Archival Report

Biological Psychiatry: CNNI

Neural Processing During Fear Extinction Predicts Intrusive Memories

Stephan F. Miedl, Julina A. Rattel, Laila K. Franke, Jens Blechert, Martin Kronbichler, Victor I. Spoormaker, and Frank H. Wilhelm

ABSTRACT

BACKGROUND: Deficient extinction learning has been suggested as an important mechanism involved in the etiology of posttraumatic stress disorder. A key feature of posttraumatic stress disorder, reexperiencing the trauma in form of intrusions, may be linked to deficient extinction learning. This link is investigated in a novel, functional magnetic resonance imaging–compatible fear conditioning procedure that uses trauma films. Based on previous results, we expected deficient fear extinction indexed by exaggerated responding in the anterior insula and dorsal anterior cingulate cortex to predict subsequent intrusions.

METHODS: A total of 58 healthy participants underwent acquisition and extinction learning with faces as conditioned stimuli (CS) and highly aversive 16-second films depicting interpersonal violence as unconditioned stimuli. During the subsequent 3 days, participants reported intrusive memories on their smartphone.

RESULTS: Successful fear acquisition was evidenced by differential (CS+ > CS−) activity (threat cues associated with trauma films > cues paired only with neutral films) of a widespread network, including the anterior insula and dorsal anterior cingulate cortex, whereas extinction was characterized exclusively by differential anterior insula activity. Differential conditioned responding during late extinction in the anterior insula and dorsal anterior cingulate cortex was positively related to intrusive memory frequency independent of unconditioned stimuli responding. Exploratory analysis also revealed intrusion sensitivity of the hippocampus, rostral anterior cingulate cortex, and ventromedial prefrontal cortex, among others.

CONCLUSIONS: Results support the role of extinction learning in intrusive memory formation; a failure to uncouple conditioned emotional responding from external threat cues was associated with subsequent intrusive memories, representing a potential risk marker for developing posttraumatic stress disorder symptomatology after trauma.

Keywords: Extinction memory, Fear conditioning, fMRI, Intrusions, Posttraumatic stress disorder, Trauma film

https://doi.org/10.1016/j.bpsc.2019.12.017

Up to 90% of the population experiences a traumatic event during their lifetime (1). Most people will reexperience the event by highly vivid, emotional, and involuntary recollections, although those so-called intrusions typically subside after a few days. For some people, however, intrusions will persist, constituting a hallmark symptom of posttraumatic stress disorder (PTSD) (2–4). Our current understanding of how intrusive memories develop and why they persist with such high frequency and intensity for some individuals is limited. Theoretical accounts propose that intrusions develop through associative learning mechanisms (5); however, this has not yet been sufficiently tested by experimental studies. The current study is the first to investigate whether increased neural fear acquisition and/or deficient extinction learning can explain intrusive memory formation.

Conditioning research proposes alterations in associative learning (i.e., acquisition and extinction learning) as a core mechanism linked to PTSD development (6–8). Conceptually, during fear acquisition, a neutral stimulus (e.g., a specific sound, object, or person in temporal and spatial proximity to the occurrence of the traumatic event) is paired with a traumatic event (unconditioned stimulus [US]). This turns the neutral stimulus into a conditioned stimulus (CS+) that will subsequently elicit a conditioned response (CR) in the absence of the US. Subsequently, during extinction learning, the CS+ is presented in the absence of the US, and thereby the CR should gradually decrease. Alterations in associative learning have been reported for patients with PTSD linked to heightened fear acquisition as well as deficient extinction learning (9). Deficient extinction learning has been proposed to be a key mechanism causing PTSD (10), and deficient pretrauma extinction learning has been linked to PTSD development following trauma (11). In addition, theoretical accounts propose that intrusions, a core symptom of PTSD, develop through associative learning (12); thereby, intrusions can be conceptualized as a CR to cues resembling stimuli that were present around the traumatic event and thus may continue to predict an aversive event (US) (5,13). In

SEE COMMENTARY ON PAGE 375

ISSN: 2451-9022

summary, clinical studies and theoretical accounts propose a key role of deficient extinction learning in the persistence of intrusions following trauma.

Using experimental analog designs, the etiological role of conditioning mechanisms and trauma memories in the development of PTSD-like symptoms can be studied under controlled experimental conditions in real time. Previous analog studies used either the trauma-film paradigm (14) or fear conditioning paradigms with basic aversive stimulation such as electric shock (15). Using the trauma-film paradigm, participants watch highly aversive film clips in the laboratory and subsequently report any intrusive recollections of those clips. Applying fear conditioning paradigms, associative learning mechanisms can be investigated, although the assessment of intrusions is limited when using electric shock as the US. Therefore, Wegerer *et al.* combined those two experimental approaches in their so-called conditioned-intrusion paradigm (16–18). Highly aversive film clips served as the US and were paired with a neutral stimulus right before film onset during acquisition (CS+). Using this paradigm, both deficient evaluative and physiological extinction learning have been related to intrusive memories (16). Moreover, Rattel *et al.* (19) revealed that although both acquisition and extinction learning were related to intrusions, deficient extinction learning was the driving mechanism. In summary, first experimental studies corroborate claims from theoretical accounts proposing that intrusions develop through deficient extinction learning.

Recent models on functional neuroanatomical networks in PTSD propose that intrusions develop through pathological neural encoding of trauma. These models propose that PTSD is characterized by general hyperactivity in threat processing structures such as the amygdala, anterior insula, and dorsal anterior cingulate cortex (dACC) and by hypoactivity in the ventral medial prefrontal cortex (VMPFC) (20–23). However, a more recent meta-analysis by Stark *et al.* (24) proposed hyperactivity in the anterior insula in PTSD compared with trauma-naïve control subjects as the core region linked to PTSD development. Using the trauma-film paradigm, networks around the ACC have been linked to the formation of analog intrusions (25–27). Furthermore, the neural pathophysiology proposed by neuroscientific models of PTSD has also been linked to increased fear acquisition and deficient extinction learning (23). In patients with PTSD compared with trauma-exposed healthy control subjects, fear acquisition and extinction learning have been related to the insula, dACC, and amygdala (28). However, in healthy subjects, fear acquisition (29) and extinction learning (30) have particularly been linked to brain regions implicated in threat appraisal—the anterior insula and dACC—but without consistent findings for the amygdala or VMPFC. This is further supported by a recent review proposing the dACC and anterior insula as a common core of areas affected across most categories of psychiatric illness (31). What is unclear so far is whether enhanced neural activity in these core regions, particularly during fear extinction, is linked to intrusive memory formation.

The Current Study

Using the conditioned intrusion paradigm adapted for magnetic resonance imaging (MRI), the current study set out to investigate whether associative neural responses are linked to subsequent intrusion formation. During acquisition, using neutral faces as the CS and film clips depicting severe interpersonal violence as the US, we examined the neural activity to learned threat cues associated with trauma films (CS+) in comparison with cues that were paired only with neutral films (CS−). During extinction, when the CS+ was not paired with trauma films anymore, the same differential (CS+ > CS−) neural activity was traced. We primarily expected CR activations of the anterior insula and dACC during extinction learning. Moreover, additional analyses also checked for similar neural activation patterns during fear acquisition. In line with findings by Rattel *et al.* (19), we expected increased differential activity of the anterior insula and dACC particularly during late extinction to be linked to the frequency of intrusive memories on subsequent days.

METHODS AND MATERIALS

Participants

A total of 60 healthy female participants were recruited for this study. Exclusion criteria were blood injection injury phobia, self-report of psychosis, psychotropic medication use, substance abuse/dependency, bipolar disorder, serious medical conditions, anxiety, depression, PTSD, or history of traumatic head injury. Further exclusion criteria were extensive media consumption of violent and/or medical content (more than three times a week) and poor sleep quality [score of 7 or lower on the Pittsburgh Sleep Quality Index (32)]. Participants answering "no" to the question "Are you currently mentally and physically resilient?" were excluded. Having experienced traumatic events in the past was not an exclusion criterion because this is quite common. For functional MRI (fMRI), exclusion criteria were pregnancy, ferromagnetic implants, other nonremovable metal objects, and claustrophobia. Two participants needed to be excluded owing to technical problems at the MRI scanner. Thus, 58 participants (mean age = 22.8 years, SD = 4.1) were included in the final analyses. The study was approved by the local ethics committee. Participants provided informed consent before participation.

Materials and Procedure

One week before fMRI scanning, participants completed the State-Trait Anxiety Inventory [German version (33)], the Center for Epidemiologic Studies Depression Scale [German version: Allgemeine Depressionsskala (34)], and a questionnaire assessing habitual consumption of television and film footage/video games depicting severe violence. Participants also prerated 30 emotionally neutral faces [taken from the Radboud Faces Database (35)], with the two most neutral faces in valence and least arousing faces being used later in the conditioning procedure (individually chosen for each participant).

The current study used an adaption of the conditioned intrusion paradigm previously developed by Wegerer *et al.* (16,18). The two neutral face stimuli were randomly assigned to serve as the CS+ and CS− (lasting 4 seconds each); CS+ was followed by an aversive film clip, and CS− was followed by a neutral film clip. To prevent habituation to a single film clip over repeated presentations, six film scenes depicting severe interpersonal violence served as the US. Aversive and neutral clips were extracted from commercial movies (see

Supplement). Each film clip lasted 16 seconds. Stimulus presentation and behavioral data acquisition were controlled by E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA).

Fear Conditioning Task. The fear conditioning task consisted of acquisition and immediate extinction without delay between acquisition and extinction. Trials were pseudorandomized, with no more than three stimuli of the same type (CS+ or CS–) presented consecutively. Intertrial intervals varied between 10 and 14 seconds.

Acquisition Phase. Similar to other studies [e.g., (9)], participants were informed that one of the two presented faces might be followed by an aversive film clip, whereas the other face would not be followed by an aversive film clip. In total, 16 CS+ and 16 CS− trials were presented, of which 4 CS+ and 4 CS– trials were presented without the film (75% reinforcement rate). No pause was inserted after unreinforced CS+ to compensate for the omission of the film clips. Thus, each of the six film clips was presented two times in pseudorandom order.

Extinction Phase. In total, 16 CS+ and 16 CS– trials were presented without subsequent films (no pause was inserted to compensate for the omission of the film clips).

Ratings. Ratings for US expectancy ("How much do you expect this face to be followed by an aversive film clip during its next presentation?" from 1 = very low expectancy to 9 = very high expectancy) and for negative valence ("How unpleasant does this face appear to you?" from 1 = not unpleasant at all to 9 = very unpleasant) were acquired at the end of early (after 8 CS+ and 8 CS– trials) and late (after 16 CS+ and 16 CS– trials) acquisition and extinction, respectively.

Ambulatory Assessment of Intrusive Memories. After scanning, participants were instructed to report any intrusive memories of the film scenes and faces seen in the experiment over the following 3 days. Intrusive memories were defined as recurring images or thoughts about the faces or films but also as recurring thoughts or feelings that had been present during watching (20–23). Participants were asked to report involuntary memories only and no deliberate recall (e.g., recall directly prompted by the diary questions); intrusions during the night (e.g., dreams, during awakenings) were also counted. Moreover, for each intrusion, participants were instructed to record the content and to indicate whether it was experienced as visual, auditive, thoughts, or feelings; participants were allowed to indicate more than one modality per intrusion. The total number of intrusions across the 3 days was summed as an index of intrusion frequency. Intrusions were assessed via a customized e-diary application (PsyDiary) installed on participants' smartphones.

fMRI Recording

MRI data of the experimental task were acquired on a 3T system (Magnetom TrioTim syngo; Siemens, Erlangen, Germany) with a 12-channel head coil. A total of 758 volumes, aligned to the anterior and posterior commissure plane, were acquired for each session and the first 6 volumes were discarded to allow for stabilization of the blood oxygen level–dependent signal. Functional images were acquired with a T2*-weighted gradient echo-planar imaging sequence (repetition time = 2250 ms, echo time = 30 ms, matrix 64 × 64, field of view = 192 mm, flip-angle = 70°). A total of 36 slices with a slice thickness of 3 mm and a slice gap of 0.3 mm were acquired within the repetition time. In addition, a gradient echo field map (repetition time = 532 ms, echo time 1 = 5.17 ms, echo time 2 = 7.63 ms) and a high-resolution (1 × 1 × 1.2 mm) structural scan with a T1-weighted magnetization prepared rapid acquisition gradient-echo sequence were acquired from each participant. Participants viewed the films projected to a screen outside the scanner bore through a head coil–mounted mirror, and sounds were presented via noise-shielding headphones.

Data Analysis

Statistical analyses of ratings were performed using SPSS Statistics, version 21 (IBM Corp., Armonk, NY). For ratings of US expectancy and CS valence, repeated-measures analyses of variance, including CS type (CS+ or CS−) and time (early or late) as within-participant factors, were calculated for each conditioning phase (acquisition or extinction).

fMRI data preprocessing and analysis were performed using SPM12 (Wellcome Department of Cognitive Neurology, London, UK). Details can be found in the Supplement. Our primary analysis was a region-of-interest (ROI) analysis on the dACC and insula, as follows from the introductory paragraphs. Parameter estimates of ROIs were extracted with MarsBaR using the WFU PickAtlas implemented in SPM, and the dACC mask was built with the WFU PickAtlas toolbox using the procedures described by Cascio *et al.* (36). For the anterior insula, we used an online atlas of functional ROIs (37). ROI analyses were run for both early and late acquisition and extinction by entering parameter estimates (average across ROI) of the anterior insula and dACC to paired *t* tests comparing CS+ with CS–. In addition, ROI analyses were performed for the anterior insula and dACC by correlating parameter estimates of the CS+ > CS– difference during late extinction with intrusion frequency (log transformed). The alpha level for analyses was set to .05. Because left and right anterior insula activity yielded similar results (*p*s < .031), we report only mean activity for the bilateral anterior insula. Exploratory whole-brain analyses were run entering CS+ and CS– contrast images per phase (early or late acquisition; early or late extinction) into a second-level random effects model applying a flexible factorial design with the factors CS type (CS+ or CS−) and time (early or late) as within-participant factors for each conditioning phase (acquisition or extinction). The threshold for exploratory was set at $p < .001$ uncorrected, $k = 10$ for acquisition and $p < .005$ uncorrected, $k = 10$ for extinction because the goal of these analyses was to ensure that our ROI approach targeted the right regions and that we did not miss large clusters of activity.

RESULTS

Ratings/Manipulation Check

Participants reported on average 2.12 intrusions (SD = 2.60, range = 0–13) across the 3 days of assessment, with an

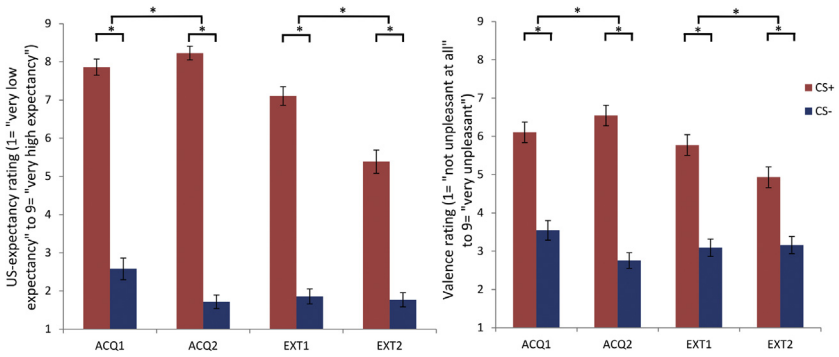

Figure 1. Unconditioned stimulus (US) expectancy rating (left) and valence rating (right) during early and late acquisition (ACQ1 and ACQ2, respectively) and early and late extinction (EXT1 and EXT2, respectively). Error bars represent standard error of the mean. *significant $p < .05$. CS, conditioned stimuli.

exponential decay pattern across days typical for trauma-film studies confirmed by a repeated-measures analysis of variance showing both significant linear effects ($F_{1,57} = 19.23$, $p < .001$) and quadratic effects ($F_{1,57} = 7.95$, $p = .007$) of time (day 1, 2, or 3) on intrusion frequency. Visual intrusions (55%) and thoughts (26%) were the most reported modalities (feelings: 12%; auditive: 7%). Depressive symptoms and trait anxiety were within the normal range (Allgemeine Depressionsskala score: mean = 9.93, SD = 5.88, range = 0–23; State-Trait Anxiety Inventory trait score: mean = 35.19, SD = 6.84, range = 25–47).

US expectancy and valence ratings demonstrated successful acquisition and extinction (see Supplement and Figure 1). However, US expectancy ($p = .591$) and valence ($p = .591$) ratings during late extinction were not correlated with intrusion frequency.

fMRI Results

ROI analysis of CS+ > CS– during acquisition showed significant effects in the anterior insula and dACC, in line with our expectations, whereas ROI analysis of CS+ > CS– during extinction showed a significant trend only in the anterior insula during late extinction (Table 1).

Exploratory whole-brain analyses entering CS+ and CS– contrast images per phase confirmed effects of the ROI analyses (Figure 2) and also revealed hippocampus, amygdala, ventral striatum, and rostral ACC (rACC) activation, among other areas, during acquisition (see Supplemental Tables S1 and S2), whereas during extinction mainly the anterior insula was active (see Supplemental Tables S10 and S11). Analyses comparing early and late phases of acquisition as well as extinction are reported in Supplemental Tables S3, S4, S12, and S13.

ROI analyses revealed that activity in the dACC ($r = .264$, $p = .041$) and anterior insula ($r = .298$, $p = .023$) during late extinction (CS+ > CS– contrast) was positively correlated with intrusion frequency (see Figure 3). At the request of the reviewers, we ran an additional exploratory whole-brain analysis adding intrusion frequency as a covariate to a *t* test of the CS+ > CS– contrast. This analysis also revealed correlations of the rACC, hippocampus, and occipital gyrus, among others, with intrusion frequency (for results, see Supplemental Tables S14 and S15). To confirm that dACC and anterior insula effects are contingent on CS responding and not US responding, we included US responding (US > US control) and CS responding (CS+ > CS– late extinction) of both regions in multiple linear regression models to predict intrusion frequency. Results for the dACC showed significant effects of both US responding ($p = .014$) and CS responding ($p = .025$), whereas anterior insula showed an effect for CS responding ($p = .023$) but not for US responding ($p = .251$).

DISCUSSION

The current study investigated whether deficient fear extinction learning predicts reexperiencing of analog trauma in the form of intrusions. Combining classical fear conditioning with trauma-film stimulation in healthy individuals, the current findings revealed anterior insula activity to trauma reminders during extinction. Moreover, individual differences in both anterior insula and dACC activity during late extinction predicted higher intrusion frequency during subsequent days. Results support the role of extinction learning in intrusive memory formation; that is, individuals who failed to uncouple conditioned emotional responding in the anterior insula and dACC from external threat cues subsequently showed more intrusive memories, further supporting the assumption that intrusions are—at least in part—CRs (12).

The trauma-film conditioning proved to be effective during acquisition, reflected by differential activity of the dACC and anterior insula as core nodes of the central autonomic–interoceptive network proposed by Fullana *et al.* (29). The authors suggested that during acquisition the anterior insula plays a crucial role in integrating awareness of one's own cognitive, affective, and physical state that is re-represented in the dACC for triggering homeostatic autonomic and behavioral responses. A recent study of our group (38) demonstrated enhanced anterior insula (including dACC and thalamus) activity during negative social evaluation (compared with neutral and positive social evaluation), providing additional evidence of anterior insula involvement during affective state integration (39). Exploratory whole-brain analysis also revealed a widespread differential activation pattern during acquisition, including the amygdala, rACC, hippocampus, and ventral striatum, in agreement with trauma-film studies linking these regions to analog intrusion formation (25–27) as well as with studies demonstrating amygdala involvement in heightened encoding and processing of emotional events (40,41). In

Table 1. ROI Analyses (Averaged Across Voxels Within ROI; Paired *t* Tests) During Early and Late Acquisition/Extinction

Contrast	Region	*t* Value	*p*	Cohen's *d*
Early Acquisition				
CS+ > CS−	dACC	4.68	<.001	.678
	R anterior insula	4.77	<.001	.715
	L anterior insula	4.60	<.001	.733
Late Acquisition				
CS+ > CS−	dACC	5.39	<.001	.799
	L anterior insula	4.90	<.001	.752
	R anterior insula	3.79	<.001	.584
Early Extinction				
CS+ > CS–	No significant effects			
Late Extinction				
CS+ > CS−	R anterior insula	1.82	.074	.275

CS, conditioned stimuli; dACC, dorsal anterior cingulate cortex; L, left; R, right; ROI, region of interest.

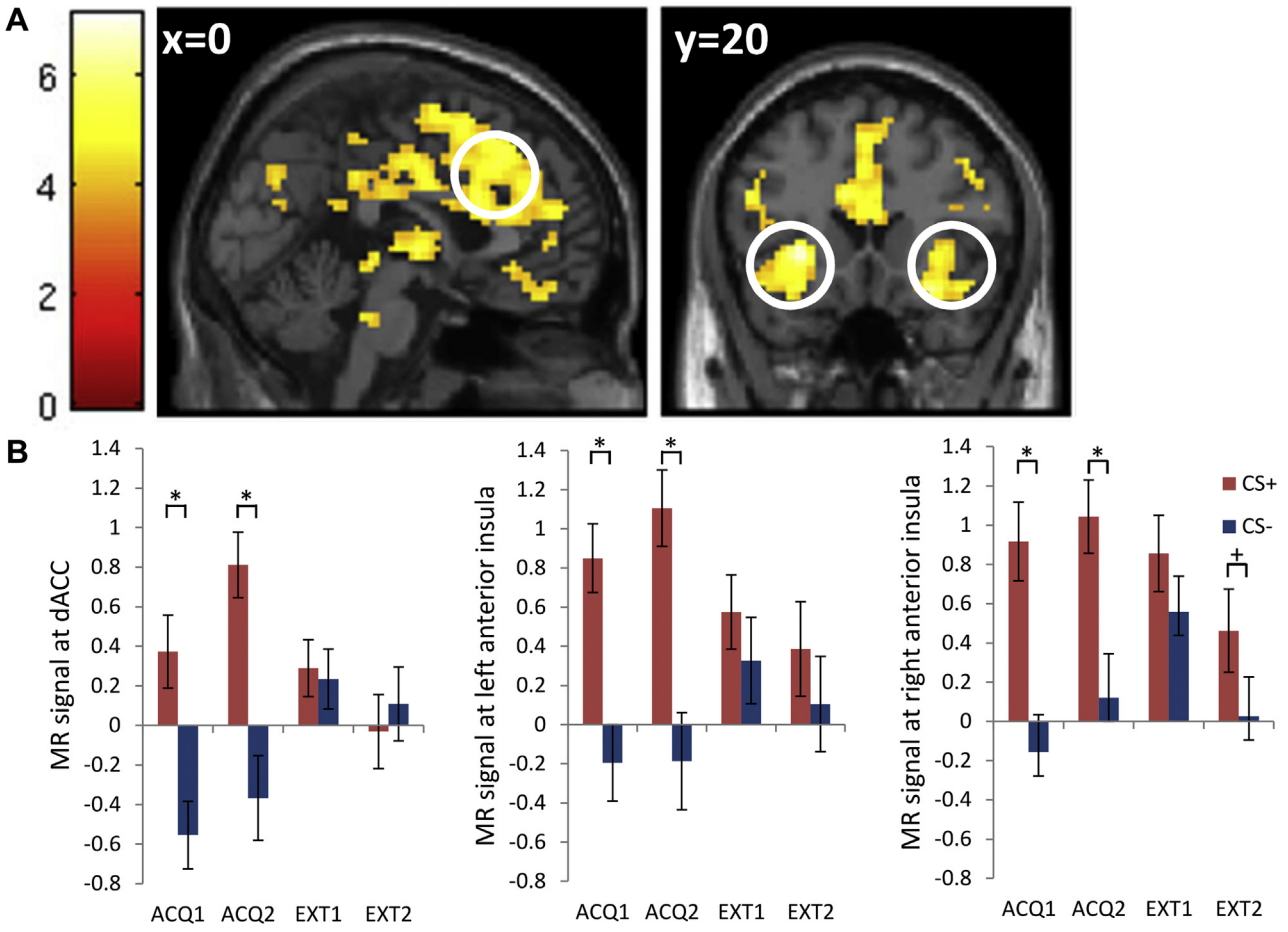

Figure 2. **(A)** An exploratory whole-brain analysis yielded similar results as the region-of-interest analysis: CS+ > CS– during late acquisition revealed widespread activity, including the dorsal anterior cingulate cortex (dACC) (left) and anterior insula (right) (display threshold: $p < .0001$ uncorrected) (see Supplemental Table S2). **(B)** Parameter estimates of the dACC and anterior insula (averaged across voxels within region of interest) during all phases of the experiment (*significant $p < .05$; +significant trend $.1 > p > .05$; see Table 1). Error bars represent standard error of the mean. ACQ1, early acquisition; ACQ2, late acquisition; CS, conditioned stimuli; EXT1, early extinction; EXT2, late extinction; MR, magnetic resonance.

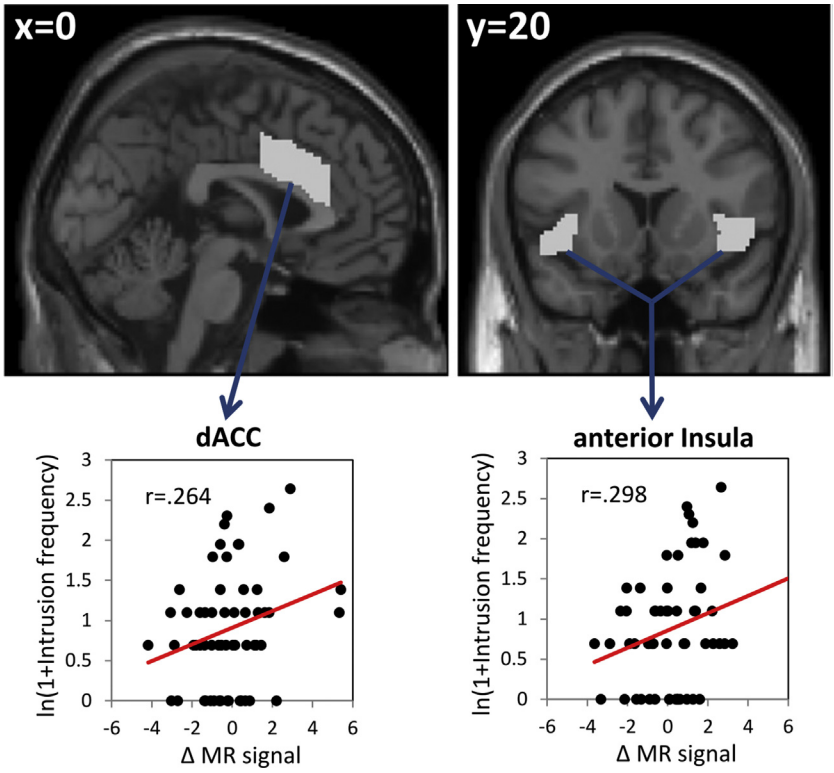

Figure 3. Positive correlation of CS+ > CS– activity (averaged across voxels within region of interest) during late extinction in the dorsal anterior cingulate cortex (dACC) (left) and bilateral anterior insula (right) with intrusion frequency. CS, conditioned stimuli; MR, magnetic resonance.

contrast, a recent meta-analysis showed no consistent amygdala involvement or hippocampus deactivations during fear acquisition (29), which could be related to differences in the design relative to our conditioning study. While fMRI trauma-film studies consistently demonstrated amygdala and hippocampus activity (25–27,42), the current design combined trauma films as the US with classical fear conditioning, where participants probably transferred emotional responding to the complex trauma films to the CS+, which might be less pronounced and/or qualitatively different in designs using, for example, electric shock as the US.

During extinction, differential CS+ > CS– activity was predominantly observed in the anterior insula. Anterior insula activation during extinction could reflect persistent learned threat-cue reactivity surviving from acquisition and is well in agreement with studies showing central autonomic–interoceptive network activity during fear acquisition (29) as well as fear extinction (30) and is also in agreement with a meta-analysis demonstrating insula hyperactivity during extinction learning in patients with PTSD (28). The absence of differential amygdala activation during extinction in the current study could point to a swift intensity reduction of CS+ response during extinction (43) that could have been enhanced by a procedural shift (i.e., reinforcement of both CS+ and CS– by film clips) during acquisition, followed by complete film US absence during extinction. Relatedly, Fullana *et al.* demonstrated in meta-analyses a lack of robust amygdala involvement during human fear extinction (30). Furthermore, it has been shown that the amygdala was activated during both negative and positive social evaluation processing (38), suggesting broader amygdala involvement including relevance detection (44) rather than exclusively fear processing. Therefore, differential amygdala activation during extinction could have been masked by predominantly signaling the absence of both aversive and neutral film clips irrespective of their inherent valence—which, however, appeared to still be discriminable in the anterior insula. In addition, the VMPFC showed no main effect during extinction, in line with a recent meta-analysis (30).

The observed positive correlation between enhanced differential activity in the anterior insula and dACC during late extinction with intrusion frequency provides a link to previous results of deficient extinction learning in patients with PTSD (9,45) as well as recent fMRI studies on intrusion formation using the trauma-film paradigm (26,27) and current neural models of PTSD (20,21,23). It might reflect intrusion-predictive hyperactivity of anterior insula, also reported to be involved in the integration of the perceived threat value of stimuli (46), enhanced empathy processing (47), and affective interoceptive prediction signaling (48) of homeostatically relevant stimuli (49,50). Importantly, dACC activity in combination with anterior insula activity strengthens the role of these core areas affected

across most categories of psychiatric illness (31) and points to involvement of central autonomic–interoceptive processing (29) in intrusion formation.

The anterior insula was the only region activated during late extinction in the CS+ > CS– contrast, but the dACC, without showing a main effect in the same contrast, was positively correlated with intrusion frequency. This effect is partly driven by deactivations in the CS+ > CS– contrast, canceling out activity in the same contrast over participants in the dACC. In general, a significant CS+ > CS– main effect for a specific brain region is not necessarily a prerequisite for demonstrating a meaningful predictive relationship of this region with a behavioral variable (here intrusion frequency); between-individual variance in (de) activation and subthreshold task-related activity can be highly informative for detecting individual vulnerability markers even in the absence of task main effects. Exploratory whole-brain analysis also revealed increased activity in the hippocampus and rACC, extending to the VMPFC and occipital/parietal and temporal regions during late extinction, being linked to heightened intrusion frequency. This is in line with research showing that hippocampal hyperactivity is linked to PTSD symptoms, probably reflecting enhanced processing of arousing memories (51,52). rACC hyperactivity related to intrusion frequency is in line with results of Bourne *et al.* (26), demonstrating heightened rACC activity to intrusion-eliciting versus non-intrusion-eliciting film scenes. Relatedly, rACC hyperactivity has been reported in PTSD (53,54) and might be associated with increased attention to salient information (49,55,56). Although VMPFC hyporesponsiveness in PTSD (57) has been linked to deficient emotion regulation, the current study indicates a positive correlation between VMPFC activity and subsequent intrusion frequency, which in combination with the rACC could reflect unsuccessful regulation (58).

Importantly, dACC and anterior insula effects on intrusion frequency were independent of US (i.e., trauma-film) responding, which further strengthens the argument of a particular role of deficient extinction learning in intrusion formation. Moreover, occipital activity in the CS+ > CS– contrast during late extinction correlating with intrusion frequency could be related to the fact that visual reexperiencing (55%) was the most reported intrusion modality. Unexpectedly, although differential US expectancy and valence ratings proved successful acquisition and extinction learning, they were unrelated to intrusion frequency. It appears that explicit ratings could not capture the fast and automatic dACC and anterior insula responding during late extinction. Stimulus generalization processes probably explain how non-extinguished CS+ responses lead to subsequent intrusions (16,59); although the exact CS+ face is, of course, never again encountered in daily life, a variety of facial features similar to the CS+ may have triggered intrusions in daily life—probably driven by compromised selectivity of threat detection in the anterior insula (60). Despite the obvious complexity of neurocognitive processes involved, our novel, naturalistic experimental task was able to capture neural processes relating to the conditioning origin of intrusive memories, with relevance for better understanding how peritraumatic and posttraumatic neural processing may result in PTSD in some individuals.

The current results expand previous findings of behavioral and physiological conditioned responses predicting intrusions (16,19) on a neural level. It could be suggested that enhanced autonomic–interoceptive responding during extinction not only reflects deficient extinction learning (7,16) but also represents a pervasive effect of enhanced sensation-based episodic memory representation in response to cues (20), delineating a PTSD vulnerability mechanism that is strengthening subsequent intrusive memory formation. Future studies should also include men because it has been demonstrated that men respond to aversive film clips differently than women (61) and that this may mediate differences in intrusion formation (19). Because the current sample included naturally cycling women (n = 27) as well as women taking hormonal contraceptives of different classes (n = 31), we cannot be sure whether the effects demonstrated here will generalize to all gonadal hormone status subgroups of women (different cycle phases combined vs. progestogen-only oral contraceptive use).

Conclusions

Our results demonstrate that in healthy women, stronger differential anterior insula and dACC activity during extinction predicted the subsequent occurrence of intrusive memories. This extends findings in PTSD showing hyperactivity in these regions by employing an experimental analog etiology model of PTSD, possibly providing a first glimpse into peritraumatic and early posttraumatic neural processes during PTSD symptom development. It appears that particularly the sustained conditioned responding of core nodes of the central autonomic–interoceptive network to external threat cues not associated with actual threat anymore was predictive of intrusions. This type of dysfunctional neural processing in response to trauma cues might represent a vulnerability factor for the development of persistent intrusions after traumatic incidents.

ACKNOWLEDGMENTS AND DISCLOSURES

Julina A. Rattel was financially supported by the doctoral program "Imaging the Mind" of the Austrian Science Fund (Grant No. W1233-G17 [to FWF; principal investigator: Frank H. Wilhelm]).

We thank Lisa M. Grünberger for her help with data acquisition.

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Clinical Stress and Emotion Laboratory (SFM, JAR, LKF, JB, FHW), Division of Clinical Psychology, Psychotherapy, and Health Psychology, Department of Psychology, University of Salzburg; Centre for Cognitive Neuroscience (JB, MK), University of Salzburg; and Neuroscience Institute (MK), Christian Doppler Clinic, Paracelsus Medical University Salzburg, Salzburg, Austria; and Department of Translational Research in Psychiatry (VIS), Max Planck Institute of Psychiatry, Munich, Germany.

Address correspondence to Stephan Miedl, Ph.D., Division of Clinical Psychology, Psychotherapy, and Health Psychology, University of Salzburg, 5020 Salzburg, Austria; E-mail: stephan.miedl@sbg.ac.at.

Received Sep 20, 2019; revised Dec 12, 2019; accepted Dec 23, 2019.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsc.2019.12.017.

REFERENCES

1. Kilpatrick DG, Resnick HS, Milanak ME, Miller MW, Keyes KM, Friedman MJ (2013): National estimates of exposure to traumatic events and PTSD prevalence using DSM-IV and DSM-5 criteria. J Trauma Stress 26:537–547.
2. American Psychiatric Association (2013): Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Arlington, VA: American Psychiatric Publishing.
3. Bonanno GA, Diminich ED (2013): Annual research review: Positive adjustment to adversity-trajectories of minimal-impact resilience and emergent resilience. J Child Psychol Psychiatry 54:378–401.
4. Norris FH (1992): Epidemiology of trauma: Frequency and impact of different potentially traumatic events on different demographic groups. J Consult Clin Psychol 60:409–418.
5. Ehlers A, Clark DM (2000): A cognitive model of posttraumatic stress disorder. Behav Res Ther 38:319–345.
6. Ross DA, Arbuckle MR, Travis MJ, Dwyer JB, van Schalkwyk GI, Ressler KJ (2017): An integrated neuroscience perspective on formulation and treatment planning for posttraumatic stress disorder: An educational review. JAMA Psychiatry 74:407–415.
7. Orr SP, Metzger LJ, Lasko NB, Macklin ML, Peri T, Pitman RK (2000): De novo conditioning in trauma-exposed individuals with and without posttraumatic stress disorder. J Abnorm Psychol 109:290–298.
8. Pittig A, Treanor M, LeBeau RT, Craske MG (2018): The role of associative fear and avoidance learning in anxiety disorders: Gaps and directions for future research. Neurosci Biobehav Rev 88:117–140.
9. Blechert J, Michael T, Vriends N, Margraf J, Wilhelm FH (2007): Fear conditioning in posttraumatic stress disorder: Evidence for delayed extinction of autonomic, experiential, and behavioural responses. Behav Res Ther 45:2019–2033.
10. Rothbaum BO, Davis M (2003): Applying learning principles to the treatment of post-trauma reactions. Ann N Y Acad Sci 1008:112–121.
11. Graham BM, Milad MR (2011): The study of fear extinction: Implications for anxiety disorders. Am J Psychiatry 168:1255–1265.
12. Duits P, Cath DC, Lissek S, Hox JJ, Hamm AO, Engelhard IM, *et al.* (2015): Updated meta-analysis of classical fear conditioning in the anxiety disorders. Depress Anxiety 32:239–253.
13. Lommen MJJ, Engelhard IM, Sijbrandij M, van den Hout MA, Hermans D (2013): Pre-trauma individual differences in extinction learning predict posttraumatic stress. Behav Res Ther 51:63–67.
14. Holmes EA, Bourne C (2008): Inducing and modulating intrusive emotional memories: A review of the trauma film paradigm. Acta Psychol 127:553–566.
15. Lissek S, Powers AS, McClure EB, Phelps EA, Woldehawariat G, Grillon C, *et al.* (2005): Classical fear conditioning in the anxiety disorders: A meta-analysis. Behav Res Ther 43:1391–1424.
16. Wegerer M, Blechert J, Kerschbaum H, Wilhelm FH (2013): Relationship between fear conditionability and aversive memories: Evidence from a novel conditioned-intrusion paradigm. PLoS One 8: e79025.
17. Wegerer M, Blechert J, Wilhelm FH (2013): Emotional learning: A naturalistic experimental paradigm for investigating fear acquisition and extinction using aversive films. Z Psychiatr Psychol Psychother 61:93–103.
18. Wegerer M, Kerschbaum H, Blechert J, Wilhelm FH (2014): Low levels of estradiol are associated with elevated conditioned responding during fear extinction and with intrusive memories in daily life. Neurobiol Learn Mem 116:145–154.
19. Rattel JA, Wegerer M, Miedl SF, Blechert J, Grünberger LM, Craske MG, *et al.* (2019): Peritraumatic unconditioned and conditioned responding explains sex differences in intrusions after analogue trauma. Behav Res Ther 116:19–29.
20. Brewin CR, Gregory JD, Lipton M, Burgess N (2010): Intrusive images in psychological disorders: Characteristics, neural mechanisms, and treatment implications. Psychol Rev 117:210–232.
21. Patel R, Spreng RN, Shin LM, Girard TA (2012): Neurocircuitry models of posttraumatic stress disorder and beyond: A meta-analysis of functional neuroimaging studies. Neurosci Biobehav Rev 36:2130–2142.
22. Pitman RK, Rasmusson AM, Koenen KC, Shin LM, Orr SP, Gilbertson MW, *et al.* (2012): Biological studies of post-traumatic stress disorder. Nat Rev Neurosci 13:769–787.
23. Rauch SL, Shin LM, Phelps EA (2006): Neurocircuitry models of posttraumatic stress disorder and extinction: Human neuroimaging research—Past, present, and future. Biol Psychiatry 60:376–382.
24. Stark EA, Parsons C, Van Hartevelt T, Charquero-Ballester M, McManners H, Ehlers A, *et al.* (2015): Post-traumatic stress influences the brain even in the absence of symptoms: A systematic, quantitative meta-analysis of neuroimaging studies. Neurosci Biobehav Rev 56:207–221.
25. Battaglini E, Liddell B, Das P, Malhi G, Felmingham K, Bryant RA (2016): Intrusive memories of distressing information: An fMRI study. PLoS One 11:e140871.
26. Bourne C, Mackay CE, Holmes EA (2013): The neural basis of flashback formation: The impact of viewing trauma. Psychol Med 43:1521–1532.
27. Clark IA, Holmes EA, Woolrich MW, Mackay CE (2016): Intrusive memories to traumatic footage: The neural basis of their encoding and involuntary recall. Psychol Med 46:505–518.
28. Suarez-Jimenez B, Albajes-Eizagirre A, Lazarov A, Zhu X, Harrison BJ, Radua J, *et al.* (2019): Neural signatures of conditioning, extinction learning, and extinction recall in posttraumatic stress disorder: A meta-analysis of functional magnetic resonance imaging studies [published online ahead of print Jul 1]. Psychol Med.
29. Fullana M, Harrison B, Soriano-Mas C, Vervliet B, Cardoner N, Avila-Parcet A, *et al.* (2016): Neural signatures of human fear conditioning: An updated and extended meta-analysis of fMRI studies. Mol Psychiatry 21:500–508.
30. Fullana MA, Albajes-Eizagirre A, Soriano-Mas C, Vervliet B, Cardoner N, Benet O, *et al.* (2018): Fear extinction in the human brain: A meta-analysis of fMRI studies in healthy participants. Neurosci Biobehav Rev 88:16–25.
31. Downar J, Blumberger DM, Daskalakis ZJ (2016): The neural crossroads of psychiatric illness: An emerging target for brain stimulation. Trends Cogn Sci 20:107–120.
32. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ (1989): The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. Psychiatry Res 28:193–213.
33. Laux L, Glanzmann P, Schaffner P, Spielberger CD (1981): State-Trait-Angstinventar STAI: Manual. Göttingen, Germany: Hogrefe.
34. Hautzinger M, Bailer M (1993): ADS: Allgemeine Depressionsskala. Göttingen, Germany: Hogrefe Verlag.
35. Langner O, Dotsch R, Bijlstra G, Wigboldus DHJ, Hawk ST, van Knippenberg A (2010): Presentation and validation of the Radboud Faces Database. Cogn Emot 24:1377–1388.
36. Cascio CN, Konrath SH, Falk EB (2015): Narcissists' social pain seen only in the brain. Soc Cogn Affect Neurosci 10:335–341.
37. Shirer WR, Ryali S, Rykhlevskaia E, Menon V, Greicius MD (2012): Decoding subject-driven cognitive states with whole-brain connectivity patterns. Cereb Cortex 22:158–165.
38. Miedl SF, Blechert J, Klackl J, Wiggert N, Reichenberger J, Derntl B, *et al.* (2016): Criticism hurts everybody, praise only some: Common and specific neural responses to approving and disapproving social-evaluative videos. NeuroImage 132:138–147.
39. Gu X, Hof PR, Friston KJ, Fan J (2013): Anterior insular cortex and emotional awareness. J Comp Neurol 521:3371–3388.
40. Phelps EA, LeDoux JE (2005): Contributions of the amygdala to emotion processing: From animal models to human behavior. Neuron 48:175–187.
41. Pessoa L, Adolphs R (2010): Emotion processing and the amygdala: From a "low road" to "many roads" of evaluating biological significance. Nat Rev Neurosci 11:773–783.
42. Miedl SF, Wegerer M, Kerschbaum H, Blechert J, Wilhelm FH (2018): Neural activity during traumatic film viewing is linked to endogenous estradiol and hormonal contraception. Psychoneuroendocrinology 87:20–26.

43. LaBar KS, Gatenby JC, Gore JC, LeDoux JE, Phelps EA (1998): Human amygdala activation during conditioned fear acquisition and extinction: A mixed-trial fMRI study. Neuron 20:937–945.
44. Sander D, Grafman J, Zalla T (2003): The human amygdala: An evolved system for relevance detection. Rev Neurosci 14:303–316.
45. Norrholm SD, Jovanovic T, Olin IW, Sands LA, Karapanou I, Bradley B, *et al.* (2011): Fear extinction in traumatized civilians with posttraumatic stress disorder: Relation to symptom severity. Biol Psychiatry 69: 556–563.
46. Wiech K, Lin CS, Brodersen KH, Bingel U, Ploner M, Tracey I (2010): Anterior insula integrates information about salience into perceptual decisions about pain. J Neurosci 30:16324–16331.
47. Singer T, Seymour B, O'Doherty J, Kaube H, Dolan RJ, Frith CD (2004): Empathy for pain involves the affective but not sensory components of pain. Science 303:1157–1162.
48. Paulus MP, Stein MB (2006): An insular view of anxiety. Biol Psychiatry 60:383–387.
49. Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, *et al.* (2007): Dissociable intrinsic connectivity networks for salience processing and executive control. J Neurosci 27:2349–2356.
50. Lovero KL, Simmons AN, Aron JL, Paulus MP (2009): Anterior insular cortex anticipates impending stimulus significance. NeuroImage 45:976–983.
51. Shin LM, Shin PS, Heckers S, Krangel TS, Macklin ML, Orr SP, *et al.* (2004): Hippocampal function in posttraumatic stress disorder. Hippocampus 14:292–300.
52. Werner NS, Meindl T, Engel RR, Rosner R, Riedel M, Reiser M, *et al.* (2009): Hippocampal function during associative learning in patients with posttraumatic stress disorder. J Psychiatr Res 43:309–318.
53. Bryant RA, Felmingham KL, Kemp AH, Barton M, Peduto AS, Rennie C, *et al.* (2005): Neural networks of information processing in posttraumatic stress disorder: A functional magnetic resonance imaging study. Biol Psychiatry 58:111–118.
54. Felmingham KL, Williams LM, Kemp AH, Rennie C, Gordon E, Bryant RA (2009): Anterior cingulate activity to salient stimuli is modulated by autonomic arousal in posttraumatic stress disorder. Psychiatry Res 173:59–62.
55. Bishop S, Duncan J, Brett M, Lawrence AD (2004): Prefrontal cortical function and anxiety: Controlling attention to threat-related stimuli. Nat Neurosci 7:184–188.
56. Whalen PJ, Bush G, McNally RJ, Wilhelm S, McInerney SC, Jenike MA, *et al.* (1998): The emotional counting Stroop paradigm: A functional magnetic resonance imaging probe of the anterior cingulate affective division. Biol Psychiatry 44:1219–1228.
57. Etkin A, Wager TD (2007): Functional neuroimaging of anxiety: A meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. Am J Psychiatry 164:1476–1488.
58. Etkin A, Egner T, Kalisch R (2011): Emotional processing in anterior cingulate and medial prefrontal cortex. Trends Cogn Sci 15:85–93.
59. Lopresto D, Schipper P, Homberg JR (2016): Neural circuits and mechanisms involved in fear generalization: Implications for the pathophysiology and treatment of posttraumatic stress disorder. Neurosci Biobehav Rev 60:31–42.
60. Onat S, Büchel C (2015): The neuronal basis of fear generalization in humans. Nat Neurosci 18:1811–1818.
61. Wilhelm FH, Rattel JA, Wegerer M, Liedlgruber M, Schweighofer S, Kreibig SD, *et al.* (2017): Attend or defend? Sex differences in behavioral, autonomic, and respiratory response patterns to emotion-eliciting films. Biol Psychol 130:30–40.