Research Paper

Brain-based Classification of Negative Social Bias in Adolescents With Nonsuicidal Self-injury: Findings From Simulated Online Social Interaction

Irene Perini a,⁎, Per A. Gustafsson a,b, J. Paul Hamilton a, R. Kämpe a, Leah M. Mayo a, Markus Heilig a,c,1, Maria Zetterqvist a,b,1

a Center for Social and Affective Neuroscience, Department of Clinical and Experimental Medicine, Linköping University, 581 85 Linköping, Sweden
b Department of Child and Adolescent Psychiatry, Region Östergötland, 581 85 Linköping, Sweden
c Department of Psychiatry, Region Östergötland, 581 85 Linköping, Sweden

ABSTRACT

Background: Interpersonal stress and perceived rejection have been clinically observed as common triggers of nonsuicidal self-injury (NSSI), with self-injury behavior regulating both affective and social experiences. We investigated whether the subjective interpretation of social interaction in a simulated online environment might be biased in the NSSI group, and the brain mechanisms underlying the experience.

Methods: Thirty female adolescent patients with NSSI and thirty female age-matched controls were investigated in this case–control study. In our novel task that simulates interaction on current social media platforms, participants indicated whether they liked or disliked pictures of other players during a functional magnetic resonance imaging (fMRI) scan. Participants also viewed positive and negative feedback directed toward them by others. The task also assessed the subjective effects of the social interaction. Finally, subjects underwent a separate facial electromyography session, which measured facial expressions processing.

Outcomes: Behavioral, the NSSI group showed a negative bias in processing social feedback from others. A multi-voxel pattern analysis (MVPA) identified brain regions that robustly classified NSSI subjects and controls.

Regions in which mutual activity contributed to the classification included dorsomedial prefrontal cortex and subgenual anterior cingulate cortex, a region implicated in mood control. In the NSSI group, multi-voxel classification scores correlated with behavioral sensitivity to negative feedback from others. Results remained significant after controlling for medication, symptoms of depression, and symptoms of borderline personality disorder.

Interpretation: This study identified behavioral and neural signatures of adolescents with NSSI during social interaction in a simulated social media environment. These findings highlight the importance of understanding social information processing in this clinical population and can potentially advance treatment approaches.

© 2019 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Nonsuicidal self-injury (NSSI) is defined as the direct, deliberate destruction of one’s own body tissue without suicidal intent [1], typically including behaviors such as cutting, burning, or hitting oneself. The risk of engaging in NSSI is particularly high during adolescence, with prevalence rates around 17% in community samples [2] and between 40 and 80% in clinical samples [3]. NSSI is more common in females, especially in clinical samples [4]. The behavior is often related to distress and/or functional impairment, and can occur together with or independently of psychiatric diagnoses, including depression and anxiety [5]. Diagnostically, it is currently a symptom of borderline personality disorder (BPD) [6]. NSSI is far more prevalent than BPD in adolescents [2,2], suggesting that NSSI can exist independently of or co-occur with BPD. NSSI has been suggested as a diagnostic entity of its own and included in the third section of the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM–5) [6] as a condition requiring further study, highlighting the importance of research in this area.

Interpersonal stress, perceived criticism, and social rejection are common triggers of NSSI. Social interactions are essential for well-being across the lifespan, but particularly so during adolescence, when socialization into peer-groups increases the importance of social interactions [7]. The age of onset for NSSI is around 12–14 years [3] and peaks during adolescence, with rates declining in adulthood [8]. Sensitivity to interpersonal stress, high emotional distress and in some cases chronic romantic stress [9] have been suggested to play a role in the development and maintenance of NSSI, with self-injury behavior regulating both affective and social experiences [1]. The fact that...
Research in context

Evidence before this study

Although interpersonal stress has been clinically observed as a common trigger of nonsuicidal self-injury (NSSI), behavioral and neural mechanisms underlying social processing in adolescents with NSSI remain relatively unexplored in experimental settings. To identify previous studies which addressed behavioral and brain correlates to social processing in individuals with NSSI, we entered the following search items in PubMed: “NSSI”, “fMRI” and “social”. We found three articles, of which two included brain imaging studies. In the two brain imaging studies, which used the same NSSI adolescent sample, participants were selected based on the psychiatric diagnoses of depression with or without borderline personality disorder. In the current study, which provides a larger sample size than the aforementioned studies, we selectively included individuals with NSSI, independent of psychiatric diagnosis.

Added values of this study

In this study, individuals with NSSI experienced evaluative social feedback more negatively than controls following participation in a task which simulated online social interaction. In addition, multivoxel pattern analysis (MVPA) of neural-response data yielded a significant classification of individuals with NSSI and classification indices correlated selectively in individuals with NSSI with elevated sensitivity to negative social feedback. To our knowledge, this is the first study to identify significant behavioral and brain differences in processing social information in individuals with NSSI (independent of psychiatric diagnoses), compared to controls.

Implications of all the available evidence

Using an ecologically valid social interaction task, we provide novel behavioral and neural-functional insights into how individuals with NSSI experience social interactions.

interpersonal difficulties commonly precede NSSI has also been noted in the proposed DSM-5 criteria [6]. A role of interpersonal stress in triggering NSSI is supported by studies using semi-structured interviews and the proposed DSM-5 criteria[6]. A role of interpersonal stress in triggering NSSI is supported by studies using semi-structured interviews and the proposed DSM-5 criteria[6].

2. Materials and Methods

2.1. Participants

Thirty healthy controls (mean age, 16.4 years, SD 0.9) were recruited through advertisement in schools and on Facebook. Thirty clinical participants (mean age, 15.9 years, SD 0.8) were recruited from the child and adolescent psychiatric (CAP) clinic at Linköping University Hospital, Sweden. Participants completed two sessions on separate days, in a counterbalanced order. One session included an fMRI scan, and the other assessed behavior and facial electromyography (EMG). Three participants per group were excluded from the MRI session because of stress associated with the session or use of dental braces. Twenty-seven patients and 27 age-matched controls were thus included in the MRI analysis, and 30 patients and controls completed the facial EMG session.

Inclusion criteria for the clinical participants were: NSSI, independent of psychiatric diagnosis, being a female between 15 and 18 years, and having engaged in five or more instances of NSSI during the last six months. Exclusion criteria were: current or lifetime diagnosis of schizophrenia, bipolar or psychotic disorder and/or alcohol/drug dependence, and IQ below 80. Healthy controls were included if they had no DSM Axis I or II disorder during the last year and no lifetime presence of NSSI. Patients taking psychotropic medications were included provided that they were ongoing and unchanged for at least three months; sensitivity analyses were subsequently carried out to examine the potential confounding role of medication. Eleven patients were taking psychotropic medications (Table 1). Adolescents meeting inclusion criteria were approached with oral and written information about the study. Participants (and parents, if the participant was <18) gave written informed consent. Clinical recruitment occurred from June 2016 to March 2018. Controls and clinical participants did not differ in age, IQ or handedness. Additional information is presented in Table 1. The study was approved by the Regional Ethical Board of Linköping (Dnr 2015/273-31; 2016/224-32). To ensure that any adverse reactions, such as worry or anxiety, were managed according to clinical routine, all participants were accompanied to the sessions by the same clinician that had provided their initial diagnosis.
2.2. Psychometric Measures

We used the clinical interview Schedule for Affective Disorders and Schizophrenia for School-Age-Children-Present and Lifetime (K-SADS-PL) [16] for DSM-IV diagnoses; selected questions from the semi-structured Self-injurious thoughts and behaviors interview (SITBI [17, 18]) to obtain detailed information about suicidal behavior; the Structured Clinical Interview for DSM-IV Personality Disorders (SCID-II) [19] for symptoms of borderline personality disorder; Children’s Depressive Rating Scale – Revised (CDRS-R) [20] for depressive symptoms; an abbreviated version of Wechsler Intelligence Scales, fourth edition for children [21] or adults [22], depending on participants’ age, for intelligence. For the clinical sample only Clinical Assessment of Nonsuicidal Self-Injury Disorder (CANDI) [23] assessed NSSI characteristics (including frequency and means of NSSI) and whether NSSI participants met criteria for a diagnosis of NSSI disorder. Assessments were performed by the last author, a clinical psychologist with extensive experience in psychiatric assessment, together with the second author, a child psychiatrist. Final psychiatric diagnoses for the clinical sample were based on all available information from diagnostic interviews and medical records, using DSM-5 [6].

2.3. fMRI Task

In the scanner, subjects participated in a simulated online game aimed at identifying neural regions involved in the processing of self-relevant information during social interaction [14]. Briefly, participants were presented with pictures of other adolescents, and asked to indicate whether they liked or disliked the person. Similarly, participants viewed their own picture being judged by other simulated players. The event-related design consisted of trials comprising three epochs: question, anticipation and outcome. For full details on the design see Perini et al. [14].

Questionnaires were completed immediately following the MRI scan to assess subjective perception of the social interaction and its effects on

Table 1
Participant demographics.

| Demographic characteristics | NSSI (n = 26–30) | Healthy controls (n = 26–30) | Comparison statistic |
|-----------------------------|------------------|-------------------------------|---------------------|
| Age (m (sd))                | 15.9 (0.79)      | 16.4 (1.0)                    | n.s.                |
| IQ (m (sd))                 | 96.6 (9.83)      | 100.9 (10.94)                 | n.s.                |
| Sex                         | Female 30 (100%) | 30 (100%)                     |                     |
| Parental education          | University/college 23 (41.8%) | 32 (60.4%) | n.s. |
| Current family structure    | Married/co-habitant 12 (40.0%) | 18 (62.1%) | n.s. |
| Depressive symptoms (CDRS-R (m (sd)) | 45.7 (13.36) | 22.2 (4.95) | p < 0.001 Cohens d = 2.33 |
| DSM-5 NSSID diagnosis       | 18 (62.1%)       |                               |                     |
| Age of onset (m (sd))       | 13.2 (1.25)      |                               |                     |
| Number of methods (m (sd))  | 3.8 (2.13)       |                               |                     |
| Cutting frequency (12 months) (m (sd)) | 54.6 (55.7) | 3.5 (5.15) |                     |
| Latest NSSI episode (weeks) (m (sd)) | 3.5 (5.15) |                     |                     |
| Suicidal behaviors          | Suicidal ideation 30 (100%) |                     |                     |
| Ever inpatient psychiatric care | 7 (23.3%) |                     |                     |
| SCID-II self-report (m (sd)) | 6.00 (2.8) | 1.07 (1.5) | p < 0.001 Cohens d = 2.19 |
| SCID-II interview (m (sd))  | 3.23 (2.6)       | 0.13 (0.51)                  | p < 0.001 Cohens d = 1.65 |
| Psychiatric diagnoses*      | Depression 15 (50.0%) |                     |                     |
| Anxiety disorder            | 13 (43.3%)       |                               |                     |
| Posttraumatic stress disorder | 1 (3.3%)    |                               |                     |
| Borderline traits           | 13 (43.3%)       |                               |                     |
| Eating disorder             | 6 (20.0%)        |                               |                     |
| ADHD/ADD                    | 15 (50.0%)       |                               |                     |
| High functioning autism     | 4 (13.3%)        |                               |                     |
| ODD/CD                      | 3 (10.0%)        |                               |                     |
| Medications**               | SSRi/SNRI 8 (26.7%) |                     |                     |
| SSRI/SNRI + methylphenidate | 1 (3.3%)        |                               |                     |
| Neuroleptic                 | 1 (3.3%)         |                               |                     |
| SSRI/SNRI + neuroleptic     | 1 (3.3%)         |                               |                     |
| No medication               | 19 (63.3%)       |                               |                     |

Note. *each participant could have several diagnoses **medication at time of fMRI.
participants. Question one assessed the perceived ratio of negative feedback
(“How often were you disliked?” 0% “never”-100% “always”) when, in fact, the task was balanced so that the participants received equal amounts of positive and negative feedback. Question two asked:
“How much did you like to see your own face?” (0 “not at all”-10 “very much”). Finally, we investigated the emotional effects of being liked or disliked by others with the following questions: “How bad did it feel to be disliked?” (0 “not at all” – 10 “very much”) and “How good did it feel to be liked?” (0 “not at all” – 10 “very much”). Scores’ residuals were tested for normality. Depending on whether the assumption of normality was met or not, individual scores for each group were compared using respectively a univariate ANOVA or a Mann–Whitney test, respectively. Either Pearson’s or Spearman’s correlation coefficients were calculated depending on whether data met assumptions of normality. Behavioral statistics were calculated using the Statistical Package for the Social Sciences (SPSS) version 25.

2.4. Image Acquisition

Imaging was performed using a Philips Ingenia 3 Tesla MRI scanner (Philips Healthcare, Best, The Netherlands) equipped with a 32-channel Philips d5 Head head-coil. Six dummy volumes were acquired before each scan to allow the spin system to reach steady-state longitudinal magnetization and reduce possible effects of partial saturation. Blood oxygen-level-dependent (BOLD) data were acquired with an echo-planar imaging (EPI) sequence: TR = 2000 ms; TE = 30 ms; flip angle = 77°; field-of-view = 220 × 220 mm²; in-plane resolution = 3.4 × 3.4 mm; slice thickness = 4 mm, no slice gap; number of axial slices (angled with the AC-PC line) = 32; number of volumes = 195. Two functional runs were collected and each run lasted for 6 min and 45 s. A high-resolution 3D T1-weighted Turbo Field Echo scan was acquired before the EPI data acquisitions: TR = 7.0 ms; TE = 3.2 ms; flip angle = 8°; field-of-view = 256 × 254 × 170 mm; voxel resolution = 1 × 1 × 1 mm; no slice gap; plane: sagittal; number of sagittal slices = 170.

2.5. Image Preprocessing

Preprocessing was performed with the Analysis of Functional Neuro Images (AFNI) software v16.2.12 [24]. BOLD images were de-spiked and slice-time corrected. For motion correction and co-registration purposes, each EPI volume was registered to the volume with the minimum outlier fraction (using the AFNI outlier definition). Functional images were then warped to Talairach template space using a combination of affine and non-linear transformations [25]. Nuisance effects due to head motion (estimated from the motion correction procedure) were accounted for by adding the motion parameters (and their derivatives) as regressors of no interest in the main regression. A motion censoring threshold of 0.3 mm per TR was implemented in combination with an outlier fraction threshold of 0.1. Volumes violating either of these thresholds were subsequently ignored in the time-series regression. There was no statistical difference in the amount of volumes ignored in the two groups (t = -1.1, p = 0.2, two-tailed). For the participants with volume censoring over 30% of the total volume number, we found no significant difference in the censoring per condition for either anticipation (p = 0.1) and outcome (p = 0.7) intervals. Table S1 shows percentage of removed volumes for each participant.

2.6. Univariate Analysis

A univariate general linear model (GLM) analysis described elsewhere [14] was initially performed to identify brain correlates of self-relevant processing and to assess replication of previous findings [14]. The AFNI program 3dClustSim was used to determine cluster-size thresholds necessary for identifying effects significant at alpha = 0.05, family-wise-error corrected (per voxel p < 0.002, two-sided; cluster threshold = 30 voxels), according to current recommendations by AFNI developers [26,27]. Average spatial smoothness estimates, across all participants, and entered into 3dClustSim were obtained using the 3dFWHMx function with the ACF option.

2.7. Multi-voxel Pattern Analysis

We applied a multivariate machine-learning approach, first, to determine whether brain activity during the online-game task could be used to classify group membership and, second, to identify brain regions that most significantly contributed to multivariate discrimination of groups. We applied a support vector machine (SVM) to voxel-wise indices of task-activation using in-house software developed in MATLAB 10.6 (www.mathworks.com) calling both MATLAB and AFNI functions [24].

The machine-learning approach was applied, voxel-wise, to neural-response indices derived from traditional GLM “task-based” analysis ap- proaches that were applied to BOLD time-series data from both anticipation and outcome epochs. We used leave-one-out cross validation to determine the diagnostic accuracy of the approach. With this method, all participants but one is used to train the classifier, and the remaining subject’s data is used to test the classifier. We used a simple, linear-kernel SVM with margin softness parameter, c, set to 1. We used random permutation testing (1000 permutations) to determine the significance of classifier accuracy achieved. To depict classifier performance, a receiver operating characteristic (ROC) curve was generated in SPSS using classification scores. To identify brain regions that reliably contributed to discrimination between healthy and NSSI, we performed the following steps: The resulting brain maps from the permutation testing were transformed into z-scores by subtracting the average permutation weight scores from the classification weight scores and dividing the resulting difference by the standard deviation of the permutation weight scores. We used AFNI’s 3dClustSim (spatial smoothness estimated with 3dFWHMx, ACF option) to determine cluster-size thresholds necessary for identifying effects significant at per-voxel p = 0.002 (two-sided; z = 2.89), family-wise-error corrected at alpha = 0.05 (cluster size = 6). We conducted this MVPA approach twice, once on voxels from a whole-brain, grey-matter mask and, again, using this same mask but without regions activated for the self-versus-other contrast during anticipation (i.e. anterior insula, dorsal anterior cingulate cortex, and supplementary motor area). We did this to determine if brain regions that are “silent” in the univariate, task-activation sense might, at a multivariate level, still effectively discriminate disordered from healthy samples.

Finally, to evaluate the clinical relevance of our machine-learning approach, we examined in the NSSI group the association between sensitivity to rejection as measured by post-game self-report and the classification score generated for each member of the NSSI group (i.e. Euclidian norm distance between individual-subject data and the trained support vector) when their data were classified during leave-one-out cross-validation. We used partial correlations to estimate the association between the MVPA classification score and rejection sensitivity controlling for BPD traits, assessed using the SCID II interview, depression scores, assessed using the CDRS-R scale, and medication use.

2.8. Facial Electromyography

Emotions sequences included happiness, sadness, surprise, fear, anger, and disgust comprising 15 static morphed images displayed for 200 ms each that conveyed from 5% emotion to 100% emotion. Stimuli were created as previously described [28]. Briefly, young men and women (n = 12 each gender) posed each emotion and were photographed; composite images of each of the six emotions were created for each gender [29,30]. Each gender-by-emotion sequence was displayed four times for a total of 48 trials presented in pseudorandom order. Participants were instructed to press a button as soon as they
could identify the emotion (‘sensitivity’) and then given a list of all six emotions to choose from to identify the emotion (‘accuracy’). Facial EMG recordings of the corrugator and zygomatic muscles were obtained simultaneously.

Recordings were obtained as previously described [31]. Briefly, sensors consisted of bipolar 4 mm silver/silver chloride electrodes filled with electrode gel placed on the left side of the face and a ground electrode on the forehead. EMG signals were amplified, filtered through a 10–500 Hz band pass, digitized at 1 kHz, rectified, and integrated over 20 ms using the MP150 Data Acquisition system, and Acqknowledge software from Biopac Systems (Biopac Systems Incorporated, Camino Goleta, CA, USA). To control for the varied duration of trial length (e.g. how long it took the participant to identify the emotion), we analyzed the mean EMG activity during the final 1000 ms of the stimulus display compared to the mean activity during the 1000-millisecond baseline immediately prior to trial onset [32].

Behavioral outcomes (sensitivity, accuracy) and mean facial EMG activity (corrugator, zygomatic) were analyzed using a repeated-measures analysis of variance (ANOVA) with the within-subject factor of emotion and the between-subject factor of group. A secondary analysis assessed reactions to each emotion individually, accounting for the emotional intensity at the time of trial termination [29].

2.9. Data Statement
We do not have participants’ consent to share raw data publicly. The data that has been used is confidential.

3. Results

3.1. Behavioral Findings

Based on post-scan queries, NSSI patients felt rejected significantly more often than controls ([U = 172, p = 0·009, r = 0·37]; Fig. 1) and, at trend level, also disliked others, more often than controls (F (0,52) = 11·5, p = 0·08, η2p = 0·05). Compared to controls, NSSI individuals were also significantly more sensitive to being rejected and disliked significantly more to see their own face during the game (sensitivity to rejection $F(0,47) = 11·5, p = 0·001, η^2_p = 0·19$; face likeness $U = 16, p < 0·001, r = 0·82$). In contrast, we did not observe a group difference with regard to the positive effects of being liked by others ($F(0,47) = 180·3, p = 0·78, η^2_p = 0·002$; Fig. 1).

We observed significant correlations between NSSI characteristics and behavioral data. The average cutting frequency, which reflected the number of cuts during the past 12 months was 54.6 (SD 55.7). One patient was an extreme outlier, with an estimated cutting frequency of 400, and was therefore removed from the analysis. Cutting frequency was negatively correlated with how much participants liked to see their own face ($r_s(23) = −0·47, p = 0·024$) and positively...
correlated to perceived rejection scores ($r_c(23) = 0.46, p = 0.026$).

Further, the recency of the latest NSSI episode was negatively correlated with perceived rejection, indicating that the more recent the NSSI episode, the more the patient felt rejected ($r_c(24) = -0.48, p = 0.016$).

The pattern and significance of results did not change when excluding NSSI patients on medication (see Supplemental data). Overall, our behavioral findings suggest that NSSI patients have a negative-interpretation bias for social evaluation, and when presented with negative feedback, are more sensitive to it than controls. This bias correlates with a critically important clinical parameter, the frequency of NSSI events.

### 3.2. Univariate Analysis: Brain Correlates to Self-relevant Evaluation

We previously reported that right anterior insula and dorsal anterior cingulate cortex (rAI and dACC) are key regions for processing self-relevant information in healthy adolescents [14]. These findings replicated in both healthy and NSSI individuals in the present study. For the main effect of Perspective, clusters in bilateral AI and ACC merging with orbitofrontal cortex, together with a cluster in the visual cortex were significantly activated ($p < 0.002$, family-wise-error corrected; Fig. 2a, Table 2). No between-group differences were identified in this analysis (Table 2). β values within the right AI, showed a significant effect of Perspective with "self" condition values higher than "other" condition values ($F(1,52) = 51.3, p < 0.001, \eta^2_p = 0.49$). As significant effect of outcome ($F(1,52) = 8.73, p = 0.003, \eta^2_p = 0.14$) driven by higher values for rejection ($t = 2.9, p = 0.003$) and a perspective-by-outcome interaction was observed ($F(1,52) = 9.3, p = 0.001, \eta^2_p = 0.15$), with the self-rejection condition significantly higher than self-acceptance ($t = 4.3, p < 0.001$; Fig. 2b). β values extracted from the ACC showed higher values for "self" compared to "other" condition ($F(1,52) = 18.9, p < 0.001, \eta^2_p = 0.26$; Fig. 2b).

**Table 2**

| Analysis         | Region                      | Talairach coordinates | Voxels |
|------------------|-----------------------------|------------------------|--------|
| Anticipation interval | Self > other | Supplementary motor area (peak) | 2 2 53 | 231 |
|                  | Dorsal anterior cingulate cortex | -3 9 51 |        |
|                  | Postcentral gyrus           | -4 -14 38 |        |
|                  | Anterior insula             | 35 17 8 93 |        |
|                  |                             | -31 14 11 69 |        |
| Outcome interval | Self > other | Inferior frontal gyrus (peak) | 44 14 -10 | 210 |
|                  | Right anterior insula       | 31 18 5 |        |
|                  | Right mid-anterior insula   | 41 9 1 |        |
|                  | Rostral anterior cingulate cortex (peak) | 2 35 11 | 46 |
|                  | Dorsal anterior cingulate cortex | 2 15 11 |        |
|                  | Inferior frontal gyrus (peak) | -28 11 -16 | 92 |
|                  | Left anterior insula        | 31 17 5 |        |
|                  | Left mid-anterior insula    | 37 11 5 |        |
|                  | Fusiform gyrus              | 29 -43 -13 | 39 |
|                  | Middle occipital gyrus      | -28 -64 2 | 69 |

For the anticipation interval, the whole-brain, grey-matter analysis showed activation of the rAI and the dACC and supplementary motor area (SMA) bilaterally for the "self" versus "other" conditions contrast in...
both groups ($p < 0.002$, family-wise-error corrected; Fig. 3c; Table 2). Again, no significant between-group effects were observed (Table 2).

### 3.3. Multivariate Analysis: Neural Response During Anticipation of Self-relevant Evaluation

The whole-brain, grey-matter, multivariate analysis rendered statistically-significant classification of subjects during anticipation, with accuracy of 68% (the percentage of patients that were correctly classified), sensitivity = 0.74, specificity = 0.59 (Fig. 3a); permutation $p = 0.031$; AUC ROC = 0.77, $p = 0.001$ (Fig. 3b). Classification scores remained significantly different between groups after excluding NSSI patients on medication ($U = 117, p = 0.007, r = 0.32$). Classification scores derived from the multivariate analysis, significantly correlated with rejection sensitivity scores in the NSSI patients but not in controls [patients $r_c(25) = 0.42, p = 0.04$, controls $r_c(23) = 0.12, p = 0.57$; $r$ scores comparison not significant $p = 0.1$, two-tails] (Fig. 3d, circles). The correlation remained significant after controlling for BDP traits $r_c(22) = 0.46, p = 0.02$, depression traits $r_c(20) = 0.47, p = 0.03$, and medication use $r_c(22) = 0.42, p = 0.04$.

Brain regions significantly contributing to the discrimination between groups included dorsomedial prefrontal cortex (dmPFC; three clusters), posterior cingulate cortex (PCC) and subgenual anterior cingulate cortex (sgACC) (Fig. 3c; Table 3). Excluding from the analysis mask regions activated for the self-versus-other contrast during anticipation yielded the same outcome, with classification accuracy = 68% and $p = 0.029$. Classification scores significantly correlated with rejection sensitivity [patients $r(21) = 0.42, p = 0.04$ (Fig. 3d, triangles)]. To investigate the composition of the multivariate effects we extracted $\beta$-coefficients from the aforementioned regions and compared them between groups (see Fig. S1). Average-$\beta$ scores for all regions were significantly lower in NSSI [sgACC ($U = 221, p = 0.005$), PCC ($U = 253, p = 0.02$), dmPFC1 ($U = 172, p < 0.001$), dmPFC2 ($U = 225, p = 0.006$), dmPFC ($U = 227, p = 0.007$)] (Fig. S1).

The multivariate analyses targeting the outcome intervals did not yield above-chance classification accuracy scores (outcome self-rejection accuracy = 46%, outcome self-acceptance accuracy = 31%).

### 3.4. Facial EMG Findings

Behaviorally, there was a main effect of emotion on reaction time ($F(5,270) = 11.7, p < 0.001$) and accuracy ($F(5,285) = 29.2, p < 0.001$), but this did not differ between patients and controls (Reaction time: group, $p = 0.37$; group*emotion, $p = 0.83$; Accuracy: group, $p = 0.38$; emotion* group, $p = 0.32$). Specifically, all participants were less sensitive (e.g. slower) and accurate to detect fear ($p < 0.05$ versus other emotion, Bonferroni corrected).

There was a main effect of emotion on corrugator reactivity such that faces conveying negative emotions (fear, anger, disgust) elicited an increase in corrugator reactivity, while happy faces elicited a reduction in activity ($F(5,275) = 26.9, p < 0.001$). There was no effect of group ($p = 0.76$) or a group-by-emotion interaction ($p = 0.80$). Similarly, happy faces elicited significantly more zygomatic activity than surprise or the negative emotions ($F(5,275) = 5.99, p < 0.001$), but these...
findings were consistent across patients and controls (group, $p = 0.95$; emotion$x$group, $p = 0.17$).

When including sensitivity (e.g. reaction time) in the analysis, we still found no effect of group on corrugator reactivity (happy: $p = 0.73$; surprise: $p = 0.74$; sad: $p = 0.92$; fear: $p = 0.78$; anger: $p = 0.43$; disgust: $p = 0.57$) or zygomatic reactivity (happy: $p = 0.35$; surprise: $p = 0.30$; sad: $p = 0.22$; fear: $p = 0.27$; anger: $p = 0.82$; disgust: $p = 0.32$).

### 4. Discussion

This study is to our knowledge the first to examine neural functioning during a simulated online interaction in NSSI CAP patients. We assessed the largest sample of adolescents with NSSI up to date, both behaviorally and neurally, to characterize processing of social stimuli. We found that adolescents with NSSI showed a negative bias in interpreting evaluative social feedback from others. We assessed the largest sample of adolescents with NSSI up to date, both being during a simulated online interaction in NSSI CAP patients. We first to examine neural functioning during a simulated online interaction in NSSI CAP patients. We assessed the largest sample of adolescents with NSSI up to date, both behaviorally and neurally, to characterize processing of social stimuli. We found that adolescents with NSSI showed a negative bias in interpreting evaluative social feedback from others.

We found that this bias was associated with distinct neural-response patterns to social evaluation, as identified by multivoxel pattern analysis. The clinical relevance of the negative social-evaluation bias we observed is indicated by its correlation with NSSI. We found that adolescents with NSSI showed a negative bias in interpreting feedback. This bias was assessed during a simulated online interaction in NSSI CAP patients. We found that this bias was as-
A univariate approach can fail to identify potentially relevant information from spatially distributed effects [39]. Using a multivariate approach, we identified a pattern of brain regions in which activity during anticipation of judgment by others robustly classified subjects into patient and control groups. Most interestingly, the identified regions, which included sgACC, dmPFC and PCC are outside the salience network. The sgACC is considered to be involved in the generation of affective states and has been shown to be anatomically and functionally altered in psychiatric conditions, in particular in mood disorders [40]. The other regions, which include PCC and portions of dmPFC have been shown to be involved with self-referential processing and autobiographical memory [41]. A possible interpretation of the pattern that emerges is that attribution of salience to the self-referential social feedback stimuli by the ACC and insular cortex is not altered in NSSI, but that activity of other brain areas subsequently results in a negative interpretation bias in the patient group.

The validity and clinical relevance of our multi-variate approach is supported by the finding that multivariate classification scores correlated significantly with rejection sensitivity, suggesting a potential neurobiological basis for the reported behavioral differences in response to social feedback. Importantly, this correlation remained significant after controlling for symptoms of BPD, a condition in which great efforts are made to avoid perceived or real abandonment [6], and also after controlling for symptoms of depression, where negative self-evaluation is pronounced [6]. In our view, this finding provides preliminary support for the validity and utility of NSSI as an independent diagnostic entity.

There are limitations to our study. As is the case in all cross-sectional fMRI studies, our ability to infer a causal relationship between any of the brain responses and the behavioral findings is limited. Finally, because our finding is the first of its kind, our data need to be replicated in an independent sample. Nevertheless, our results have several potentially important clinical implications. Although the identified behavioral results are task-specific and were not investigated using a standardized self-report, they point to a vulnerability and negative affective bias that requires emphasis in clinical practice. Treatment approaches that make adolescents with NSSI aware of the existence of their affective biases can help making social interactions less painful, and promote more favorable interpretations of social feedback, thus reducing a potential trigger of NSSI. Some of the measures employed in our study have the potential to serve as biomarkers of therapeutic response to interventions, and future research should be directed to exploring this potential.

In summary, this study used a simulation of online social interaction to provide novel insight into the behavioral and neurological mechanisms of NSSI. The findings from this ecologically valid paradigm can potentially advance effective treatment, such as behavioral training using multivariate real-time fMRI neurofeedback.

Author Contributions

I.P., M.Z., P.G., M.H. designed research; I.P., R.K., M.Z., L.M. performed research; I.P., R.K. P.H., L.M. analyzed data; I.P. made figures; I.P., P.H., R.K., M.Z., L.M., P.G, M.H., wrote the paper.

Declaration of Competing Interest

The authors declare no competing interests.

Acknowledgements

The Authors would like to thank professor Martin Paulus for thoughtful contribution to the study design. This research was supported by The Swedish Research Council (538-2013-7434) and the ALF Grants, Östergötland County (LIO-53931; LIO-520131). The authors gratefully acknowledge staff at the Center for Medical Imaging Analysis and Visualization (CMIV) and the Child- and Adolescent Psychiatric clinic, Linköping University Hospital, Sweden.

Appendix A. Supplementary Data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eclinm.2019.06.016.

References

[1] Nock MK. Self-injury. Annu Rev Clin Psychol 2010;6:339–63.
[2] Muehlenkamp JJ, Claes L, Havertape L, Plener PL. International prevalence of adolescent non-suicidal self-injury and deliberate self-harm. Child Adolesc Psychiatry Ment Health 2012;6:10.
[3] Klonsky ED, Muehlenkamp JJ. Self-injury: a research review for the practitioner. J Clin Psychol 2007;63(11):1045–56.
[4] Bresin K, Schoenleber M. Gender differences in the prevalence of nonsuicidal self-injury: a meta-analysis. Clin Psychol Rev 2015;38:55–64.
[5] Klekens GA-O, Hasing P, Claes L, et al. The DSM-5 nonsuicidal self-injury disorder among incoming college students: prevalence and associations with 12-month mental disorders and suicidal thoughts and behaviors. Depress Anxiety 2018;35 (7):629–37.
[6] APA. Diagnostic and statistical manual of mental disorders. 5th ed.Washington DC: American Psychiatric Publishing; 2013.

[7] Steinberg L. Cognitive and affective development in adolescence. Trends Cogn Sci 2005;9(2):69–74.

[8] Moran P, Coffey C, Fau-Romanu i, H, Romanu i H, Fau-Olsson C, et al. The natural history of self-harm from adolescence to young adulthood: a population-based cohort study. Lancet 2012;379:236–43.

[9] Miller AB, Linthicum KP, Helms SW, et al. Reciprocal associations between adolescent girls’ chronic interpersonal stress and nonsuicidal self-injury: a multi-wave prospective investigation. J Adolesc Health 2018;63:694–700.

[10] Kim KL, Cushman GK, Weissman AB, et al. Behavioral and emotional responses to interpersonal stress: a comparison of adolescents engaged in non-suicidal self-injury to adolescent suicide attempters. Psychiatry Res 2015;228:899–906.

[11] Santangelo PS, Koenig J, Funke V, et al. Ecological momentary assessment of affective experiences of adverse life events and trauma symptoms. J Youth Adolesc 2013;42:1997–2009.

[12] Santangelo PS, Koenig J, Funke V, et al. Ecological momentary assessment of affective experiences of adverse life events and trauma symptoms. J Youth Adolesc 2013;42:1997–2009.

[13] Steinberg L. Cognitive and affective development in adolescence. Trends Cogn Sci 2005;9(2):69–74.

[14] Nock MK, Prinstein MJ, Sterba SK. Revealing the form and function of self-injurious thoughts and behaviors: a real-time ecological assessment study among adolescents and young adults. J Abnorm Psychol 2009;118:816–27.

[15] Brown RC, Plener PL, Groen G, Neff D, Bonenberger M, Abler B. Differential neural processing of social exclusion and inclusion in adolescents with non-suicidal self-injury and young adults with borderline personality disorder. Front Psych 2017;8:267.

[16] Perini I, Gustafsson PA, Hamilton JP, Kampe R, Zetterqvist M, Heilig M. The salience of stress: a comparison of adolescents engaged in non-suicidal self-injury to adolescent suicide attempters. Psychiatry Res 2015;228:899–906.

[20] Poznanski E, Mokros H. Children’s depression rating scale. 2019.03.

[20] Poznanski E, Mokros H. Children’s depression rating scale. 2019.03.

[21] Talairach J, Tournoux P. Co-planar stereotactic atlas of the human brain. New York: Thieme Medical Publishers; 1988.

[22] APA. Diagnostic and statistical manual of mental disorders. 5th ed.Washington DC: American Psychiatric Publishing; 2013.

[23] Cox RW, AfNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 1996;29(1):162–73.

[24] Cox RW, AfNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 1996;29(1):162–73.

[25] Talairach J, Tournoux P. Co-planar stereotactic atlas of the human brain. New York: Thieme Medical Publishers; 1988.

[26] Cox RW, Chen G, Glen DR, Reynolds RC, Taylor PA. FMRI clustering and false-positive rates redux. Brain Connect 2017;7;3:152–71.

[27] Cox RW, Chen G, Glen DR, Reynolds RC, Taylor PA. FMRI clustering and false-positive rates. Proc Natl Acad Sci U S A 2017;114(17):E3370–E1.

[28] Bamford S, Penton-Voak I, Pinkney V, Baldwin DS, Munafò MR, Garner M. Early effects of duloxetine on emotion recognition in healthy volunteers. J Psychopharmacol (Oxford, England) 2015;29(5):534–41.

[29] Tiddeman B, Burt M, Perrett D. Prototyping and transforming facial textures for perception research. IEEE Comput Graph Appl 2001;21(5):42–50.

[30] Skinner AL, Benton CP. Anti-expression aftereffects reveal prototype-referenced coding of facial expressions. Psychol Sci 2010;21(9):1248–53.

[31] Mayo LM, de Wit H. Acquisition of conditioned responses to a novel alcohol-paired cue in social drinkers. J Stud Alcohol Drugs 2016;77(2):317–26.

[32] Wardle MC, de Wit H, Penton-Voak I, Lewis G, Munafò MR. Lack of association between COMT and working memory in a population-based cohort of healthy young adults. Neuropsychopharmacology 2013;38:1253–63.

[33] Prinstein MJ, Guerry JD, Browne CB, Rancourt D. Interpersonal models of nonsuicidal self-injury. In: Nock MK, editor. Understanding nonsuicidal self-injury origins, assessment, and treatment. Washington D.C.: American Psychological Association; 2009. p. 79–98.

[34] Sherman LE, Hernandez LM, Greenfield PM, Dapretto M. What the brain ‘likes’: neural correlates of providing feedback on social media, vol. 13; 2018; 669–707.

[35] Sherman LE, Payne AO, Hernandez LM, Greenfield PM, Dapretto M. The power of the like in adolescence: effects of peer influence on neural and behavioral responses to social media. Psychol Sci 2016;27(7):1027–35.

[36] Labar KS, Crupain MJ Fau-Voyvodic JF, Voyvodic JF Fau-McCarthy G, McCarthy G. Dynamic perception of facial affect and identity in the human brain. Cereb Cortex 2003; 13:1023–33.

[37] Seymour KE, Jones RN, Cushman GK, et al. Emotional face recognition in adolescent suicide attempters and adolescents engaging in non-suicidal self-injury. Eur Child Adolesc Psychiatry 2016;25:247–59.

[38] Neta M, Norris CJ Fau-Whalen PJ, Whalen PJ. Corrugator muscle responses are associated with individual differences in positivity-negativity bias. 2009;9:640–8.

[39] Sherman LE, Hernandez LM, Greenfield PM, Dapretto M. The power of the like in adolescence: effects of peer influence on neural and behavioral responses to social media. Psychol Sci 2016;27(7):1027–35.

[40] Drevets WC, Savitz J, Trimble M. The subgenual anterior cingulate cortex in mood disorders. CNS Spectr 2008;13:1023–33.

[41] Spreng RN, Mar Ra Fau-Kim ASN, Kim AS. The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. 2008;3:489–510.