Received: 6 June 2019 | Revised: 21 November 2019 | Accepted: 11 December 2019

DOI: 10.1002/hbm.24915

RESEARCH ARTICLE

WILEY

Amygdala and nucleus accumbens involvement in appetitive extinction

Onno Kruse[1,2] | Sanja Klein[2,3] | Isabell Tapia León[1,2] | Rudolf Stark[2,3] | Tim Klucken[1,2]

[1]Department of Clinical Psychology, University Siegen, Siegen, Germany

[2]Bender Institute for Neuroimaging (BION), Justus Liebig University Giessen, Giessen, Germany

[3]Department of Psychotherapy and Systems Neuroscience, Justus Liebig University Giessen, Giessen, Germany

Correspondence
Dr. Onno Kruse, Department of Psychotherapy and Systems Neuroscience, Justus Liebig University Giessen, Otto-Behaghel-Strasse 10H, 35394 Giessen, Germany.
Email: onno.kruse@psychol.uni-giessen.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: KL2500/5-1

Abstract

Extinction of appetitive conditioning is regarded as an important model for the treatment of psychiatric disorders like addiction. However, very few studies have investigated its neural correlates. Therefore, we investigated neural correlates of appetitive extinction in a large human sample including all genders (N = 76, 40 females) to replicate and extend results from a previous study. During differential appetitive conditioning, one stimulus (CS+) was paired with the chance to win a monetary reward, whereas another stimulus (CS−) was not. During appetitive extinction on the next day, neither the CS+ nor the CS− were reinforced. After successful acquisition of appetitive conditioning, the extinction phase elicited significant reductions of valence and arousal ratings toward the CS+ and a significant reduction in skin conductance responses to the CS+ from early to late extinction. On a neural level, early extinction showed significant differential (CS+ − CS−) activation in dACC and hippocampus, whereas involvement of the vACC and caudate nucleus did not replicate. The differential activation of amygdala and nucleus accumbens during late extinction was replicated, with the amygdala displaying significantly higher differential activation during the late phase of extinction as compared to the early phase of extinction. We show discernible signals for reward learning and extinction in subregions of amygdala and nucleus accumbens after extinction learning. This successful replication underlines the role of nucleus accumbens and amygdala in neural models of appetitive extinction in humans that was previously only based on animal findings.

KEYWORDS

amygdala, conditioning, fMRI, nucleus accumbens, reward

1 | INTRODUCTION

Appetitive extinction is regarded as a key mechanism for the treatment of psychiatric disorders like addiction (Millan, Marchant, & McNally, 2011). During extinction training, a stimulus (CS+; e.g., blue rectangle) associated with a reward (unconditioned stimulus or UCS; e.g., money) during an acquisition training is no longer paired with that reward. Subjects are then assumed to form an extinction memory trace that inhibits the CS-UCS association (Quirk & Mueller, 2008). However, studies of its neural correlates in humans are scarce (Konova & Goldstein, 2019). In a previous study, we identified the nucleus accumbens (NAcc) and the amygdala as key structures

involved in extinction learning when it has taken place (Kruse, Tapia León, Stark, & Klucken, 2017). This study was the first to investigate main effects of appetitive extinction with functional magnetic resonance imaging (fMRI); therefore, its findings need to be replicated in an independent sample to build a neural model of appetitive extinction and to further study neural mechanisms and moderators.

Acquisition of appetitive conditioning is investigated by repeatedly pairing one neutral stimulus (CS+; e.g., blue rectangle) with the chance to win a reward (UCS; e.g., money). Another neutral stimulus (CS−; e.g., yellow rectangle) is never paired with reward. After few pairings, presenting the CS+ elicits conditioned responses as compared to the CS−. These include higher subjective ratings of valence and arousal, increased skin conductance responses (SCRs), and increased BOLD responses in brain areas associated with conditioning and reward processing (Andreatta & Pauli, 2015; Tapia León, Kruse, Stalder, Stark, & Klucken, 2018).

Brain regions associated with appetitive conditioning mainly include the amygdala, nucleus accumbens (NAcc), dorsal and ventral anterior cingulate cortex (dACC/vACC), orbitofrontal cortex (OFC), and caudate nucleus (Chase, Kumar, Eickhoff, & Dombrovski, 2015; Klucken, Tabbert et al., 2009; Klucken, Wehrum et al., 2013; Martin-Soelch, Linthicum, & Ernst, 2007). In the context of acquisition of appetitive conditioning, the amygdala is thought to encode the association of the CS with the unconditioned reaction, whereas the NAcc has been associated with subjective CS/UCS-association and the reward prediction error (O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003; Tapia León et al., 2018). Within the anterior cingulate cortex, the vACC is thought to mediate early discriminatory learning in contrast to the dACC, which is thought to encode outcome expectancy (Alexander & Brown, 2011). The OFC has been linked to the subjective value of the CS and was found to be a key region in reversal learning (Finger, Mitchell, Jones, & Blair, 2008). The caudate as part of the dorsal striatum has been associated with instrumental reward learning (O'Doherty et al., 2004).

In the context of extinction learning, central roles have been ascribed to the amygdala, NAcc, and vACC (or ventromedial prefrontal cortex, depending on the exact location), whereas the dACC is regarded a key region in the recall of conditioning (Konova & Goldstein, 2019). In our first study, we found an involvement of the vACC, caudate nucleus, and hippocampus during the early phase of extinction, whereas during the late phase of extinction, amygdala and NAcc were involved (Kruse et al., 2017). As mentioned before, the brain areas associated with extinction learning are also involved in the acquisition of appetitive conditioning. Animal studies using multi-cell-recordings have shown that although some neurons in amygdala and NAcc fire during anticipation of conditioned reward and cease to fire after extinction, other populations of neurons in these regions begin to fire with successful extinction learning (Janak, Chen, & Caulder, 2004; Tye, Cone, Schairer, & Janak, 2010).

The aim of the present study is to replicate and extend the previous results in a large, independent sample of all genders, as opposed to the smaller sample of male subjects in the previous study. During the early phase of extinction, we assumed increased activation of vACC and dACC to the CS+ as compared to the CS−. In the late phase, we assumed increased activation in NAcc and amygdala to the CS+ as compared to the CS−. In an extension of the previous results, we also tested for increased differential activation in the late phase as compared to differential activation during the early phase, utilizing the increased power of the greater sample.

2 | MATERIALS AND METHODS

2.1 | Subjects

The final sample included in the analyses consisted of 76 subjects (40 female, 36 male; age: M = 23.76, S.D. = 3.73). A total of 90 subjects took part, 14 subjects were excluded because (a) they could not correctly name the CS+ immediately before extinction in a free recall question which colored rectangle had been associated with a chance to win a reward on the previous day (n = 6), (b) they expected rewards after the CS− as well as indicated in equally high UCS-expectancy ratings immediately following the acquisition phase (n = 1), (c) technical difficulties during data collection (n = 2), (d) anomalies in the fMRI-data indicated by more than 10% outlying volumes (n = 2), and (e) reported familiarity with the paradigm after finishing the study (n = 3). All subjects were right-handed, German native speakers and had normal or corrected-to-normal vision. Past or current mental or neurological problems, consumption of psychotropic drugs, chronic illnesses or treatments and conditions preventing them from entering the MRI scanner were exclusion criteria. There was no overlap between this sample and the data reported in Kruse et al. (2017). All subjects gave written informed consent. The study was conducted in accordance with the Declaration of Helsinki (2008) and approved by the local ethics committee.

2.2 | Procedure

The appetitive conditioning paradigm took place on two consecutive days with acquisition on the first day and extinction on the second, roughly about 24 hr later. Data collection took place throughout the day between 8 a.m. and 7 p.m.

2.2.1 | Acquisition training

We used an adapted version of the monetary incentive delay task (MID; Knutson, Westdorp, Kaiser, & Hommer, 2000) as a conditioning procedure in the MRI. The same paradigm has been used before and is described in detail in Kruse et al. (2017). In short, it consisted of 21 CS+ with partial reinforcement (62% reinforcement) and 21 CS− trials. Subjects were not instructed about the explicit CS-UCS contingencies. The first trials of the CS+ and the CS− were later excluded from the analyses because learning could not have taken place yet (Kruse, Tapia León, Stalder, Stark, & Klucken, 2018). In each trial, a CS

+ or CS− (blue or yellow rectangle) was presented for 6 s, followed by a fixation cross for a variable (1–3 s) interstimulus interval. Then, a target (white square) was presented for at least 16 ms up to a maximum of 750 ms. Subjects were instructed to press a reaction button every time the target was presented regardless of the CS presented before. Pressing the reaction button following a CS+ while the target was visible always resulted in a win of 0.50€ (UCS). Directly following the target, feedback was presented for 2 s. The presentation time of the target was adjusted according to individual reaction times to ensure a similar reinforcement for all subjects (aim: 6.50€ for wins in 62% of CS + trials). Individual mean reaction times and standard deviations used to calculate presentation times (win: M_{RT} + 2x SD_{RT}; loss: M_{RT} − 2x SD_{RT}) were determined from a practice session with different stimuli. If subjects won unplanned or did not win in scheduled reinforcement trials, the target presentation time was corrected online (subtracting or adding 20 ms to the presentation time, respectively) to ensure reinforcement as planned in future trials. CS+ trials that did not result in wins as planned or vice versa were adaptively repeated in scheduled CS+ trials with the according duration of target presentation.

2.2.2 | Extinction

On the day after acquisition training, the extinction was conducted in the MRI scanner. Subjects were asked to freely recall the reinforced stimulus ("Which colored rectangle was followed by the chance to win a reward?"), and the collected data were excluded from this analysis, if they failed. The extinction consisted of 20 CS+ and 20 CS− trials. As before, subjects were instructed to always press the button when they saw the target but in contrast to the acquisition phase, this never resulted in winning money, regardless of how fast subjects reacted.

2.3 | Subjective ratings

Subjects completed ratings of the CS+ and CS− on three scales: arousal, valence, and UCS-expectancy. Rating collection took place before and after acquisition training as well as after extinction training. Nine-point self-assessment manikin scales were used for the affective ratings of arousal and valence (Bradley & Lang, 1994), whereas UCS-expectancy was rated in 10% steps from 0 to 100%. Ratings were analyzed in 2 (CS: CS+, CS−) × 3 (time: pre-acquisition, post-acquisition, post-extinction) × 2 (gender: male, female) analyses of variance (ANOVA) using SPSS 23. Significant interactions were followed up with paired *t*-tests, which were corrected for multiple comparisons using Bonferroni-correction (α = 0.05).

2.4 | Skin conductance measuring

Skin conductance was measured during the acquisition and extinction training with reusable Ag/AgCl electrodes with 13/8 mm outer/inner diameter filled with isotonic (0.05 M NaCl) electrolyte medium placed proximal and distal on the hypothenar eminence on the non-dominant left hand. Data were collected with a sampling rate of 1 kHz. For preprocessing and data analysis, Ledalab 3.4.4 was used (Benedek & Kaernbach, 2010). First, the data was downsampled to 100 Hz and smoothed with a 32 sample FWHM Gaussian kernel. As each picture was presented for 6 s, the time window from 1 to 6 s was defined as analysis window (entire interval response; Pineles, Orr, & Orr, 2009). The extracted response was defined as the largest difference between a maximum and the minimum that directly preceded it. The preceding minimum had to be within the analysis window for the response to be counted. Responses smaller than 0.01 μS were considered zero responses. All maximum responses were log(μS + 1) transformed to correct for violation of normal distribution of the data.

Mean SCRs for CS+ and CS− were calculated subsequently. Skin conductance data were analyzed in separate 2 (CS: CS+, CS−) × 2 (time: early phase, late phase) × 2 (gender: male, female) ANOVAs for acquisition and extinction training. Significant interactions were followed up with paired *t*-tests and corrected for multiple comparisons, using Bonferroni-correction (α = 0.05).

2.5 | Functional magnetic resonance imaging

All MRI images were acquired using a 3 Tesla whole-body tomograph (Siemens Prisma, Siemens Healthineers, Erlangen) with a 64-channel head coil. The structural images consisted of 176 T1-weighted sagittal slices (slice thickness 0.9 mm; FoV = 240 mm; TR = 1.58 s; TE = 2.3 s). For the functional images, a total of 440 images was acquired with a T2*-weighted gradient echo-planar imaging (EPI) with 36 slices covering the whole brain (voxel size = 3 × 3 × 3.5 mm; gap = 0.77 mm; descending slice acquisition; TR = 2 s; TE = 30 ms; flip angle = 75; FoV = 192 × 192 mm; matrix size = 64 × 64; GRAPPA = 2; phase encoding direction: anterior–posterior). The field of view was positioned automatically relative to the AC-PC line with an orientation of −30°. Preprocessing, first and second level analysis were conducted using SPM 12 (Wellcome Department of Cognitive Neurology, 2014) implemented in Matlab (The MathWorks Inc., 2012).

For preprocessing, all EPI images were coregistered to an EPI template, realigned, and unwarped using field maps recorded directly before the EPI images, slice time corrected, normalized to MNI standard space via segmentation of the structural T1-image coregistered to a T1-template and smoothed with a Gaussian kernel at 6 mm FWHM. Functional data were analyzed for outlying volumes using a distribution free approach for skewed data (Schweckendiek et al., 2013). For this approach, after realignment each volume is compared with the preceding and following volume in order to calculate deviation scores. Deviation scores are compared against a threshold calculated based on all collected data, to identify outliers. If more than 10% of volumes of a time series were marked as outliers, the time series was discarded from analysis. Each resulting outlying volume was later modeled within the general linear model as a separate regressor of no interest.

TABLE 1 Mean (SD) subjective ratings of CS+ and CS−

		Pre-acquisition		Post-acquisition		Post-extinction	
Arousal	CS+	3.45	(1.98)	5.91	(2.14)*,†	4.16	(1.63)*,†
	CS−	3.47	(1.96)	3.30	(2.03)	3.39	(1.80)
Valence	CS+	6.03	(1.80)	6.79	(1.81)*,†	4.63	(1.73)*
	CS−	5.74	(2.05)	4.93	(1.84)*	4.71	(1.82)
UCS-expectancy	CS+	6.04	(1.72)	8.99	(1.89)*,†	1.63	(1.63)*,†
	CS−	5.99	(1.51)	1.13	(0.62)*	1.36	(1.02)

*indicates a significant difference to the mean rating of the same CS at the previous time point ($p < 0.05$).; †indicates a significant difference to the mean rating of the CS− at the same time point ($p < 0.05$).

TABLE 2 Main effects and interaction effects from 2 (CS: CS+, CS−) × 3 (time: pre-acquisition, post-acquisition, post-extinction) × 2 (gender: female, male) ANOVA for subjective ratings of arousal, valence, and UCS-expectancy with *F*-value, *p*-value. **$p < 0.01$; ***$p < 0.001$

	Effect	*F*-value	*p-value*
Arousal	CS	31.05	<0.001***
	Time	20.08	<0.001***
	Gender	1.60	0.210
	CS × time	23.78	<0.001***
	CS × gender	0.01	0.907
	Time × gender	1.70	0.191
	CS × time × gender	0.53	0.467
Valence	CS	12.67	0.001**
	Time	28.87	<0.001***
	Gender	0.26	0.609
	CS × time	12.10	<0.001***
	CS × gender	0.27	0.601
	Time × gender	0.38	0.684
	CS × time × gender	0.42	0.610
UCS-expectancy	CS	546.55	<0.001***
	Time	441.55	<0.001***
	Gender	0.18	0.666
	CS × time	317.83	<0.001***
	CS × gender	0.001	0.978
	Time × gender	0.51	0.603
	CS × time × gender	1.49	0.232

For the acquisition phase, the CS+, CS−, UCS+ (win feedback following a CS+), NoUCS+ (no win feedback following a CS+), and UCS− (no win feedback following a CS−) were modeled as regressors of interest. Although the UCS+ models the feedback that money was won after the CS+, NoUCS+ and UCS− model the feedback that no money is won after CS+ or CS−, respectively. The first CS+ and the first CS− were modeled separately as learning could not have taken place at that time. For extinction training, the regressors were similar to acquisition, excluding the UCS+ because the CS+ was no longer reinforced and not modeling the first CS separately.

TABLE 3 Main effects and interaction effects from 2 (CS: CS+, CS−) × 2 (time: early phase, late phase) × 2 (gender: female, male) ANOVA for skin conductance responses during acquisition and extinction training with *F*-value, *p*-value. **$p < 0.01$; ***$p < 0.001$

	Effect	*F*-value	*p-value*
Acquisition	CS	36.92	<0.001***
	Time	2.91	0.092
	Gender	0.35	0.552
	CS × time	0.03	0.872
	CS × gender	0.22	0.638
	Time × gender	2.29	0.134
	CS × time × gender	0.01	0.937
Extinction	CS	10.89	0.001**
	Time	26.63	<0.001***
	Gender	0.05	0.831
	CS × time	1.51	0.222
	CS × gender	2.19	0.143
	Time × gender	0.98	0.325
	CS × time × gender	0.05	0.824

CS regressors were split into an early ($CS+_{early}/CS-_{early}$) and a late phase ($CS+_{late}/CS-_{late}$) to enable clear differentiation between early and late effects (Kruse et al., 2017; Kruse et al., 2018; LaBar, Gatenby, Gore, LeDoux, & Phelps, 1998). Events were modeled as stick functions and were convolved with the canonical hemodynamic response function. Six movement parameters were entered as covariates alongside regressors of no interest for the identified outlying volumes. The time series was then filtered with a high pass filter (time constant = 128 s). For acquisition training, a CS+ − CS− contrast was calculated, while for the extinction training separate $CS+_{early} - CS-_{early}$ and $CS+_{late} - CS-_{late}$ contrasts were calculated. In addition to these main analyses, analyses were extended to include a $(CS+_{late} - CS-_{late}) - (CS+_{early} - CS-_{early})$ contrast for the extinction phase.

On the group level, one-sample *t*-tests were performed for the first-level contrasts to examine neural differences in appetitive conditioning and extinction. Region of interest (ROI) analyses on the voxel level were conducted using the small volume correction in SPM12 with $p < 0.05$ (FWE). The coordinates of peak voxels of the first study on neural correlates of appetitive extinction (Kruse et al., 2017) were

used as centers of 6 mm-spheres, which were used as ROIs for small volume correction. In addition, whole brain analyses were performed with $p < 0.05$ (FWE), $k > 10$ voxel for the extinction training and, exploratively, for a two-sample t-test comparing male to female subjects. As no whole brain results emerged for the test for gender differences, we do not further include this analysis in the Results section.

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

3 | RESULTS

3.1 | Subjective ratings

ANOVA of subjective ratings of valence, arousal, and UCS-expectancy (Table 1) revealed main effects of CS, time, and a CS × time interaction (see Table 2 for detailed statistics). There was neither a significant main effect of gender, nor any significant gender × CS, gender × time, nor gender × CS × time interactions.

3.2 | Skin conductance responses

For acquisition training, ANOVA of SCRs revealed a significant main effect of CS (F[1, 74] = 36.92; $p < 0.001$). For detailed statistics, see Table 3. The main effect of CS was driven by significantly higher responses to the CS+ as compared to the CS− for the early phase (t[75] = 5.33; $p < 0.001$), as well as the late phase (t[75] = 5.29; $p < 0.001$). Similarly, for the extinction phase, we found a significant main effect of CS (F[1, 74] = 10.90; $p = 0.001$) and time (F[1, 74] = 26.63; $p < 0.001$). However, the main effect of CS was driven by significantly higher SCRs to the CS+ as compared to the CS− in the early phase (t[75] = 2.55; $p = 0.013$), but not in the late phase (t[75] = 1.55; $p = 0.126$). Neither during acquisition nor extinction there were any main effects or interactions qualified by gender (all $p > 0.13$).

3.3 | Hemodynamic responses

3.3.1 | Acquisition

Analysis of BOLD-responses during the acquisition phase shows increased responses to the CS+ as compared to the CS− throughout key areas associated with acquisition of appetitive conditioning (see Table 4).

3.3.2 | Extinction

Analysis of the early phase of extinction revealed increased BOLD contrast in dACC and hippocampus (see Table 5). A whole brain analysis further revealed a significant differential BOLD response (CS+ − CS−) in the right frontal operculum ($k = 16$; $x = 36$; $y = 16$; $z = 10$; $Z_{max} = 5.26$; $p_{FWE} = 0.007$). Crucially, during the late phase, increased differential BOLD (CS+ − CS−) emerged in the right amygdala and the left NAcc (see Figure 1). Taken together, during both early and late phase, results from the first study on neural correlates of appetitive extinction were replicated with the exception of vACC and caudate nucleus during the early phase. To extend the previous results, activation during the late phase was compared to activation during the early phase. This revealed increased differential BOLD contrast (CS+ − CS−) in the right amygdala ($k = 102$; $x = 18$; $y = 2$; $z = -20$; $Z_{max} = 3.40$; $p_{FWE} = 0.026$) during the late phase as compared to the early phase. In an exploratory analysis, we looked for significant activations in the contrast CS− − CS+, but found no significant results in neither phase. There were no significant differences between genders.

To further discern amygdala and NAcc activation during acquisition and extinction learning, we extracted contrast estimates at the locations of the peak voxels identified during acquisition and the late

TABLE 4 Region of interest (ROI) activations during acquisition (CS+ − CS−). Localization, cluster size (k), and statistics (FWE-corrected) of the peak voxel in the respective ROI

Structure	Side	k	x	y	z	z_{max}	p_{corr}
Amygdala	L	123	−20	−2	−12	5.93	<0.001
	R	123	16	−4	−8	7.23	<0.001
Caudate	L	123	−12	12	−2	6.30	<0.001
	R	123	10	10	−2	7.05	<0.001
dACC	L	123	−4	8	44	Inf	<0.001
	R	123	8	16	40	Inf	<0.001
Midbrain		123	8	−28	−6	6.26	<0.001
NAcc	L	123	−12	12	−2	6.30	<0.001
	R	123	10	4	−4	7.60	<0.001
OFC	L	123	16	16	−12	6.28	<0.001

TABLE 5 Region of interest (ROI) activations during the early and late extinction phase (CS+ − CS−). Localization, cluster size (k), and statistics (FWE-corrected) of the peak voxel in the respective ROI

Contrast	Structure	Side	k	x	y	z	z_{max}	p_{corr}
CS+ (early) − CS− (early)	dACC	L	39	−6	6	36	3.39	0.010
	Hippocampus	R	52	18	−34	6	2.89	0.038
CS+ (late) − CS− (late)	Amygdala	R	110	18	0	−20	3.70	0.003
	NAcc	L	100	−14	8	−6	3.00	0.027
		R	79	12	12	−8	2.46	0.095

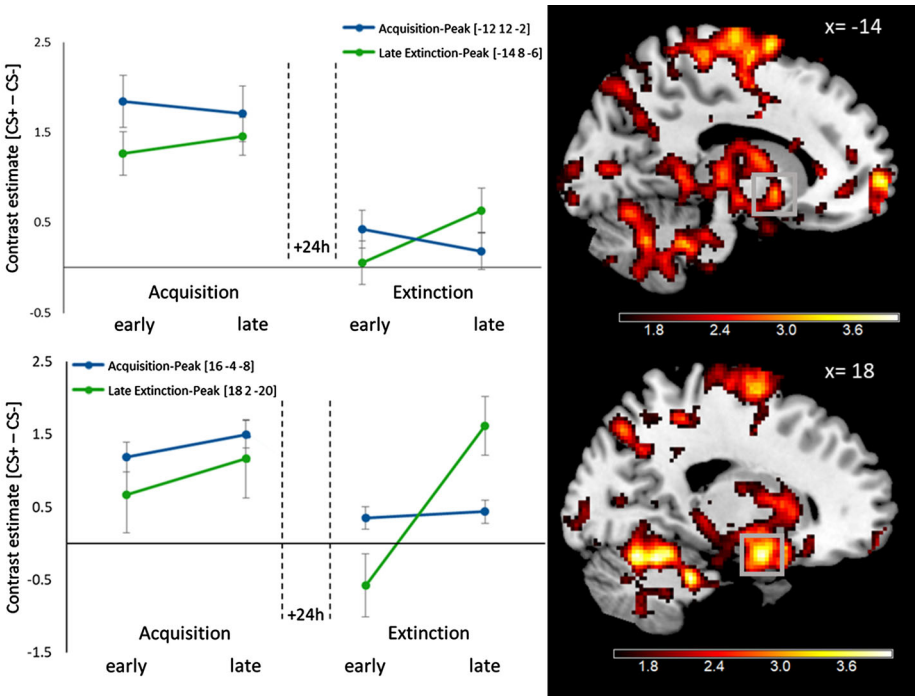

FIGURE 1 (left) Time course of contrast estimates (CS+ – CS–) at the location of the significant peak voxels during acquisition (blue) and late extinction (green) in the left NAcc (above) and right amygdala (below). The line plots show the contrast estimates for the location of each of the four peak voxels (acquisition/extinction peak in NAcc & acquisition/extinction peak in amygdala) for the early and late phase of acquisition as well as the early and late extinction phases. At the location of acquisition-peak voxels, there is significant differential BOLD-contrast during acquisition but a marked reduction of differential BOLD-contrast during extinction. At the location of extinction-peak voxels, there is differential BOLD-contrast similar to that at the acquisition peak voxels during acquisition but discernible patterns of activation during extinction. (right) Significant activation in NAcc (above) and amygdala (below) during late extinction

phase of extinction for both acquisition and extinction (Figure 1). For illustrative purposes, we created a new model also separating early and late acquisition phase. At the location of peak voxels showing greatest differentiation during acquisition training, there is a marked reduction in responding during extinction training. However, while at the location of peak voxels later showing greatest differentiation during late extinction, the pattern of activation at these locations seems similar during acquisition, responding at these locations increases during late extinction. In addition, at the location of the late extinction peak voxel in the amygdala, there is markedly reduced BOLD-contrast during the early phase of extinction.

4 | DISCUSSION

The aim of the study was to examine neural correlates of appetitive extinction in humans. To this date, only one study with a small male-only sample investigated the neural correlates of appetitive extinction (Kruse et al., 2017). The present study replicates and extends previous results in several ways. Subjects of all genders were included, as opposed to an all-male sample. This independent sample consists of 76 instead of 21 subjects, and data were collected throughout the day instead of data collection solely in the afternoon. This change in protocol ensures that there is no time of day effects due to circadian rhythms (e.g., cortisol levels). In addition, we looked at the time-course of responding for amygdala and NAcc substructures activated during acquisition or extinction training and tested for increased differential activations during the late phase of extinction as compared to the early phase. As in the first study, subjects acquired appetitive conditioning on the first day and returned 1 day later for extinction training. We will shortly discuss the results of the acquisition phase before focusing on the extinction phase.

4.1 | Acquisition training

Appetitive acquisition resulted in increased subjective ratings of arousal and valence to the CS+ as compared to ratings obtained directly before acquisition training and as compared to the CS–. In addition, UCS expectancy increased towards the CS+ as compared to pre-acquisition ratings and compared to CS–. SCRs to the CS+ were significantly increased as compared to the CS– throughout acquisition training. fMRI showed significant differential BOLD contrast (CS+ – CS–) in the amygdala, the dorsal and ventral striatum, midbrain, dorsal ACC, and the OFC. Taken together, the results are in line with previous research on acquisition of appetitive conditioning (Andreatta & Pauli, 2015; Klucken, Kruse et al., 2015; Kruse et al., 2017; Tobler, Fletcher, Bullmore, & Schultz, 2007). In line with our findings, Chase

et al. (2015) showed in their meta-analysis that involvement of the amygdala is specific to acquisition of classical conditioning in contrast to tasks focusing instrumental learning. The NAcc has mainly been associated with acquiring the subjective CS/UCS-association and contingency awareness (Klucken, Schweckendiek et al., 2009; Tapia León et al., 2018).

4.2 | Appetitive extinction

Extinction of appetitive conditioning was indicated by significant decreases in subjective ratings of valence and arousal toward the CS+ from post-acquisition to post-extinction and as compared to the CS−. UCS-expectancy towards the CS+ was also reduced significantly from post-acquisition to post-extinction. However, even post-extinction, it remained slightly but significantly higher than UCS-expectancy following the CS−. Regarding SCRs, there were significant main effects of CS and time during extinction training. Post hoc tests revealed that the main effect of CS was mainly driven by significantly higher SCRs to the CS+ during the early phase of extinction. There was no significant difference between SCRs to the CS+ as compared to the CS− during the late phase. In general, this is in line results from the first study. There is a trend toward higher UCS-expectancy to the CS+ following extinction training was visible as well, despite overall successful extinction learning.

4.3 | Early phase of appetitive extinction

Early extinction training was expected to be characterized by recall of the acquired CS/UCS association from the day before. This is in line with significantly increased SCRs following the CS+ as compared to the CS−. Regarding neural correlates, we were able to show differential activation of the dACC. The dACC is assumed to encode outcome expectancy and seems to be a neural correlate of the retrieval of a consolidated acquisition memory trace. Similarly, Ebrahimi et al. (2017) reported dACC activation during appetitive extinction in an exploratory analysis. In addition, we were able to show differential activation of the hippocampus, which might be a correlate of extinction learning being context specific (Bouton, 2002). Differential activation of caudate nucleus or vACC, that were shown to be activated during early appetitive extinction in the first study, did not replicate. This might be the case for a variety of reasons. First, nonsignificance can obviously not be interpreted as the absence of an effect. Higher in-group variance in the sample of the present sample might have decreased the power to detect these effects more than the increased sample size increased power. Second, the previous study applied a more strict protocol, only collecting data in the afternoon to have constant baseline cortisol levels for a secondary research question (Kruse et al., 2018). As this is known to affect emotional learning, it is possible, that variation in involvement of vACC and caudate nucleus during early extinction is moderated by the time of day, for example, due to variation in baseline cortisol levels (Merz, Stark, Vaitl, Tabbert, & Wolf, 2013). While established for fear conditioning, future research should assess these factors in appetitive conditioning.

4.4 | Late phase of appetitive extinction

The late phase of extinction training was expected to be characterized by successful extinction of appetitive conditioning. Differences in SCRs between CS+ and CS− were no longer significant in this phase and subjective ratings collected directly after indicated significantly reduced ratings of valence and arousal toward the CS+. Crucially, as expected, we found differential activation (CS+ − CS−) of amygdala and NAcc during the late phase of extinction. This replicates the main finding of our previous study and generalizes the assumed involvement of amygdala and NAcc to a sample consisting of male and female subjects. In line with these findings, another study, investigating the appetitive extinction of drug cues in (mainly male) cocaine users on the same day as acquisition, also found increased activation of the striatum and the amygdala during appetitive extinction (Konova et al., 2019). Extending the previous results, we analyzed for potentially higher differential activation during the late phase as compared to the early phase. We found increased differential activation of the right amygdala during the late phase. Notably, we compared the time course of contrast estimates at the location of the significant peak voxels from acquisition of conditioning to activation at the locations of peak voxels from extinction of conditioning. At the location of peak voxels of the acquisition phase, there was a marked decrease of activation throughout the extinction phase. This is in line with the expected reduction of the acquired conditioned responses. In contrast, at the location of peak voxels of the late extinction phase, there was a differentiation of acquisition and extinction signaling with a markedly increased response during the late phase of extinction. This pattern suggests discernible substructures in amygdala and NAcc which differentiate to signal extinction learning. This is in line with neuronal recordings in animal studies also showing a subpopulation of neurons specifically active after successful extinction learning (Janak et al., 2004; Tye et al., 2010). Interestingly, particularly in the amygdala, at the location of the peak voxel identified in late extinction, there also was a marked decrease during early extinction. This level even lies below the mean level of activation in the peak voxel of acquisition. This unexpected finding suggests that the amygdala plays a specific role in both early and late extinction which needs to be studied in more detail. These exploratory, descriptive results suggest that future studies should try to focus on activation patterns within these regions using pattern analyses like representational similarity analysis (RSA; Jin, Zelano, Gottfried, & Mohanty, 2015), which the present design did not permit because the employed ITI was too short (Visser et al., 2016). This might allow to capture how subnuclei within amygdala and NAcc work in parallel during acquisition but show discernible patterns during extinction.

4.5 | Limitations

Despite the higher sample size, the higher number of subjects does not allow for a complete assessment of possible boundary conditions like the effect of the time of day on variation regarding extinction

learning. Nevertheless, we were able to show that several neural correlates of appetitive extinction were robust and can be replicated with a less strict protocol in a sample consisting of male and female subjects, therefore increasing generalizability to the population. In addition, although we did not find sex differences regarding appetitive extinction, we were not able to include exact assessments of the female hormonal cycle, which is assumed to affect emotional learning, and the sample also includes women taking hormonal contraceptives (Merz, Kinner, & Wolf, 2018).

4.6 | Conclusion

In the light of the recent replicability crisis especially regarding fMRI research, it seems particularly important to investigate robustness and replicability. We set out to replicate previously reported neural correlates of appetitive extinction in an independent sample, while extending our findings to a sample consisting of subjects of all genders. Replication of the main finding, an involvement of NAcc and amygdala in the late phase of extinction, was successful. The absence of significant differences in neural activations between genders underscores the generalizability of the results. In addition, for the first time, it was possible to show subregions of amygdala and NAcc displaying separate acquisition and extinction coding signals which follow a similar course during acquisition but differentiate during extinction. As extinction of appetitive conditioning is regarded as a model for the treatment of psychiatric disorders like addiction, this allows to translate animal models that have built on the involvement of these areas in extinction learning and form the basis for pharmacological research, to humans.

ACKNOWLEDGMENTS

This study was supported by the German Research Foundation (KL2500/5-1).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Onno Kruse https://orcid.org/0000-0003-3098-4955
Isabell Tapia León https://orcid.org/0000-0002-7888-5270

REFERENCES

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, *14*(10), 1338–1344. https://doi.org/10.1038/nn.2921

Andreatta, M., & Pauli, P. (2015). Appetitive vs. aversive conditioning in humans. *Frontiers in Behavioral Neuroscience*, *9*, 128. https://doi.org/10.3389/fnbeh.2015.00128

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028

Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry*, *52*(10), 976–986. https://doi.org/10.1016/S0006-3223(02)01546-9

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Chase, H. W., Kumar, P., Eickhoff, S. B., & Dombrovski, A. Y. (2015). Reinforcement learning models and their neural correlates: An activation likelihood estimation meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(2), 435–459. https://doi.org/10.3758/s13415-015-0338-7

Ebrahimi, C., Koch, S. P., Friedel, E., Crespo, I., Fydrich, T., Ströhle, A., … Schlagenhauf, F. (2017). Combining D-cycloserine with appetitive extinction learning modulates amygdala activity during recall. *Neurobiology of Learning and Memory*, *142*(Pt B), 209–217. https://doi.org/10.1016/j.nlm.2017.05.008

Finger, E. C., Mitchell, D. G. V., Jones, M., & Blair, R. J. R. (2008). Dissociable roles of medial orbitofrontal cortex in human operant extinction learning. *NeuroImage*, *43*(4), 748–755. https://doi.org/10.1016/j.neuroimage.2008.08.021

Janak, P. H., Chen, M.-T., & Caulder, T. (2004). Dynamics of neural coding in the accumbens during extinction and reinstatement of rewarded behavior. *Behavioural Brain Research*, *154*(1), 125–135. https://doi.org/10.1016/j.bbr.2004.02.003

Jin, J., Zelano, C., Gottfried, J. A., & Mohanty, A. (2015). Human amygdala represents the complete spectrum of subjective valence. *The Journal of Neuroscience*, *35*(45), 15145–15156.

Klucken, T., Kruse, O., Wehrum-Osinsky, S., Hennig, J., Schweckendiek, J., & Stark, R. (2015). Impact of COMT Val158Met-polymorphism on appetitive conditioning and amygdala/prefrontal effective connectivity. *Human Brain Mapping*, *36*(3), 1093–1101. https://doi.org/10.1002/hbm.22688

Klucken, T., Schweckendiek, J., Merz, C. J., Tabbert, K., Walter, B., Kagerer, S., … Stark, R. (2009). Neural activations of the acquisition of conditioned sexual arousal: Effects of contingency awareness and sex. *The Journal of Sexual Medicine*, *6*(11), 3071–3085. https://doi.org/10.1111/j.1743-6109.2009.01405.x

Klucken, T., Tabbert, K., Schweckendiek, J., Merz, C. J., Kagerer, S., Vaitl, D., & Stark, R. (2009). Contingency learning in human fear conditioning involves the ventral striatum. *Human Brain Mapping*, *30*(11), 3636–3644. https://doi.org/10.1002/hbm.20791

Klucken, T., Wehrum, S., Schweckendiek, J., Merz, C. J., Hennig, J., Vaitl, D., & Stark, R. (2013). The 5-HTTLPR polymorphism is associated with altered hemodynamic responses during appetitive conditioning. *Human Brain Mapping*, *34*(10), 2549–2560. https://doi.org/10.1002/hbm.22085

Knutson, B., Westdorp, A., Kaiser, E., & Hommer, D. (2000). FMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage*, *12*(1), 20–27. https://doi.org/10.1006/nimg.2000.0593

Konova, A. B., & Goldstein, R. Z. (2019). The emerging neuroscience of appetitive and drug cue extinction in humans. *Psychopharmacology*, *236*(1), 407–414. https://doi.org/10.1007/s00213-018-5098-y

Konova, A. B., Parvaz, M. A., Bernstein, V., Zilverstand, A., Moeller, S. J., Delgado, M. R., … Goldstein, R. Z. (2019). Neural mechanisms of extinguishing drug and pleasant cue associations in human addiction: Role of the VMPFC. *Addiction Biology*, *24*(1), 88–99. https://doi.org/10.1111/adb.12545

Kruse, O., Tapia León, I., Stalder, T., Stark, R., & Klucken, T. (2018). Altered reward learning and hippocampal connectivity following psychosocial stress. *NeuroImage*, *171*, 15–25. https://doi.org/10.1016/j.neuroimage.2017.12.076

Kruse, O., Tapia León, I., Stark, R., & Klucken, T. (2017). Neural correlates of appetitive extinction in humans. *Social Cognitive and Affective Neuroscience*, *12*(1), 106–115. https://doi.org/10.1093/scan/nsw157

LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E., & Phelps, E. A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: A mixed-trial fMRI study. *Neuron*, *20*(5), 937–945. https://doi.org/10.1016/S0896-6273(00)80475-4

Martin-Soelch, C., Linthicum, J., & Ernst, M. (2007). Appetitive conditioning: Neural bases and implications for psychopathology. *Neuroscience and Biobehavioral Reviews*, *31*(3), 426–440. https://doi.org/10.1016/j.neubiorev.2006.11.002

Merz, C. J., Kinner, V. L., & Wolf, O. T. (2018). Let's talk about sex … differences in human fear conditioning. *Current Opinion in Behavioral Sciences*, *23*, 7–12. https://doi.org/10.1016/j.cobeha.2018.01.021

Merz, C. J., Stark, R., Vaitl, D., Tabbert, K., & Wolf, O. T. (2013). Stress hormones are associated with the neuronal correlates of instructed fear conditioning. *Biological Psychology*, *92*(1), 82–89. https://doi.org/10.1016/j.biopsycho.2012.02.017

Millan, E. Z., Marchant, N. J., & McNally, G. P. (2011). Extinction of drug seeking. *Behavioural Brain Research*, *217*(2), 454–462. https://doi.org/10.1016/j.bbr.2010.10.037

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, *304*(5669), 452–454. https://doi.org/10.1126/science.1094285

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329–337. https://doi.org/10.1016/S0896-6273(03)00169-7

Pineles, S. L., Orr, M. R., & Orr, S. P. (2009). An alternative scoring method for skin conductance responding in a differential fear conditioning paradigm with a long-duration conditioned stimulus. *Psychophysiology*, *46*(5), 984–995.

Quirk, G. J., & Mueller, D. (2008). Neural mechanisms of extinction learning and retrieval. *Neuropsychopharmacology*, *33*(1), 56–72.

Schweckendiek, J., Klucken, T., Merz, C. J., Kagerer, S., Walter, B., Vaitl, D., & Stark, R. (2013). Learning to like disgust: Neuronal correlates of counterconditioning. *Frontiers in Human Neuroscience*, *7*, 346. https://doi.org/10.3389/fnhum.2013.00346

Tapia León, I., Kruse, O., Stalder, T., Stark, R., & Klucken, T. (2018). Neural correlates of subjective CS/UCS association in appetitive conditioning. *Human Brain Mapping*, *39*, 1637–1646. https://doi.org/10.1002/hbm.23940

Tobler, P. N., Fletcher, P. C., Bullmore, E. T., & Schultz, W. (2007). Learning-related human brain activations reflecting individual finances. *Neuron*, *54*(1), 167–175. https://doi.org/10.1016/j.neuron.2007.03.004

Tye, K. M., Cone, J. J., Schairer, W. W., & Janak, P. H. (2010). Amygdala neural encoding of the absence of reward during extinction. *The Journal of Neuroscience*, *30*(1), 116–125. https://doi.org/10.1523/jneurosci.4240-09.2010

Visser, R. M., Haan, M. I. C. d., Beemsterboer, T., Haver, P., Kindt, M., & Scholte, H. S. (2016). Quantifying learning-dependent changes in the brain: Single-trial multivoxel pattern analysis requires slow event-related fMRI. *Psychophysiology*, *53*(8), 1117–1127.

How to cite this article: Kruse O, Klein S, Tapia León I, Stark R, Klucken T. Amygdala and nucleus accumbens involvement in appetitive extinction. *Hum Brain Mapp*. 2020; 41:1833–1841. https://doi.org/10.1002/hbm.24915