE* **8**, e51149 (2013).
105. Porreca, F. & Navratilova, E. Reward, motivation, and emotion of pain and its relief. *Pain* **158**(Suppl 1), S43–S49 (2017).
106. Dour, H. J., Brown, L. A. & Craske, M. G. Positive valence reduces susceptibility to return of fear and enhances approach behavior. *J. Behav. Ther. Exp. Psychiatry* **50**, 277–282 (2016).
107. Elsenbruch, S., Roderigo, T., Enck, P. & Benson, S. Can a brief relaxation exercise modulate placebo or nocebo effects in a visceral pain model? *Front. Psychiatry* **10**, 144 (2019).
108. Elsenbruch, S. et al. How positive and negative expectations shape the experience of visceral pain: an experimental pilot study in healthy women. *Neurogastroenterol. Motil.* **24**, 914–e460 (2012).
109. Enck, P., Chae, Y. & Elsenbruch, S. Novel designs and paradigms to study the placebo response in gastroenterology. *Curr. Opin. Pharm.* **37**, 72–79 (2017).
110. Dirikx, T., Hermans, D., Vansteenwegen, D., Baeyens, F. & Eelen, P. Reinstatement of extinguished conditioned responses and negative stimulus valence as a pathway to return of fear in humans. *Learn Mem.* **11**, 549–554 (2004).
111. Labrenz, F. et al. From Pavlov to pain: How predictability affects the anticipation and processing of visceral pain in a fear conditioning paradigm. *Neuroimage* **130**, 104–114 (2016).
112. Labrenz, F., Icenhour, A., Benson, S. & Elsenbruch, S. Contingency awareness shapes acquisition and extinction of emotional responses in a conditioning model of pain-related fear. *Front. Behav. Neurosci.* **9**, 318 (2015).
113. Dunlap, W. P., Cortina, J. M., Vaslow, J. B. & Burke, M. J. Meta-analysis of experiments with matched groups or repeated measures designs. *Psychol. Methods* **1**, 170–177 (1996).
114. Friston, K. J., Holmes, A. P., Price, C. J., Buchel, C. & Worsley, K. J. Multisubject fMRI studies and conjunction analyses. *Neuroimage* **10**, 385–396 (1999).

115. Nichols, T., Brett, M., Andersson, J., Wager, T. & Poline, J. B. Valid conjunction inference with the minimum statistic. *Neuroimage* **25**, 653–660 (2005).
116. Deen, B., Pitskel, N. B. & Pelphrey, K. A. Three systems of insular functional connectivity identified with cluster analysis. *Cereb. Cortex* **21**, 1498–1506 (2011).
117. Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R. & Kievit, R. A. Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Res.* **4**, 63 (2019).

Acknowledgements

We would like to thank Alexandra Kornowski and Dr Marcel Gratz for their excellent technical support. We also thank Sopiko Knuf-Rtveliashvili, Nelly Hazzan, Dr. Alexandra Labanski, Lea Schliephake and Katharina Krawczyk (née Fleischer) for support in data acquisition. The study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation; project number 316803389; SFB 1280, subprojects A10 and A12). The funding organisation was not involved in study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

Author contributions

L.R.K., R.J.P. and L.P. performed the research; S.E., H.E., A.I. and L.R.K. designed the research study; L.R.K., R.J.P., L.P., K.F. and N.T. analysed the data; L.R.K., A.I. and S.E. wrote the first draft of the paper; S.E. and H.E. acquired funding; all authors contributed to the interpretation of the data, revised the manuscript for critical content, and approved the final version of the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-021-02008-1>.

Correspondence and requests for materials should be addressed to S.E.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



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

© The Author(s) 2021, corrected publication 2021