



ARTICLE

Reduced safety processing during aversive social conditioning in psychosis and clinical risk

Megan Quarmley¹, Ruben C. Gur¹, Bruce I. Turetsky¹, Anna J. Watters¹, Warren B. Bilker², Mark A. Elliott³, Monica E. Calkins¹, Christian G. Kohler¹, Kosha Ruparel¹, Petra Rupert¹, Raquel E. Gur¹ and Daniel H. Wolf¹

Social impairment occurs across the psychosis spectrum, but its pathophysiology remains poorly understood. Here we tested the hypothesis that reduced differential responses (aversive vs. neutral) in neural circuitry underpinning aversive conditioning of social stimuli characterizes the psychosis spectrum. Participants age 10–30 included a healthy control group (HC, analyzed $n = 36$) and a psychosis spectrum group (PSY, $n = 71$), including 49 at clinical risk for psychosis and 22 with a frank psychotic disorder. 3T fMRI utilized a passive aversive conditioning paradigm, with neutral faces as conditioned stimuli (CS) and a scream as the unconditioned stimulus. fMRI conditioning was indexed as the activation difference between aversive and neutral trials. Analysis focused on amygdala, ventromedial prefrontal cortex, and anterior insula, regions previously implicated in aversive and social-emotional processing. Ventromedial prefrontal cortex activated more to neutral than aversive CS; this “safety effect” was driven by HC and reduced in PSY, and correlated with subjective emotional ratings following conditioning. Insula showed the expected aversive conditioning effect, and although no group differences were found, its activation in PSY correlated with anxiety severity. Unexpectedly, amygdala did not show aversive conditioning; its activation trended greater for neutral than aversive CS, and did not differ significantly based on group or symptom severity. We conclude that abnormalities in social aversive conditioning are present across the psychosis spectrum including clinical risk, linked to a failure of safety processing. Aversive and safety learning provide translational paradigms yielding insight into pathophysiology of psychosis risk, and providing potential targets for therapeutic and preventative interventions.

Neuropsychopharmacology (2019) 44:2247–2253; <https://doi.org/10.1038/s41386-019-0421-9>

INTRODUCTION

Schizophrenia is a devastating neuropsychiatric disorder characterized by hallucinations, delusions, and deficits in cognitive and social-emotional abilities. Much of the disability in schizophrenia relates to social impairment [1], which is driven by a combination of negative symptoms such as asociality [2, 3], positive symptoms such as paranoia, [4] and deficits in social cognition [5]. While most prior work examining social impairment in psychosis has focused on chronic schizophrenia, social dysfunction associated with negative symptoms is also common at the first episode, and evident in youth at clinical risk for psychosis [6, 7]. Furthermore, while positive symptoms are the strongest contributors to risk for conversion to frank psychosis, negative symptoms and social impairment also increase conversion risk [8]. However, the neurobiology of social dysfunction and associated symptom domains remains poorly understood in schizophrenia, and even less is known in at-risk populations. A wide range of possible neurobehavioral mechanisms have been considered, from basic perceptual processes to high-level social cognition [5, 9].

One promising candidate mechanism contributing to social dysfunction involves hyper-reactivity of threat-processing circuitry, associated with inappropriate aversive reactions to neutral stimuli and a failure to distinguish threat from nonthreat. Such overgeneralization of aversive processing would impact social

inferences and avoidance behavior leading to social dysfunction. To date, most of the evidence supporting this theory comes from functional neuroimaging studies of emotion identification in schizophrenia, where overactivation to neutral faces in amygdala is associated with reduced differential activation to aversive versus neutral faces [10]. Amygdala hyperactivation to threatening facial expressions has also been linked to severity of negative symptoms or positive symptoms in schizophrenia [11, 12].

Aversive conditioning provides an appealing paradigm to investigate this putative pathophysiology. Unlike the emotion identification paradigms, conditioning paradigms capture the learning process by which aversive value is assigned to neutral stimuli, and also leverage a more extensive animal model literature [13, 14]. Both animal models and human neuroimaging have identified a set of interconnected regions involved in aversive conditioning, including the amygdala, anterior insular cortex (AIC), and ventromedial prefrontal cortex (vmPFC) [15, 16]. The amygdala and AIC generally respond preferentially to aversive stimuli, while vmPFC responds preferentially to the neutral “safety” stimulus, and is thought to modulate threat responses [17, 18]. Overlapping circuitry is implicated in emotion regulation and social cognition [15, 16], providing a plausible biological connection between aberrant aversive processing and social dysfunction.

¹Department of Psychiatry, University of Pennsylvania, Philadelphia, PA 19104, USA; ²Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, PA 19104, USA and ³Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, USA
Correspondence: Daniel H. Wolf (danwolf@penmedicine.upenn.edu)

Received: 5 February 2019 Revised: 29 April 2019 Accepted: 8 May 2019
Published online: 21 May 2019

Table 1. Demographic and clinical characteristics of the analyzed fMRI sample

Variable	HC (n = 36)	PSY (n = 71)	CR (n = 49)	PD (n = 22)	t/χ^2 value	p value
Age, mean (SD)	18.8 (3.8)	20.2 (3.5)	19.9 (3.3)	20.8 (3.9)	-1.9	0.06 ^{a,b}
Sex (% Female)	56%	51%	51%	50%	0.2	0.63 ^c
Race (% Caucasian)	56%	39%	33%	55%	2.5	0.11
Education, mean (SD)	11.8 (3.6)	12.1 (2.5)	12.1 (2.7)	12.2 (2.1)	-0.5	0.59
Parental education, mean (SD)	14.7 (2.1)	14.2 (2.7)	14.0 (2.7)	14.7 (2.6)	1.0	0.30
Negative symptoms, mean (SD) ^d	7.2 (5.7)	14.9 (10.0)	12.4 (7.7)	20.6 (12.2)	-4.2	<.0001
Positive symptoms, mean (SD) ^e	1.1(1.8)	9.2 (6.0)	6.9 (4.6)	15.1 (5.4)	-7.7	<.0001
Trait anxiety, mean (SD) ^f	30.8 (7.2)	40.7 (13.2)	38.4 (12.2)	45.6 (14.3)	-4.1	<.0001
Cognition, mean (SD) ^g	0.51 (0.65)	-0.26 (1.05)	-0.06 (0.95)	-0.79 (1.13)	4.0	0.0001
Antipsychotic use (% Yes)	—	21%	6%	59%	—	—

^aTwo-tailed p values, comparing healthy control (HC) and psychosis (PSY) groups
^bStudent's t test used to compare group means for dimensional variables
^cChi-square test was used to compare proportions for categorical variables (race, sex)
^dTotal score on the Clinical Interview for Negative Symptoms
^eTotal of positive symptom items on the Scale of Prodromal Symptoms
^fTrait score from the Stait-Trait Anxiety Inventory
^gComputerized Neurocognitive Battery overall accuracy, z scores

Several studies have examined aversive conditioning in schizophrenia, and consistently identify reduced differential conditioning, as indexed by autonomic responses [19–23] and fMRI activation [19, 20, 22]; one study in a small sample did not find any abnormal response to the cue [24]. This reduction in differential conditioning has been related to reduced responses to aversive stimuli [21], increased responses to neutral stimuli [23], and a combination of these effects [19, 20, 22].

Limited work with face emotion paradigms indicates that exaggerated responsivity in threat-related circuitry is also present in psychosis-risk populations [25, 26]. This finding is consistent with broader work suggesting that psychosis occurs on a spectrum, with those at risk showing abnormalities similar to but less severe than those seen in frank psychosis [27]. However, to date there have been no fMRI studies investigating aversive conditioning in a psychosis spectrum sample that includes individuals at risk for psychosis.

Here we attempt to fill this gap and increase understanding of the pathophysiology of social dysfunction in psychosis, using a novel social aversive conditioning fMRI paradigm. We hypothesized that a reduction in differential aversive conditioning responses would be present across the psychosis spectrum from risk to frank schizophrenia, in regions associated with both fear conditioning and social cognition including the amygdala, vmPFC, and AIC. We expected that conditioning phenotypes in the clinical risk group would be intermediate between the healthy and frank psychosis groups.

METHODS

Participants

Participants were recruited from several sources including the Philadelphia Neurodevelopmental Cohort [28, 29], Penn Psychosis Evaluation and Recovery Clinic (www.med.upenn.edu/bbl/penn-perc.html), iConnect (University of Pennsylvania recruitment database), and the surrounding community. The fMRI study sample included 118 individuals between the ages of 10–30 meeting clinical and demographic criteria: 37 healthy comparison (HC) individuals without any history of psychosis in themselves or their first-degree relatives, and 81 individuals within the psychosis spectrum (PSY). Within this psychosis spectrum group, 55 individuals were at clinical risk for psychosis (CR) and 26 individuals had a DSM-IV psychotic disorder (PD). After complete

description of the study to the participants, written informed consent and assent (age < 18) were obtained. All study procedures were approved by the University of Pennsylvania’s Institutional Review Board. fMRI analysis included data from 107 participants (36 HC, 49 CR, 22 PD; see Supplementary Methods for exclusions). Groups exhibited expected differences in symptom severity and overall cognition, but did not differ demographically (Table 1).

Clinical assessment

A clinical diagnostic interview and cognitive assessment were performed at the initial study visit. fMRI and EEG were conducted on the second and third visits, with the order of fMRI and EEG visits counterbalanced (EEG reported separately [30]). On the second visit, dimensional clinical measures were obtained including negative symptoms, which were our a priori primary symptom focus, positive symptoms, and trait anxiety. See Supplementary Methods for additional assessment details.

MRI methods

3T Blood Oxygen Level Dependent (BOLD) fMRI was performed during a single run of initial aversive conditioning, reported here, as well as a run of reversal conditioning and a resting-state scan, to be reported separately. See Supplementary Methods for details of acquisition and image preprocessing.

Social aversive conditioning fMRI paradigm. The novel fMRI task used a passive social aversive delay conditioning paradigm, integrating aspects of various tasks used previously by others [22, 31, 32]. Two color photographs of affectively neutral human male faces served as conditioned stimuli (CS). One face was paired with an aversive auditory unconditioned stimulus (US), yielding an aversively conditioned face (CS+), and the other face was not paired with any sound, yielding a nonaversively or neutrally conditioned face (CS–). The single US was a 1 s mixture of two human screams (male and female, without words), presented binaurally at 100 dB SPL (Supplementary Audio file) via noise-cancelling headphones (OptoAcoustics, Israel). The same two conditioned stimuli (CS) were used for each subject, with the pairing to aversive US counterbalanced across subjects. The temporal sequence of CS+ and CS– was also counterbalanced across subjects, by switching all CS+ trials to CS– and vice versa. The CS+ was paired with the US on 50% of CS+ trials (CS+p; Fig. 1), and unpaired on the other half of CS+ trials (CS+u).

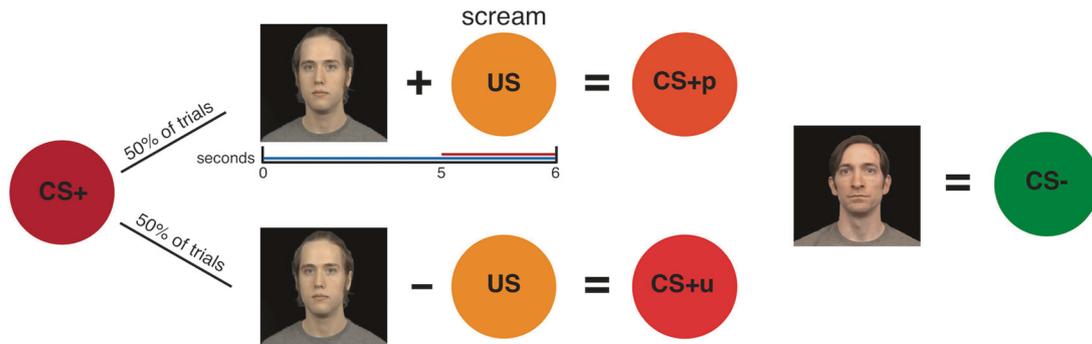


Fig. 1 Passive aversive delay conditioning fMRI paradigm. One neutral male face (CS+) co-terminated with an aversive scream US on half of the trials (CS+p) and was not paired with the scream on the other half (CS+u). The second neutral male face (CS-) was never paired with the scream

Each trial consisted of a 6 s presentation of the CS (co-terminating with the 1 s US for CS+p trials). During the inter-trial interval (mean ~4 s, range 2–12 s), a complex crosshair was displayed (scrambled face with central asterisk), designed to match the faces on lower level perceptual features including color and luminance. There were 48 trials: 12 CS+p, 12 CS+u, and 24 CS-. Trials were ordered in a fast event-related sequence, counter-balanced so that each CS was equally likely to follow the US and equally likely to follow inter-trial intervals of different durations. To ensure task attention, participants were instructed to press a response button during each of eight 2 s “catch events,” signaled by the crosshair asterisk turning from black to green at unpredictable intervals unrelated to conditioning events. Total task duration was 8 min 48 s.

Subjective ratings. Subjective ratings were also obtained for each CS prior to and following scanning. Prior to scanning, subjects rated a set of 17 face images for valence and arousal; the two CS images were included within this larger set without being identified as task relevant. Following completion of the conditioning run, subjects provided these ratings again only for the two CS images. Here we focused group comparisons on the post-scan rating of subjective feelings while viewing the faces because this measure showed the most robust conditioning effect across the full sample; we refer to this subjective rating as “behavioral conditioning.” Participants were also asked to identify which CS had been preferentially paired with the scream. At the end of the fMRI session they answered additional questions regarding subjective responses to the task in general and the US specifically. See Supplementary Methods for additional details regarding behavioral measures.

Statistical analysis

Individual-level fMRI analysis. Individual-level time-series analysis was carried out using a general linear model in FSL (<https://www.fsl.fmrib.ox.ac.uk>) version 5.0.2.1. Event-related analysis focused on comparing BOLD responses to CS- and CS+u trials, thus isolating the effects of conditioning by the US without the confound of responses to the US itself. Task regressors (CS-, CS+u, CS+p) with 6 s event durations were convolved with a canonical double-gamma hemodynamic response function. Additional regressors modeled catch events and 24 motion parameters (standard and extended). The primary contrast of interest was CS+u > CS-; this is the most common index of fMRI aversive conditioning in the literature [17].

Group-level fMRI analysis. Based on our strong a priori hypotheses regarding fear conditioning circuitry [17, 33] our primary regions of interest (ROIs) were the amygdala, anterior insula (AIC), and vmPFC. These three bilateral ROIs were then merged

into a single mask for voxelwise analyses (Supplementary Methods and Fig. S1).

Voxelwise analyses were run across all subjects to identify effects of conditioning regardless of group. Our primary group comparison was HC > PSY, testing our key hypothesis that reduced differential conditioning would be present across the psychosis spectrum. We also tested the hypothesis that differential conditioning abnormalities in CR would be intermediate between HC and PD, using a group-level regression to implement a linear trend test by coding each HC individual as +1, each CR as 0, and each PD as -1 (along with a constant mean intercept regressor). Secondary pairwise analyses compared the two clinical groups (CR and PD) to each other and to HC. Significant clusters within the ROI masks were defined as $p < 0.05$, FWE-corrected using 5000 permutations in FSL’s *randomise*, with threshold-free cluster enhancement (TFCE) [34]. Such permutation methods minimize distributional assumptions, ensuring rigorous control of multiple comparisons [35]. Additional exploratory analyses examining dimensional correlations were performed in R using contrast parameters extracted separately for each individual subject from each of the three ROIs defined above. Exploratory p values were not corrected for multiple comparisons.

RESULTS

Differential conditioning effects common across groups

Differential aversive conditioning effects (CS+u > CS-) were first examined across all subjects (Fig. 2). The insula (AIC) as expected activated more to CS+u than to CS-. Also as expected, vmPFC activated to CS- but deactivated to CS+u. Amygdala activated to both CS- and CS+u without a significant differential conditioning effect, but the trend was toward greater activation to the CS- than the CS+. The whole-brain pattern of conditioning responses (Fig. S2) was generally consistent with results from a recent meta-analysis; that study also reported nonsignificant differential conditioning effects in amygdala but did not specify the direction of any trend [17].

Between-group comparison of differential conditioning

There was significantly greater differential conditioning in vmPFC in HC than PSY (Fig. 3a). Clinical risk participants (CR) were intermediate between HC and those with frank psychotic disorders (PD), as indicated by a significant linear HC > CR > PD effect (Fig. 3b). HC showed the expected “safety” conditioning (response greater to CS- than CS+u, producing deactivation in the CS+u > CS- contrast), while PSY showed a blunting of this normative response. This group difference was driven by both blunted activation to the CS- and blunted deactivation to CS+u; there were no group differences in vmPFC response to the CS+p trials where the scream US was present (Fig. S3). In post hoc voxelwise analyses within our three ROIs, neither the two-group

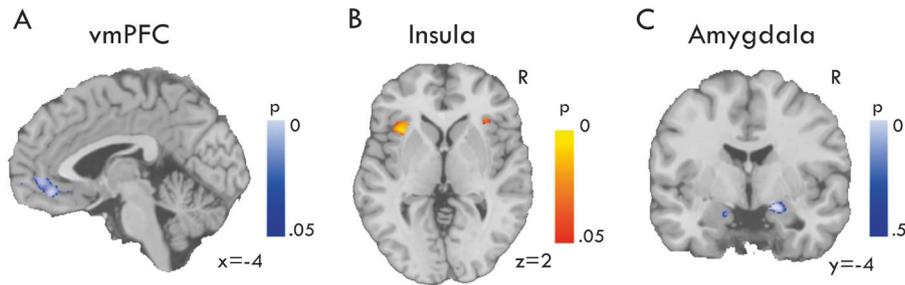


Fig. 2 Differential conditioning effects in CS+u > CS- contrast (all participants, $n = 107$). Hot colors represent CS+u > CS- activation, cold colors represent deactivation (CS- > CS+u). Significant clusters in vmPFC (a) and AIC (b) are corrected within the combined-ROI mask using threshold-free cluster enhancement (TFCE), $p < 0.05$; c shows the subthreshold trend in amygdala (peak TFCE $p = 0.07$)

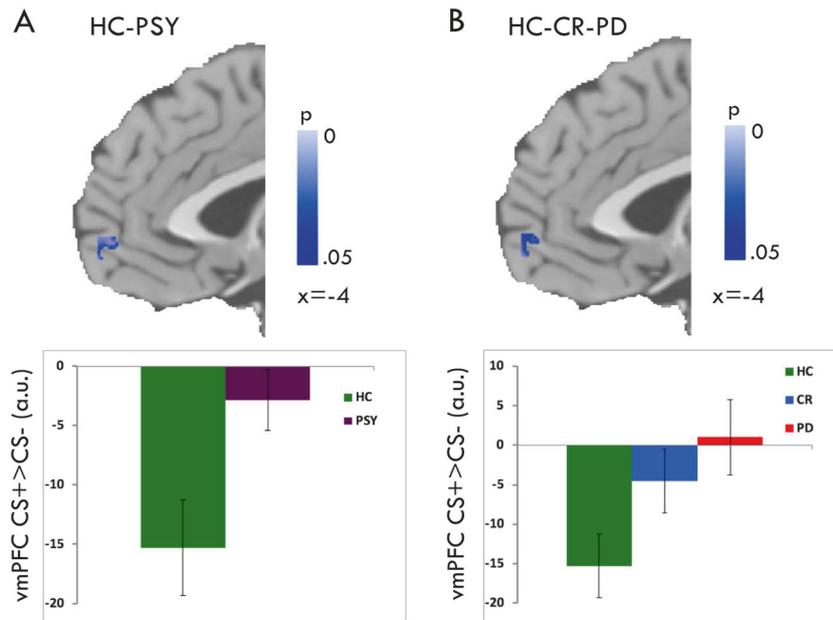


Fig. 3 Differential conditioning effects (CS+ > CS- contrast) in vmPFC for two-group comparison (a) and three-group linear trend analysis (b). Voxelwise maps show significant TFCE-corrected clusters with greater differential activation in healthy controls (HC) than in the psychosis spectrum group (PSY) (42 voxels, peak $t = -3.60$ at MNI $-4,58,-2$); the clinical risk (CR) was intermediate between these two groups as shown by the significant linear trend (38 voxels, peak $t = -3.63$ at MNI $-2,62,-10$). The cold (blue) color indicates that the vmPFC deactivation is stronger (more negative) in HC than PSY. Descriptive bar graphs show data extracted from vmPFC; errors bars are SEM

nor the linear trend analysis revealed any significant clusters with greater differential conditioning in PSY than HC, nor for CS- or CS+u conditions separately, nor for the CS+p (scream US) condition. There were no significant clusters showing HC vs. PSY group differences or linear HC > CR > PD effects within the amygdala or AIC ROIs. Consistent with the a priori analyses, post hoc voxelwise pairwise group comparisons (HC vs. CR, HC vs. PD, CR vs. PD) within the three ROIs showed a significant TFCE-corrected cluster in vmPFC for HC vs. PD (peak $p = 0.04$) and a trend for HC vs. CR (peak $p = 0.08$), without any significant results in amygdala or AIC. Examination of contrast parameters extracted from each of these ROIs separately (averaging across all voxels within an ROI) revealed significant differences only for vmPFC HC vs. PD ($p = 0.01$; HC vs. CR $p = 0.07$; CR vs. PD $p = 0.42$).

Exploratory correlations of fMRI differential conditioning with behavioral ratings

Across all subjects, there was a significant positive correlation between differential vmPFC activation and differential post-conditioning subjective rating of emotional valence (Fig. 4a, b;

Spearman $\rho = 0.23$, uncorrected $p = 0.02$). The correlation with vmPFC activation was also significant for the post-conditioning ratings of CS- ($\rho = -0.23$, $p = 0.02$), but not for the post-conditioning ratings of CS+ ($\rho = 0.13$, $p = 0.15$) or the pre-conditioning ratings ($\rho = -0.09$, $p = 0.35$). A linear trend (HC > CR > PD) was statistically significant for post-conditioning CS- ratings ($p = 0.02$) and the HC vs. PSY group difference was borderline significant ($p = 0.054$); HC showed positive ratings while CR ratings were neutral and PD actually rated the CS- negatively (Fig. 4c). There were no significant linear trends or group differences for the post-conditioning CS+ ratings, or for differential ratings (CS+ > CS-) (Fig. S4). This abnormality in post-scan CS- ratings did not reflect baseline abnormalities in subjective responses to faces, or failure in basic attention to the task, as there were no group differences in pre-conditioning subjective ratings, response accuracy for catch events, accuracy of post-conditioning identification CS-US pairings, or aversive response to the US (see Supplementary Tables S1, S2). There were no significant correlations between subjective differential ratings and differential activation in AIC or amygdala.

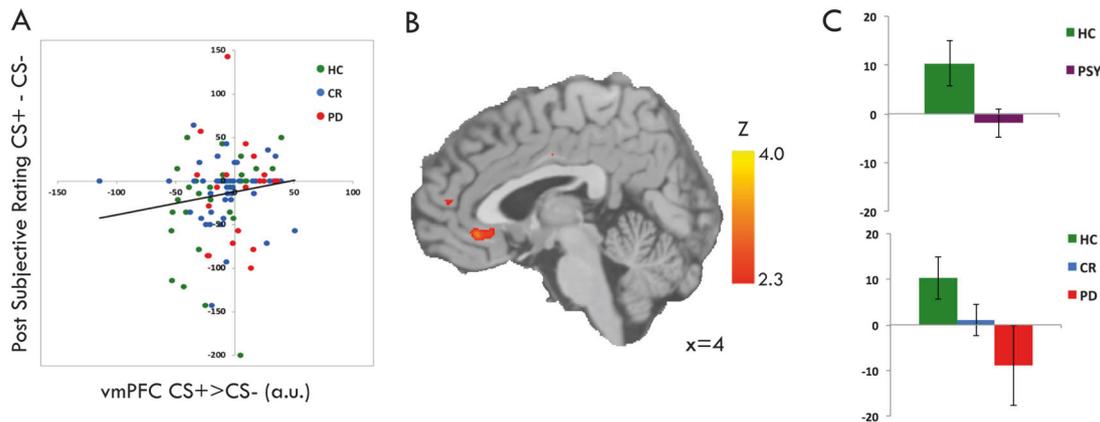


Fig. 4 **a** Differential conditioning data extracted from the vmPFC ROI correlated significantly with differential post-conditioning subjective valence ratings (participants with valid behavioral data, $n = 102$). A more normative activation pattern in the vmPFC (negative values, stronger response to CS- than CS+) was associated with a more normative subjective rating pattern (negative values, CS- rated more positive than CS+). **b** Descriptive voxelwise image shows location of this correlation in vmPFC, display threshold $z = 2.33$. **c** Bar graphs show that HC subjectively rated CS- positively, while PSY did not

Exploratory correlations with symptoms, US unpleasantness, cognition, age and sex

Across the psychosis spectrum, reduced differential conditioning in the bilateral AIC correlated with greater anxiety symptom severity (Fig. S5, $\rho = 0.27$, uncorrected $p = 0.028$), but not negative or positive symptom severity. No significant correlations were found between clinical symptom severity and activation in vmPFC or amygdala or with the behavioral measure of conditioning (Fig. S6).

Additional exploratory analyses examined the subjective ratings of US unpleasantness, overall cognitive ability, age, and sex, as described in Supplementary Results and Table S3. No relationships were found with the behavioral measure of conditioning. Significant (uncorrected) relationships to differential fMRI activation within our a priori ROIs included: (1) higher subjective US unpleasantness ratings correlated with greater insula differential conditioning and with more negative (normative) differential conditioning in vmPFC; (2) greater cognitive performance correlated with greater differential conditioning in the insula; (3) females showed greater differential conditioning than males in the insula. Importantly, controlling for these variables did not significantly change the group effects or the behavior–vmPFC correlation reported above. Also note that none of the exploratory results reported here would survive multiple comparisons correction given the many tests performed.

DISCUSSION

To our knowledge, this is the first study examining aversive conditioning across the psychosis spectrum from risk to frank disorder. Our findings reveal social aversive conditioning abnormalities in established psychosis and similar yet less robust intermediate abnormalities in those at clinical risk, providing further functional neuroimaging support for the spectrum conceptualization of psychosis [25, 27].

Across this psychosis spectrum, we found a reduction in differential conditioning, indicating a failure to distinguish between aversive and neutral conditioned stimuli. Behaviorally, this effect was driven by a reduction of the normative positive response to the neutral or “safe” stimulus. In fMRI, vmPFC showed safety conditioning effects (neutral > aversive) in healthy individuals, with abnormal blunting of this response across the psychosis spectrum, correlating with abnormal behavioral conditioning. This correlation suggests that a reduction or reversal of the normative vmPFC response may underpin abnormalities in the conscious evaluation of CS valence and value.

The vmPFC findings are consistent with prior fMRI work showing activation of vmPFC to safety stimuli [17, 18], as well as a broader involvement of vmPFC in processing rewards and positive emotions [36, 37], and in downregulating negative emotions and threat responses [20, 33, 38]. Our findings are also consistent with prior work in schizophrenia showing overgeneralization of aversive processing to neutral stimuli [10, 19, 20]. The regions exhibiting conditioning abnormalities in schizophrenia have varied across studies, including amygdala, inferior parietal cortex, posterior cingulate, and ventral striatum. Such variation may reflect the consequences of limited statistical power or differences in the specific populations or paradigms [19, 20, 22]. To our knowledge, the present study is the first to identify such abnormalities within vmPFC in psychosis during initial aversive conditioning. However, one prior study reported reduced vmPFC responses to the safe context during fear conditioning extinction recall [20], and a failure to deactivate vmPFC has been observed in schizophrenia in other emotional paradigms [39]. Although no studies have previously examined aversive conditioning within psychosis risk, there is evidence that at-risk populations exhibit structural abnormalities in vmPFC and its temporo-limbic connections [40, 41], and abnormal stress-induced dopamine responses in vmPFC [42]. This phenotype may not be specific to the psychosis spectrum as reduced discrimination between fear and safety cues in vmPFC has also been found in anxiety disorders [33]. However, while it is plausible that the observed impairment in safety processing would lead to a tendency to interpret the environment as threatening and thus impact clinical symptoms and social function, here we did not find a relationship between vmPFC activation and severity of anxiety or other symptom dimensions. Limitations in measurement reliability and the wide multiplicity of factors impacting clinical symptoms likely impeded our ability to detect clinical correlates. It is likely that the impaired conditioning seen here has stronger effects on automatic evaluation than on explicit evaluation, and may impact neutral or ambiguous stimuli more than clearly valenced stimuli, partly limiting the impact on real-world function.

The insula showed the most robust aversive conditioning effects (aversive > neutral) in our study. This finding is consistent with prior fMRI studies [17], and with the known role of this region in processing aversive stimuli, particularly aspects of salience, consciousness of emotions based on interoception, and anticipation of threat [17, 31]. Furthermore, in our study only the insula showed a significant (uncorrected) correlation with symptom severity. Across the psychosis spectrum, higher trait anxiety was

associated with greater aversive conditioning responses in this region; however, this exploratory correlation must be interpreted with caution and bears replication in another psychosis spectrum sample.

Unlike the insula, the amygdala did not show aversive conditioning in our study, and in fact showed a trend for greater response to the neutral than the aversive CS. Contrary to prior fMRI studies from our group using emotion identification paradigms that did not involve in-scanner conditioning [11, 25], we also did not find significant correlations between heightened amygdala threat reactivity and symptom severity, although we did observe trends in this direction. These negative findings are somewhat surprising in the context of the human and animal literature demonstrating an essential and selective role for the amygdala in aversive conditioning [13, 43]. However, they are much less surprising in the context of the human fMRI literature, where aversive conditioning responses are only inconsistently observed in amygdala, and a recent meta-analysis showed neither activation nor deactivation in the amygdala [17, 44]. The amygdala also responds to rewarding stimuli [45], and different subnuclei and neuronal populations have heterogeneous and even antagonistic actions [46, 47]. Aversive conditioning responses in amygdala may also have a complex within-trial (cue-onset phase vs. outcome phase) or across-trial temporal pattern of activation (conditioning followed by habituation) that is not captured well by the standard CS+ > CS- contrast we employed [17, 31, 48]. The observed amygdala response therefore likely reflects unresolved spatial and temporal complexity rather than a true failure to encode aversive conditioning. It is likely that conditioning effects we observed occur fairly rapidly and are sustained, which is why they are captured in the average CS+ > CS- difference. Characterizing conditioning abnormalities in amygdala will benefit from improved techniques for parsing the complexity of amygdala fMRI responses, including higher spatial resolution with 7 T MRI, task paradigms that use intra-trial jitter to temporally dissociate CS onset and US onset phases of each trial, and inclusion of more trials analyzed using reinforcement learning and habituation models to capture individual differences in the time course of these processes.

Limitations

A significant limitation of this study is that the majority of participants with frank psychotic disorders were taking antipsychotic medications, while most clinical risk participants were not. While this limitation is common in studies including both individuals at risk and those with psychotic disorders, future studies with large samples including participants both on and off medication will be better able to address the potential impact of medication. Another limitation is that we did not employ a nonsocial aversive conditioning task for comparison. While our social paradigm was chosen to maximize engagement of social-emotional circuitry and increase clinical relevance to social impairment, further work will be required to determine whether the observed abnormalities are selective for social stimuli or are associated with aversive conditioning for nonsocial stimuli as well, and whether sensory modality (auditory, visual, olfactory, gustatory) impacts the outcome. It is possible that our results were impacted by group or individual differences in mood state before or during the task but we did not obtain state mood measures at the time of scan. Although the fMRI differential conditioning effects reported above were not confounded by age or sex or race, our CS stimuli only included young-adult male Caucasian faces, and future studies using a variety of actors might reveal interactions between participant and actor demographics. Efforts to relate abnormal aversive conditioning across the psychosis spectrum to abnormalities in neurotransmitter systems such as dopamine, glutamate, and GABA will also be critical [42, 49, 50].

Conclusions

Our findings provide novel evidence for social aversive conditioning abnormalities across the psychosis spectrum, linked to a failure of safety processing. In this first study of aversive conditioning to include a clinical psychosis risk population, our findings indicate that these abnormalities can arise early in the course of illness, and consequent disruption of social processing may contribute to further illness progression. Abnormalities of aversive conditioning deserve further study in psychosis risk for insight into pathophysiology and as a potential therapeutic target. Readily implemented in animal models, aversive conditioning paradigms can serve as an informative translational bridge spanning from genetics up through neural circuitry and clinical phenomenology.

FUNDING AND DISCLOSURE

This work was supported by the National Institute of Mental Health Conte Center P50MH096891, with additional support from RC2MH089983, RC2MH089924, T32MH019112, R01MH107235, and R01MH113565. The research also received support from the Lifespan Brain Institute (LiBI) of Penn Medicine and the Children's Hospital of Philadelphia. The authors declare no competing interests.

ACKNOWLEDGEMENTS

The authors thank Karthik Prabhakaran for technical assistance, and Conte Summer Undergraduate Program (C-SURE) students Joanna Kass, Lauren Carpenter, and Phillip Dmitriev for assistance with data analysis.

ADDITIONAL INFORMATION

Supplementary Information accompanies this paper at (<https://doi.org/10.1038/s41386-019-0421-9>).

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

REFERENCES

1. Bottlender R, Strauss A, Moller HJ. Social disability in schizophrenic, schizoaffective and affective disorders 15 years after first admission. *Schizophr Res.* 2010;116:9–15.
2. Marder SR, Galderisi S. The current conceptualization of negative symptoms in schizophrenia. *World Psychiatry.* 2017;16:14–24.
3. Kirkpatrick B, Fenton WS, Carpenter WT Jr., Marder SR. The NIMH-MATRICES consensus statement on negative symptoms. *Schizophr Bull.* 2006;32:214–9.
4. Pinkham AE, Harvey PD, Penn DL. Paranoid individuals with schizophrenia show greater social cognitive bias and worse social functioning than non-paranoid individuals with schizophrenia. *Schizophr Res Cogn.* 2016;3:33–38.
5. Green MF, Horan WP, Lee J. Social cognition in schizophrenia. *Nat Rev Neurosci.* 2015;16:620–31.
6. Ballon JS, Kaur T, Marks II, Cadenhead KS. Social functioning in young people at risk for schizophrenia. *Psychiatry Res.* 2007;151:29–35.
7. Lyne J, O'Donoghue B, Owens E, Renwick L, Madigan K, Kinsella A, et al. Prevalence of item level negative symptoms in first episode psychosis diagnoses. *Schizophr Res.* 2012;135:128–33.
8. Yung AR. Toward improved risk prediction in individuals at high risk of psychotic disorders. *Am J Psychiatry.* 2015;172:932–3.
9. Porcelli S, Van Der Wee N, van der Werff S, Aghajani M, Glennon JC, van Heukelum S, et al. Social brain, social dysfunction and social withdrawal. *Neurosci Biobehav Rev.* September 2018. <https://doi.org/10.1016/j.neubiorev.2018.09.012>.
10. Anticevic A, Van Snellenberg JX, Cohen RE, Repovs G, Dowd EC, Barch DM. Amygdala recruitment in schizophrenia in response to aversive emotional material: a meta-analysis of neuroimaging studies. *Schizophr Bull.* 2012; 38:608–21.
11. Gur RE, Loughhead J, Kohler CG, Elliott MA, Lesko K, Ruparel K, et al. Limbic activation associated with misidentification of fearful faces and flat affect in schizophrenia. *Arch Gen Psychiatry.* 2007;64:1356–66.
12. Pinkham AE, Gur RE, Gur RC. Affect recognition deficits in schizophrenia: neural substrates and psychopharmacological implications. *Expert Rev Neurother.* 2007;7:807–16.

13. Phelps EA, LeDoux JE. Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron*. 2005;48:175–87.
14. Finlayson K, Lampe JF, Hintze S, Wurbel H, Melotti L. Facial indicators of positive emotions in rats. *PLoS ONE*. 2016;11:e0166446.
15. Vuilleumier P, Pourtois G. Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia*. 2007;45:174–94.
16. Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. *Trends Cogn Sci*. 2000;4:223–33.
17. Fullana MA, Harrison BJ, Soriano-Mas C, Vervliet B, Cardoner N, Avila-Parcet A, et al. Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies. *Mol Psychiatry*. 2016;21:500–8.
18. Harrison BJ, Fullana MA, Via E, Soriano-Mas C, Vervliet B, Martinez-Zalacain I, et al. Human ventromedial prefrontal cortex and the positive affective processing of safety signals. *Neuroimage*. 2017;152:12–18.
19. Jensen J, Willeit M, Zipursky RB, Savina I, Smith AJ, Menon M, et al. The formation of abnormal associations in schizophrenia: neural and behavioral evidence. *Neuropsychopharmacology*. 2008;33:473–9.
20. Holt DJ, Coombs G, Zeidan MA, Goff DC, Milad MR. Failure of neural responses to safety cues in schizophrenia. *Arch Gen Psychiatry*. 2012;69:893–903.
21. Hofer E, Doby D, Anderer P, Dantendorfer K. Impaired conditional discrimination learning in schizophrenia. *Schizophr Res*. 2001;51:127–36.
22. Romaniuk L, Honey GD, King JR, Whalley HC, McIntosh AM, Levita L, et al. Midbrain activation during Pavlovian conditioning and delusional symptoms in schizophrenia. *Arch Gen Psychiatry*. 2010;67:1246–54.
23. Holt DJ, Lebron-Milad K, Milad MR, Rauch SL, Pitman RK, Orr SP, et al. Extinction memory is impaired in schizophrenia. *Biol Psychiatry*. 2009;65:455–63.
24. Linnman C, Coombs G III, Goff DC, Holt DJ. Lack of insula reactivity to aversive stimuli in schizophrenia. *Schizophr Res*. 2013;143:150–7.
25. Wolf DH, Satterthwaite TD, Calkins ME, Ruparel K, Elliott MA, Hopson RD, et al. Functional neuroimaging abnormalities in youth with psychosis spectrum symptoms. *JAMA Psychiatry*. 2015;72:456–65.
26. Bourque J, Spechler PA, Potvin S, Whelan R, Banaschewski T, Bokde ALW, et al. Functional neuroimaging predictors of self-reported psychotic symptoms in adolescents. *Am J Psychiatry*. 2017;174:566–75.
27. van Os J, Linscott RJ, Myin-Germeys I, Delespaul P, Krabbendam L. A systematic review and meta-analysis of the psychosis continuum: evidence for a psychosis proneness-persistence-impairment model of psychotic disorder. *Psychol Med*. 2009;39:179–95.
28. Calkins ME, Moore TM, Merikangas KR, Burstein M, Satterthwaite TD, Bilker WB, et al. The psychosis spectrum in a young U.S. community sample: findings from the Philadelphia Neurodevelopmental Cohort. *World Psychiatry*. 2014;13:296–305.
29. Calkins ME, Merikangas KR, Moore TM, Burstein M, Behr MA, Satterthwaite TD, et al. The Philadelphia Neurodevelopmental Cohort: constructing a deep phenotyping collaborative. *J Child Psychol Psychiatry*. 2015;56:1356–69.
30. Watters AJ, Rupert PE, Wolf DH, Calkins ME, Gur RC, Gur RE, et al. Social aversive conditioning in youth at clinical high risk for psychosis and with psychosis: an ERP study. *Schizophr Res*. 2018;202:291–6.
31. Buchel C, Morris J, Dolan RJ, Friston KJ. Brain systems mediating aversive conditioning: an event-related fMRI study. *Neuron*. 1998;20:947–57.
32. Schiller D, Levy I, Niv Y, LeDoux JE, Phelps EA. From fear to safety and back: reversal of fear in the human brain. *J Neurosci*. 2008;28:11517–25.
33. Greco JA, Liberzon I. Neuroimaging of fear-associated learning. *Neuropsychopharmacology*. 2016;41:320–34.
34. Smith SM, Nichols TE. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*. 2009;44:83–98.
35. Eklund A, Nichols TE, Knutsson H. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci USA*. 2016;113:7900–5.
36. Pujara MS, Philippi CL, Motzkin JC, Baskaya MK, Koenigs M. Ventromedial prefrontal cortex damage is associated with decreased ventral striatum volume and response to reward. *J Neurosci*. 2016;36:5047–54.
37. Haber SN, Knutson B. The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*. 2010;35:4–26.
38. Motzkin JC, Philippi CL, Wolf RC, Baskaya MK, Koenigs M. Ventromedial prefrontal cortex is critical for the regulation of amygdala activity in humans. *Biol Psychiatry*. 2015;77:276–84.
39. Park IH, Park HJ, Chun JW, Kim EY, Kim JJ. Dysfunctional modulation of emotional interference in the medial prefrontal cortex in patients with schizophrenia. *Neurosci Lett*. 2008;440:119–24.
40. Brent BK, Thermenos HW, Keshavan MS, Seidman LJ. Gray matter alterations in schizophrenia high-risk youth and early-onset schizophrenia: a review of structural MRI findings. *Child Adolesc Psychiatr Clin N Am*. 2013;22:689–714.
41. Koutsouleris N, Patschurek-Kliche K, Scheuerecker J, Decker P, Bottlender R, Schmitt G, et al. Neuroanatomical correlates of executive dysfunction in the at-risk mental state for psychosis. *Schizophr Res*. 2010;123:160–74.
42. Lataster J, Collip D, Ceccarini J, Hernaus D, Haas D, Booij L, et al. Familial liability to psychosis is associated with attenuated dopamine stress signaling in ventromedial prefrontal cortex. *Schizophr Bull*. 2014;40:66–77.
43. LeDoux J. The emotional brain, fear, and the amygdala. *Cell Mol Neurobiol*. 2003;23:727–38.
44. Sehlmeier C, Schoning S, Zwitserlood P, Pfliederer B, Kircher T, Arolt V, et al. Human fear conditioning and extinction in neuroimaging: a systematic review. *PLoS ONE*. 2009;4:e5865.
45. Zald DH. The human amygdala and the emotional evaluation of sensory stimuli. *Brain Res Brain Res Rev*. 2003;41:88–123.
46. Hrybowski S, Aghamohammadi-Sereshtki A, Madan CR, Shafer AT, Baron CA, Seres P, et al. Amygdala subnuclei response and connectivity during emotional processing. *Neuroimage*. 2016;133:98–110.
47. Kim J, Pignatelli M, Xu S, Itohara S, Tonegawa S. Antagonistic negative and positive neurons of the basolateral amygdala. *Nat Neurosci*. 2016;19:1636–46.
48. Boll S, Gamer M, Gluth S, Finsterbusch J, Buchel C. Separate amygdala subregions signal surprise and predictiveness during associative fear learning in humans. *Eur J Neurosci*. 2013;37:758–67.
49. Wolf DH, Satterthwaite TD, Loughhead J, Pinkham A, Overton E, Elliott MA, et al. Amygdala abnormalities in first-degree relatives of individuals with schizophrenia unmasked by benzodiazepine challenge. *Psychopharmacology (Berl)*. 2011;218:503–12.
50. Roalf DR, Nanga RP, Rupert PE, Hariharan H, Quarmley M, Calkins ME, et al. Glutamate imaging (GluCEST) reveals lower brain GluCEST contrast in patients on the psychosis spectrum. *Mol Psychiatry*. 2017;22:1298–305.