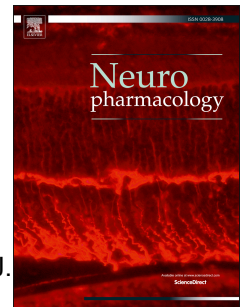


Accepted Manuscript

Habenula-prefrontal resting-state connectivity in reactive aggressive men – A pilot study

G. Gan, A. Zilverstand, M.A. Parvaz, R.N. Preston-Campbell, F. d'Oleire Uquillas, S.J. Moeller, D. Tomasi, R.Z. Goldstein, N. Alia-Klein



PII: S0028-3908(18)30812-8

DOI: <https://doi.org/10.1016/j.neuropharm.2018.10.025>

Reference: NP 7396

To appear in: *Neuropharmacology*

Received Date: 5 June 2018

Revised Date: 25 September 2018

Accepted Date: 17 October 2018

Please cite this article as: Gan, G., Zilverstand, A., Parvaz, M.A., Preston-Campbell, R.N., d'Oleire Uquillas, F., Moeller, S.J., Tomasi, D., Goldstein, R.Z., Alia-Klein, N., Habenula-prefrontal resting-state connectivity in reactive aggressive men – A pilot study, *Neuropharmacology* (2018), doi: <https://doi.org/10.1016/j.neuropharm.2018.10.025>.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Habenula-prefrontal resting-state connectivity in reactive aggressive men – a pilot study

Gan G^{1, 2}, Zilverstand A¹, Parvaz MA¹, Preston-Campbell RN¹, d'Oleire Uquillas F¹, Moeller SJ^{1,3}, Tomasi D⁴, Goldstein RZ¹, Alia-Klein N¹

¹Psychiatry and Neuroscience, Icahn School of Medicine at Mount Sinai, NY, USA;

²Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim/Heidelberg University, Mannheim, Germany

³Psychiatry, Stony Brook University School of Medicine, Stony Brook, NY, USA

⁴National Institute on Alcohol Abuse and Alcoholism, Bethesda, MD, USA

Corresponding author:

Dr. Gabriela Gan

Address: gabriela.gan@zi-mannheim.de

Keywords: reactive aggression, physical aggression, intermittent explosive disorder, resting-state fMRI, habenula, graph theory

Abstract

Disproportionate anger and reactive aggression in response to provocation are core symptoms of intermittent-explosive disorder (IED). Previous research shows a link between the propensity for aggression in healthy individuals and altered functioning of prefrontal-limbic and default-mode networks (DMN) at rest when no provocation is present. In a pilot study, we used resting-state functional magnetic resonance imaging to investigate the effects of pronounced reactive aggression in men, exemplified by IED, on the functional organization of resting-state brain networks including subcortical nodes such as the habenula previously implicated in aggression in preclinical models. Graph theory was applied to resting-state networks to determine alterations in global efficiency and clustering in high reactive aggressive men compared to low reactive aggressive men (controls). Further, we computed within-group correlations between trait aggression and graph measures, as well as within-group whole-brain seed-to-voxel regression analyses between trait aggression and habenula resting-state functional connectivity (rsFC). Reactive aggressive men compared to controls showed higher global efficiency in the left habenula, the left pulvinar in the thalamus, the left dorso-lateral prefrontal cortex, and the right temporal pole, as well as a trend for decreased clustering in DMN nodes. In the reactive aggressive group, high levels of trait aggression were linked to lower global efficiency of the left habenula, and to lower rsFC between the left habenula and the left ventro-lateral prefrontal cortex, a core region involved in inhibitory control. Together with preclinical evidence, our findings in men underline the relevance of aberrant habenula-prefrontal connectivity for the severity of aggressive behavior.

1. Introduction

Reactive aggression is an anger-based approach-related motivated behavior to a provoking or threatening event (Blair, 2012; Harmon-Jones et al., 2013). Repeated reactive aggressive outbursts are the core symptom of intermittent explosive disorder (IED), which affects about 3-7% of the US population (lifetime prevalence, Kessler et al., 2006; Scott et al., 2016). Despite the relatively high life time prevalence and devastating consequences of IED, to date the neurobiological underpinnings of IED are understudied compared to other, less prevalent mental health disorders also defined by aggressive tendencies including anti-social personality disorder (lifetime prevalence ~ 1% in the US population, Lenzenweger et al., 2007).

Neurobiological models derived from human studies suggest that disproportionate reactive aggression is mediated by a functional imbalance between hypo-active “top-down” prefrontal control and hyper-active “bottom-up” limbic areas implicating dysfunctional prefrontal-limbic circuits (Coccaro et al., 2007; Davidson et al., 2000; Siever, 2008). Specifically, individuals diagnosed with IED show greater right (McCloskey et al., 2016) and left amygdala reactivity (Coccaro et al., 2007), and lower activation in the right lateral orbitofrontal cortex in response to angry faces, as well as a reduced coupling between the left amygdala and the right ventro-medial prefrontal cortex (PFC) compared to controls (Coccaro et al., 2007). Furthermore, IED has been associated with lower grey matter volume of the right amygdala, and bilaterally in the ventro-medial PFC, the dorso-medial PFC, and the anterior insula (Coccaro et al., 2016), as well as lower white matter integrity in the left hemispheric superior longitudinal fasciculus that connects temporo-parietal association areas with the lateral and medial PFC (Lee et al.,

2016). Animal studies have underlined the role of subcortical limbic areas such as the medial hypothalamus (Falkner et al., 2014; Haller, 2013; Siegel and Victoroff, 2009) and the periaqueductal grey for defensive reactive aggressive behavior (Blair, 2012). Recently, the lateral habenula, a small epithalamic structure that plays a role in reward- and aversion-related motivated behavior (Boulos et al., 2017; Lawson et al., 2014; Matsumoto and Hikosaka, 2007), has been shown to regulate appetitive aggression in mice (Golden et al., 2016). Using a conditioned place preference paradigm with an intruder-paired “aggression” context, aggressive mice showed an increased preference for aggressive encounters while exhibiting decreased lateral habenula firing. Importantly, optogenetic photostimulation revealed that increasing lateral habenula firing (by silencing GABAergic basal forebrain projections to the lateral habenula) inhibited the aggression preference of aggressive mice. Thus, increased lateral habenula firing might represent an inhibitory mechanism to reduce the motivation to seek aggressive encounters in animals (Flanigan et al., 2017; Golden et al., 2016).

Resting-state fMRI (rsfMRI) allows for the study of the brain’s resting-state functional connectivity (rsFC) while individuals are alert but do not perform a task, thereby circumventing complex task-related variability in brain activation (Fox and Greicius, 2010). Previous research demonstrated altered rsFC in several of the aforementioned brain regions and circuits, but also in additional networks, in healthy non-clinical populations with a propensity for anger and aggression. Specifically, in healthy men, elevated trait anger and low anger control have been associated with reduced rsFC between the right amygdala and the left lateral orbitofrontal cortex (Fulwiler et al., 2012), the latter involved in behavior regulation and cognitive control (e.g., Miller and Cohen, 2001; Ochsner et al., 2004). In contrast, higher global rsFC between the

right amygdala and the right ventro-lateral PFC (vlPFC) has been reported after experimental anger induction compared to a “no-anger” resting baseline in healthy men with high trait anger (Gilam et al., 2017). Interestingly, higher rsFC between the right amygdala and the left vlPFC has also been associated with greater success in down-regulating emotions induced by angry faces in healthy participants (Morawetz et al., 2016). Moreover, the default-mode network (DMN), which is active during mind-wandering (Greicius et al., 2003; Raichle et al., 2001), has also been implicated in anger and aggression phenotypes. For example, healthy carriers of the low-activity monoamine-oxidase A (MAOA-L) genotype, a genetic risk marker for reactive aggression (Brunner et al., 1993; Buckholtz and Meyer-Lindenberg, 2008), showed increased rsFC in the DMN, but decreased rsFC in the dorsolateral prefrontal executive network (Clemens et al., 2015), a network involved in the preparation and execution of behavioral responses (Beckmann et al., 2005; Fox et al., 2006; Laird et al., 2011). Pronounced impulsivity in juvenile offenders was similarly linked to increased rsFC between the premotor cortex and the DMN, but diminished premotor rsFC with the dorsal prefrontal executive network (Shannon et al., 2011). Together, these findings suggest enhanced DMN rsFC, but reduced rsFC in medial as well as lateral prefrontal-limbic and premotor networks at an unprovoked baseline, in individuals at risk for aggressive behavior. However, to date, no systematic whole-brain investigations of rsFC in individuals with *clinically significant* reactive aggression, as exemplified by IED, have been undertaken. To that end, using a systematic whole-brain approach, here we aimed to investigate putative abnormalities in the functional organization of resting state brain networks in reactive aggressive men.

An emerging approach for investigating the organization and efficiency of brain connectivity in complex psychiatric disorders such as IED is graph theory, a mathematical framework in which a (brain) network is divided into graphs consisting of nodes (e.g., brain areas) and edges (e.g., connectivity between these areas) (Achard et al., 2006; Latora and Marchiori, 2001; Rubinov and Sporns, 2010). Healthy and efficient brain networks adhere to the “small-world” principle, in which the network contains modules (i.e., clusters of neighboring nodes) with high within-module connectivity, and robust links between these modules or clusters. Small-world networks are thus characterized by high local clustering, which indicates a high density of connections between neighboring nodes (Bassett and Bullmore, 2009), and high global efficiency, which indicates the efficient transfer of information on a global scale across the whole network through short path lengths between clusters (Achard et al., 2006; Watts and Strogatz, 1998). More specifically, the clustering coefficient metric is considered a measure of functional segregation indicating “the ability for specialized processing to occur within densely interconnected groups of neighboring brain regions” (Rubinov and Sporns, 2010). In contrast, global efficiency is considered a measure of functional integration indicating “the ease with which brain regions communicate” within the network (Rubinov and Sporns, 2010). Compromised clustering and global efficiency have been reported in resting-state brain networks across neuropsychiatric disorders including schizophrenia, Alzheimer’s disease, and attention deficit hyperactivity disorder (Bassett and Bullmore, 2009; Crossley et al., 2014; Wang et al., 2009).

By comparing high reactive aggressive men endorsing clinically significant behavioral features of IED to low aggressive healthy men, we investigated the effects of reactive aggression

on the functional organization of resting-state brain networks. We first used a graph theory approach to determine alterations in the global efficiency and the clustering coefficient across the whole brain, two central measures indicating the small-world organization of the brain. To further understand the relevance of potential alterations in the resting-state brain organization for the severity of aggressive behavior in the reactive aggressive group, we investigated associations between trait aggression (Buss and Perry, 1992) and the graph measures in nodes, that differed between the groups. Moreover, we also conducted whole brain seed-to-voxel regression analyses with those nodes differing between the groups and showing associations with trait aggression in the reactive aggressive group to determine whether rsFC of these nodes with other brain regions was modulated by trait aggression. Based on the human resting-state neuroimaging studies reviewed above, we expected that, compared to the low aggressive controls, reactive aggressive men would show altered global efficiency and clustering in the DMN, in prefrontal-limbic, and executive networks; based on the animal literature reviewed, we also expected alterations of these network properties in subcortical regions encompassing the hypothalamus, the periaqueductal grey, and the habenula.

2. Methods

2.1 Participants

Healthy low aggressive men and men with “anger issues” were recruited from the general population through newspaper advertisements and word-of-mouth (for details, Gan et al., 2016; Huang et al., 2018). We only included male participants to decrease variability in this pilot study and given a moderate overrepresentation of IED in males (Kessler et al., 2006; Scott et al.,

2016). All individuals provided written informed consent prior to study participation in accordance with Stony Brook University's Institutional Review Board.

All participants underwent a psychiatric Structured Clinical Interview (SCID, Ventura et al., 1998) to assess DSM-IV Axis-I psychiatric, and Axis-II Cluster-B personality disorders including anti-social personality disorder and borderline personality disorder (First, 1997), as well as a structured interview to assess DSM-IV and integrated research criteria for IED (IED-IR) (Coccaro, 2012; Coccaro et al., 2004). The IED-IR interview was only completed in participants who endorsed the gate question concerning verbal or physically aggressive outbursts (e.g., intense verbal arguments, temper tantrums, damage of property) within the past year. If no aggressive outbursts were endorsed, the IED-IR interview was not undertaken and IED-IR scores were not computed. In total, 11 males endorsed significant behavioral features consistent with full IED ($n=7$, IED-IR interview score ≥ 15) or subclinical IED ($n=4$; operationalized in the current study as an IED-IR interview score = 13-14), and these constituted our reactive aggressive group. Twelve healthy male controls did not endorse any aggressive outbursts in the gate question for the IED-IR interview; they also did not meet criteria for any other DSM-IV psychiatric disorder. They were matched with the reactive aggressive group on age, ethnicity, and years of education, and constituted our low aggressive group. Following our previous work (Gan et al., 2016), we included reactive aggressive participants with subclinical IED to the reactive aggressive group to acknowledge the continuous nature of aggressive behavior; this decision was also in line with the "Research Domain Criteria" (RDoC) for relevant functional domains inclusive of the negative valence system (Morris and Cuthbert, 2012). Additionally, all participants completed the Buss-Perry aggression questionnaire (BPQ) assessing overall trait

aggression, as well as physical and verbal aggression, anger, and hostility (Buss and Perry, 1992). We also assessed depressive symptoms with the Beck Depression Inventory (BDI-II) (Beck and Steer, 1984), and substance use (e.g., alcohol, cigarette, cannabis) with the SCID. Five reactive aggressive individuals reported comorbid disorders frequently found for IED (Kessler et al., 2006): current generalized anxiety disorder and a single/remitted major depressive episode (n=1); current anti-social personality disorder (n=2); remitted lifetime alcohol abuse (n=1); and remitted lifetime cannabis abuse (n=1) (see **Table 1** for substance use patterns and depressive symptoms).

Exclusion criteria were (see also, Gan et al., 2016) any: a) neurological disease, including seizures, and/or history of head trauma with loss of consciousness (>30 minutes); b) major physiological conditions (e.g., cardiovascular, endocrinological, oncological, or autoimmune disease); c) major psychiatric disorder with psychosis (e.g., schizophrenia, bipolar disorder), as well as current substance use disorders; d) use of psychoactive medication within 6 months prior to study date; e) MRI contraindications; and f) positive urine screens for psychoactive drugs or their metabolites (amphetamine/methamphetamine, cocaine, phencyclidine, benzodiazepines, cannabis, opiates, barbiturates, inhalants).

2.2 MRI data acquisition

Functional and structural MRI was performed on a 4T whole-body Varian/Siemens MRI scanner at the Brookhaven National Laboratory, Upton, New York. For rsfMRI, the blood oxygenation level dependent (BOLD)-fMRI contrast was measured with a T2*-weighted single-shot gradient-echo echo planar imaging sequence (EPI): echo time (TE)=20 ms; repetition time (TR)=1600 ms; 3.125×3.125 mm² in-plane resolution; 4 mm slice thickness; 1 mm gap; 33 coronal

slices; field of view (FOV)=200 mm; 64×64 matrix size; 90°-flip angle; 200 kHz bandwidth with ramp sampling, 4 dummy scans. A T1-weighted structural MRI scan was acquired for each participant using a 3D-MDEFT (3 dimensional modified driven-equilibrium Fourier transform) sequence (Lee et al., 1995): TE=7ms, TR=15 ms, 0.94×0.94×1.00 mm³ spatial resolution, axial orientation, 256 readout and 192×96 phase-encoding steps. All participants completed a 5-minute rsfMRI scan (195 volumes), and a 16-minute structural MRI scan. For rsfMRI, participants were instructed to lie as still as possible with eyes open. Padding, earplugs, and headphones were used to minimize motion and scanner noise (Tomasi et al., 2005).

2.3 Data analysis

2.3.1 Trait measures of anger and aggression

We computed independent t-tests or χ^2 -tests to test for group differences in demographics, depressive symptoms, substance use, the BPQ total trait aggression score and all four subscales (anger, verbal and physical aggression, hostility; see **Table 1**). The alpha-level was Bonferroni-corrected for the number of tests performed (n=15 tests; $p<.004$).

2.3.2 rsfMRI data analysis

Resting-state fMRI data were preprocessed and analyzed using MATLAB-based toolboxes (The Mathworks, Natick, MA, USA) including the CONN functional connectivity toolbox (version 16.b) (Whitfield-Gabrieli and Nieto-Castanon, 2012) (<http://www.nitrc.org/projects/conn/>), and statistical parametric mapping (SPM12, Wellcome Trust Centre for Neuroimaging, London, UK).

For *preprocessing*, fMRI BOLD data were slice-time corrected, realigned, normalized to the standard MNI-EPI template, and smoothed with an 8-mm isotropic FWHM Gaussian kernel, a smoothing kernel associated with robust performance in SPM analyses (Eklund et al., 2015; Hayasaka and Nichols, 2003). The resampled voxel size was 2x2x2 mm. We identified motion (>2 mm translation, >1.15 degree of rotation) and signal (spiking, >3 standard deviations of the signal mean, z-threshold) outlier time points for each participant using the CONN-supported Artifact Detection Toolbox (ART) (https://www.nitrc.org/projects/artifact_detect/) (Whitfield-Gabrieli and Nieto-Castanon, 2012). For each participant, we added to the CONN denoising step an array containing one binary regressor for each outlier time point, the six SPM rigid-body realignment parameters and their first derivatives (3 translation, 3 rotation parameters), and a composite measure of scan-to-scan movement computed by the ART toolbox, which estimates the maximum voxel displacement resulting from the combined effect of the individual translation and rotation displacement measures. Further, we modeled spurious low frequency fluctuations of the BOLD signal due to non-neuronal sources such as scanner artifacts in the white matter (WM) (Jo et al., 2010), and physiological artifacts (e.g., cardiac, respiratory, motion) in the cerebro-spinal fluid (CSF) (Birn et al., 2009) by extracting BOLD signal components from CSF and WM areas for each participant. These noise confounds were then added to the denoising step using the component-based noise correction method (CompCor, Behzadi et al., 2007) as implemented in CONN (Whitfield-Gabrieli and Nieto-Castanon, 2012). Further, grey matter masks were generated for each participant to ensure that only BOLD signal from the grey matter was considered in the analysis. Importantly, the CompCor method does not rely on global signal regression, another commonly used technique in rsfMRI for regressing out non-

neuronal confounds, but which has been previously criticized for inducing spurious anti-correlations due to alterations in the distribution of correlation coefficients (Chang and Glover, 2009; Murphy and Fox, 2017). All confounding factors were regressed out from the BOLD time series at each voxel using linear regression. Linear detrending was applied as part of the nuisance regression. Then, to isolate low-frequency resting-state BOLD signal fluctuations, the BOLD time series were band-pass filtered within a range of 0.008-0.09 Hz to reduce the effect of low-frequency drift and high-frequency noise (Hallquist et al., 2013).

There were no significant group differences in scan-to-scan movement (reactive aggressive group: mean = 0.19 mm \pm 0.08 mm; controls: mean = 0.27 mm \pm 0.12 mm, $t(21)=-1.695$, $p=.105$) or in the number of outlier time points (reactive aggressive: mean = 2.5% \pm 2.6%; control group: mean = 5.1% \pm 7.3%, $U=59.5$, $p=.683$, $n=23$; Mann-Whitney U-Test due to non-normal distribution).

For the *graph theory analyses*, the brain was parcellated into 638 similarly-sized brain regions (nodes) using a high-resolution anatomical template (Crossley et al., 2013), which respects grey matter landmarks based on a binarized mask of the “Automated Anatomical Labeling” template (Tzourio-Mazoyer et al., 2002; for details about the parcellation algorithm, see Zalesky et al., 2010). The Crossley template has previously been shown to be sensitive for detecting impairment of hubs across mental disorders (e.g., schizophrenia, and affective disorders) (Crossley et al., 2014). We added 5 subcortical regions of interest that were not covered by the parcellation including the periaqueductal grey [5mm sphere at $x=1$, $y=-29$, $z=12$, based on (Linnman et al., 2012)], the left/right hypothalamus [5mm sphere at $x=\pm 8$, $y=-4$, $z=-4$, based on

(Kroemer et al., 2013)], and the left/right habenula [3 mm sphere at $x=-2.8/4.8$, $y=-24.4/-24.1$, $z=2.3/2.2$; based on (Lawson et al., 2013); all coordinates given in MNI space] resulting in a total of 643 nodes.

For each participant, a 643x643 matrix of Fisher-transformed bivariate correlation coefficients between the BOLD time series of each node with all other nodes was computed. For computation of graph theory measures, the 643x643 matrix was thresholded at a fixed cost level of 0.15 including the 15% strongest positive connections. Based on the individual graph, global efficiency and a clustering coefficient metric were computed for each participant. We did not focus on local efficiency and characteristic path length as they are highly associated with the clustering coefficient or the global efficiency, respectively (Bullmore and Sporns, 2009; Rubinov and Sporns, 2010; Strang et al., 2018). We also did not focus on centrality measures such as the degree of a node (i.e., the number of edges connected to a given node), or the betweenness centrality (i.e., “the fraction of all shortest paths in the network that pass through a given node”) (Rubinov and Sporns, 2010) as the global efficiency of a single node (i.e., the average inverse shortest path distance from a single node to all other nodes) can be considered a measure of the centrality of a node (Whitfield-Gabrieli and Nieto-Castanon, 2012). Nevertheless, for thoroughness, we present group differences for the above mentioned graph measures in the supplementary material.

Following recommendations for graph theory analyses (Achard et al., 2006; Whitfield-Gabrieli and Nieto-Castanon, 2012), we show that small-world properties, indicated by higher global efficiency of our network than that of a lattice graph, as well as higher clustering/local efficiency

than that of a random graph, were present at our cost threshold of 0.15 across all participants, as well as separately for the reactive aggressive and the control group (supplementary Figure S1). Consistent with previous studies (Redcay et al., 2013; Rubinov and Sporns, 2010), we confirmed that our findings were valid at a variety of other cost thresholds from 0.05, 0.1, 0.2, to 0.25.

To identify differences in rsFC between the reactive aggressive and the control group, group differences in global efficiency and clustering coefficients were assessed across the whole network using independent *t*-tests in CONN. While we focused on group differences at a corrected whole-brain node-wise significance level of $pFDR < .05$ to minimize the risk of false-positives, we additionally report all nodes showing group differences at an uncorrected whole-brain node-wise threshold of $p < .001$ to further explore the extent of the anticipated group differences. Network-level summary measures for global efficiency and clustering coefficient averaged across all nodes of the network are also reported (Whitfield-Gabrieli and Nieto-Castanon, 2012).

To further assess the relevance of group differences in global efficiency and clustering coefficient for the severity of aggressive behavior within the reactive aggressive group, we computed within-group Spearman correlation coefficients between the graph measures of the nodes showing group differences at a corrected significance level and the BPQ total aggression score using SPSS (IBM, NY, USA). This analysis followed guidelines by Fox and Greicius (2010), who proposed to validate rsFC differences between a clinical and a healthy control group by studying associations with a variable assessing the severity of a relevant clinical symptom (such as aggressive behavior). To confirm that significant correlations were specific for the reactive

aggressive group, we compared whether correlation coefficients were significantly different between the reactive aggressive and the control group using Fisher's *r*-to-*z* transformation (R-based online tool, Diedenhofen and Musch, 2015; Fisher, 1925). Due to the dependency between the BPQ total score and the BPQ subscales (verbal and physical aggression, anger, hostility; Spearman's $r > .81$) (see also, Buss and Perry, 1992), we present exploratory correlation analyses between the graph measures and the four BPQ subscales in the supplementary material only.

Finally, we used those nodes showing significant group differences and correlations with the severity of trait aggression (BPQ total score) as seeds in within-group *post-hoc whole-brain seed-to-voxel regression analyses* in CONN; these analyses were conducted in the reactive aggressive group to determine whether these select nodes showed altered rsFC with other brain regions as a function of trait aggression. To confirm that specific correlations were not present in the control group, we (i) compared rsFC for clusters showing effects with trait aggression in the reactive aggressive group by extracting the rsFC values, (ii) assessed the same whole-brain regression model in the controls only, and (iii) formally compared whole-brain regression models between both groups. The significance level for the seed-to-voxel analyses was set to an FDR-corrected cluster-threshold of $p < .05$ at an initial voxel-wise height threshold of $p < .001$ uncorrected.

3. Results

3.1 Trait anger and aggression

Table 1 displays demographics, substance use, depressive symptoms, and trait anger, hostility and aggression measured with the BPQ for the two groups. The groups did not differ on age,

race or education. There were trends for group differences in cigarette and alcohol use, such that more reactive aggressive individuals ($n=4$) were current cigarette smokers than controls ($n=1$), and more controls ($n=6$) reported occasional alcohol use than reactive aggressive individuals ($n=1$). However, continuous measures of current and lifetime alcohol (days/years of alcohol use) did not differ between the groups. There was also no difference between the groups in cannabis use. Relative to controls, the reactive aggressive group reported significantly higher depressive symptoms on the BDI-II, as well as significantly higher aggression on the BPQ total aggression score and all subscales.

3.2 Group differences in rsFC

Graph theory analyses revealed significantly higher global efficiency and lower clustering coefficient at the network level (across all nodes, entire network) for reactive aggressive individuals compared to controls (**Table 2**). Specifically, the global efficiency of individual nodes was higher in the reactive aggressive relative to the control group in the left pulvinar (located in the thalamus), the left habenula, the left dorso-lateral PFC (dlPFC), and the right temporal pole at a corrected significance threshold, as well as in the right superior temporal gyrus at the exploratory uncorrected $p<.001$ threshold (**Figure 1A, Table 2**). The clustering coefficient of individual nodes was decreased in the left precuneus, the left dorso-medial PFC, and in left occipital areas at the exploratory uncorrected $p<.001$ threshold only (**Figure 1B, Table 2**). Additional graph theory measures (i.e., local efficiency, characteristic path length, degree, and betweenness centrality) yielded redundant results except for betweenness centrality showing group differences in a single occipital node (**Table S1**, supplementary material).

3.3 Within-group correlation analyses between node-wise global efficiency and trait aggression

Correlation analyses between the global efficiency of the left habenula, the left pulvinar, the left dlPFC, and the right temporal pole, and trait aggression (BPQ total score) revealed a significant negative correlation between trait aggression and global efficiency of the left habenula in the reactive aggressive group (Table 3, Figure 2B; significance level Bonferroni-corrected for $n=4$ tests in the reactive aggressive group, $p<.0125$). This correlation coefficient was significantly different from the correlation coefficient in the controls (Fisher's $Z=-2.389$, $p=.017$; Table 3, Figure 2B). Thus, while higher levels of trait aggression in the reactive aggressive group were associated with lower global efficiency of the left habenula, no significant association emerged in controls, and this group difference in correlation coefficients was significant. Exploratory correlation analyses between the global efficiency of the four nodes and the four BPQ subscales revealed significant negative correlations in the reactive aggressive group between the left habenula and BPQ physical aggression (Spearman's $r=-0.726$, $p=.011$; not significantly different from the controls, Fisher's $Z=-1.631$, $p=.102$), as well as between the left pulvinar and BPQ anger (Spearman's $r=-0.764$, $p=.006$; significantly different from the controls, Fisher's $Z=-2.540$, $p=.011$; Table S2) at an uncorrected significance threshold only. Across all correlation analyses between global efficiency with trait aggression, and the BPQ subscales, no correlation coefficient reached significance in the controls.

3.4 Within-group post-hoc whole-brain seed-to-voxel regression analyses with the habenula as the seed

Post-hoc whole-brain seed-to-voxel regression analyses revealed a negative correlation in the reactive aggressive group between trait aggression (BPQ total score) and rsFC of the left habenula (i.e.,

the seed) with a left-lateralized cluster encompassing the vIPFC (BA 47) and the temporal pole (BA 38) (**Table 4, Figure 2C**). While higher trait aggression was linked to a more negative habenula-vIPFC rsFC, lower trait aggression was linked to a more positive habenula-vIPFC rsFC (**Figure 2D**). In contrast, habenula-vIPFC rsFC (i.e., values extracted for the controls from the vIPFC cluster shown in **Figure 2C**) did not correlate with trait aggression in the controls (*Spearman's* $r=.224$, $p=.484$); likewise, no other correlation with the left habenula seed emerged in any other cluster for the controls in a within-group whole-brain seed-to-voxel regression analysis. Comparing regression models between the reactive aggressive and the control group in CONN revealed a significant group difference in the correlation between trait aggression and rsFC of the left habenula with occipital areas (cluster-wise $pFDR<.05$), as well as with the left vIPFC and the cerebellum at an uncorrected voxel-wise threshold ($p<.001$ uncorrected, >50 connected voxels; **Table S3**).

Due to significant group differences we further assessed the influence of depressive symptoms on our brain imaging measures (i.e., global efficiency, clustering coefficient, and the seed-based habenula-vIPFC rsFC) by computing bivariate correlations with the BDI scores. Depressive symptoms were significantly correlated with the clustering coefficient in the calcarine gyrus node and the dorso-medial PFC node in the controls, but not in the reactive aggressive group (**Table S4**). As there were no associations between depressive symptoms and global efficiency in any node that was used for follow-up correlation analyses and for the habenula-vIPFC rsFC, we did not include depressive symptoms as a covariate in the whole-brain graph theory analysis or the post-hoc whole-brain seed-to-voxel regression analysis to not further lower the power of the analysis.

4. Discussion

Using a resting-state fMRI graph theory approach, we identified larger global efficiency, specifically in the left habenula, the left pulvinar (located in the thalamus), the left dlPFC, and the right temporal pole, as well as a trend for lower clustering coefficient, specifically in DMN nodes (including the left dorso-medial PFC and left precuneus) in reactive aggressive men, showing significant behavioral features of IED, compared to low aggressive men. Moreover, correlation analyses demonstrated that the global efficiency of the left habenula correlated with the severity of aggression in the reactive aggressive group only. Follow-up post-hoc whole-brain seed-to-voxel analyses with the left habenula as the seed further determined that specifically rsFC between the left habenula and the left vlPFC was correlated to trait aggression in the reactive aggressive group. Thus, as compared to low aggressive men (controls), reactive aggressive men showed overall higher habenula global efficiency. However, within the reactive aggressive group, lower levels of trait aggression were linked to higher habenula global efficiency, as well as to a more positive habenula-vlPFC rsFC. In contrast, higher levels of trait aggression were linked to lower habenula global efficiency, as well as to a more negative habenula-vlPFC rsFC again in this group. Because the vlPFC is considered a core region for inhibitory control (Corbetta and Shulman, 2002; Swick et al., 2011), our findings in the reactive aggressive men support preclinical findings highlighting the potential relevance of habenula-cortical connectivity for the inhibitory regulation of aggressive behavior (Golden et al., 2016). While caution should be taken in interpreting our findings due to the pilot nature of our study, altered habenula-vlPFC rsFC could be a marker for the risk of engaging in premature aggressive responses during anger episodes in severely reactive aggressive men.

4.1 Heightened global efficiency at rest in reactive aggressive men

Heightened global efficiency of the left habenula, the left pulvinar, the left dlPFC, and the right temporal pole in reactive aggressive individuals partly converges with previous results implicating subcortical and prefrontal nodes in aggression across species (Alia-Klein et al., 2014; Clemens et al., 2015; Golden et al., 2016). Importantly, increased habenula firing induced by silencing GABAergic projections from the basal forebrain to the lateral habenula has been shown to inhibit a preference for seeking aggressive encounters and to decrease the severity of aggressive bouts (i.e., attack duration) in aggressive mice (Golden et al., 2016). In contrast, we observed that reactive aggressive relative to low aggressive men exhibited higher global efficiency of the left habenula at rest while no provocation and likely no aggression motivation was present. Thus, habenula functioning in reactive aggressive men might differ between rest and active aggressive behavior, a framework that merits further characterization in humans using fMRI paradigms previously shown to robustly elicit reactive aggression (Gan et al., 2016; Gan et al., 2015; Kose et al., 2015; Kramer et al., 2007). Nevertheless, given the negative correlation between habenula global efficiency with the severity of aggression, our findings also suggest that larger habenula global efficiency at rest is associated with lower trait aggression in the highly reactive aggressive men. This finding indicates that habenula activity or its integration with the rest of the brain might be needed for a lower severity of aggression, as previously shown in animal studies (Golden et al., 2016).

Heightened global efficiency of the left pulvinar in the reactive aggressive group is consistent with previous findings linking heightened reactivity of this region to high trait anger (Alia-Klein et al., 2018), and to anger reactivity in carriers of the MAOA-L genotype (Alia-Klein et al.,

2009). Moreover, poor anger regulation in response to unfair offers during an “anger-infused” ultimatum game has been associated with lower connectivity between the dorsal posterior insula and the medial thalamus, which is located close to the pulvinar (Gilam et al., 2015). While the role of the pulvinar and other thalamic nuclei for anger and aggression remains to be elucidated, it is possible that in human reactive aggression the pulvinar might be linked to altered bottom-up visual attention allocation and arousal in response to behaviorally-salient stimuli (Anders et al., 2004; Fischer and Whitney, 2012; Kanai et al., 2015; Portas et al., 1998). In line with this idea, the thalamus has recently been discussed to be involved in the responding to salient drug cues in drug addiction (Huang et al., 2018). Furthermore, enhanced global efficiency of the left dlPFC at rest in reactive aggressive men might be linked to enhanced functioning of an intrinsic connectivity network that is involved in the preparation of executed movements (Beckmann et al., 2005; Fox et al., 2006; Laird et al., 2011). This finding is partly in line with heightened left dlPFC activity during error processing in individuals diagnosed with IED, suggesting enhanced sensitivity for salient erroneous events (Moeller et al., 2014). In contrast, individuals with a propensity for aggression (MAOA-L carriers), and impulsivity (juvenile offenders), showed decreased rsFC of the dorsal prefrontal executive network (Clemens et al., 2015; Shannon et al., 2011). Differences in enhanced compared to impaired functioning of the dlPFC across studies might be explained by the heterogeneity of the aggression phenotypes that have been studied (IED vs. MAOA-L carriers, impulsive juvenile offenders), as well as by the heterogeneity of the fMRI methods employed (task-based vs. rsfMRI, graph theory vs. ICA). Finally, although the right temporal pole has not been directly associated with aggressive behavior, it belongs to an intrinsic connectivity network involved in

the discrimination of social-emotional stimuli (Laird et al., 2011), and the retrieval of social scripts (Adolphs, 2009; Frith, 2007; Olson et al., 2013), crucial functions for adaptive behavior. Taken together, as higher levels of global efficiency of nodes indicate a larger centrality within the network, it can be posited that higher global efficiency of the left habenula, the left pulvinar, the left dlPFC, and the right temporal pole may reflect a more efficient transfer of information regarding salient provoking events between subcortical and cortical nodes. This more efficient transfer could potentially be of detriment for behavior and emotion control in reactive aggressive men. In support of this interpretation, higher global efficiency has been suggested to reflect slightly more random and over-active functional brain networks in other neuropsychiatric diseases such as schizophrenia and Alzheimer's disease (Alexander-Bloch et al., 2010; Bassett and Bullmore, 2009; Lynall et al., 2010; Sanz-Arigita et al., 2010). On the other hand, and consistent with the correlation between habenula global efficiency and the severity of aggression, our findings may also suggest that higher levels of habenula global efficiency might represent a compensatory mechanism that can be deployed to decrease the severity of aggression in this population.

4.2 Altered habenula-prefrontal rsFC in reactive aggressive men

This idea is supported by the habenula-vlPFC rsFC findings in the reactive aggressive group, with lower aggression severity linked to positive habenula-vlPFC rsFC. Given limbic-prefrontal pathways in animal models (Boulos et al., 2017; Matsumoto and Hikosaka, 2007) and human tractography studies showing anatomical connectivity between the habenula and prefrontal regions (Shelton et al., 2012), the habenula-vlPFC connectivity pattern is plausible. The bilateral vlPFC has further been implicated in the control of responses towards unexpected behaviorally-

salient stimuli such as in motor inhibition tasks (Corbetta and Shulman, 2002; Duann et al., 2009; Sharp et al., 2010; Swick et al., 2011). Importantly, activation and connectivity of the vLPFC (with the amygdala) has been linked to anger in resting-state and task-based fMRI studies, suggesting a role in enhanced attention allocation to unpleasant emotional stimuli (Alia-Klein et al., 2018), anger regulation (Morawetz et al., 2016), and anger rumination (Gilam et al., 2017). We therefore speculate that in the more severely reactive aggressive men, a deficient communication between the left habenula and the left vLPFC could contribute to a diminished capacity to exert inhibitory control over aggressive responses. In contrast, enhanced (or more positive) habenula-vLPFC rsFC in the less severely reactive aggressive men could represent a protective mechanism, as indeed suggested by the negative correlation with trait aggression in the reactive aggressive group. Note that the lack of effects in controls might be explained by the overall low range in BPQ aggression scores, which would also not require a high amount of self-control. Together with findings from animal models suggesting an involvement of basal-forebrain-habenula projections specifically for the severity of aggressive behavior (Golden et al., 2016), our findings stress the role of the habenula, and its connections to the vLPFC, in the severity of aggression in human reactive aggressive men.

In the current rsFC study, we did not observe associations with the ventro-medial PFC, a brain area shown to play a role in anger and aggression across fMRI, rsFC, and structural investigations (Alia-Klein et al., 2014; Beyer et al., 2015; Coccaro et al., 2016; Coccaro et al., 2007; Gilam et al., 2015; Jacob et al., 2018). Ventro-medial PFC involvement in anger and aggression has been suggested to be related to modulating negative affect (e.g., Davidson et al., 2000). A lack of the ventro-medial PFC involvement in our study might be attributed to the absence of a

provocation that might have triggered the need to ruminate or regulate negative affect. Further, the involvement of the habenula in aggressive behavior is a relatively recently developed concept in the literature, potentially specifically linked to the severity of aggression (see also, Golden et al., 2016). Hence, previous anger and aggression studies did not specifically assess this small epithalamic region that can be easily overlooked in whole-brain analyses.

Although only a trend in the current study, lower clustering in DMN nodes in reactive aggressive men is consistent with resting-state studies linking compromised DMN functioning to a propensity for anger and aggression in non-clinical populations (Alia-Klein et al., 2014; Clemens et al., 2015; Shannon et al., 2011). Thus, with the DMN's role in self-referential processing and self-awareness during mind wandering (Alia-Klein et al., 2014; Fox et al., 2005; Gusnard et al., 2001), lower clustering predominantly in the DMN in reactive aggressive compared to low aggressive men could indicate compromised self-referential processing at rest.

4.3 Limitations

The limited spatial resolution of our rsfMRI and structural data prevented us from employing a state-of-the-art semi-automatic approach to locate the small habenula nuclei (Ely et al., 2016; Kim et al., 2016). We were therefore unable to specifically focus on the lateral habenula as a hub for motivated aggressive behavior (Golden et al., 2016), or to compare the lateral with the medial habenula, which has been implicated in social conflict in zebrafish (Chou et al., 2016). Moreover, our findings should be interpreted with caution due to the small sample of reactive and low aggressive men, particularly in generalizing our findings to a broader IED population of men and women. Due to the small sample size, we did not further investigate a trend of more current tobacco use in the reactive aggressive group, and more occasional alcohol use in the

controls. Efforts with larger sample sizes including men and women will be important in replicating the role of the habenula in reactive aggression, in addition to considering associations with varying patterns of tobacco, alcohol, and cannabis use problems, as well as depressive symptomatology, frequent comorbidities in IED (Kessler et al., 2006). High-resolution studies with a 7-T MRI scanner, or higher resolution imaging are also needed to better characterize the function of the lateral and medial habenula in human reactive aggression.

4.4 Conclusion

Despite the pilot nature of our study, our graph theory findings of higher global efficiency in the left habenula, the left pulvinar, the left dlPFC, and the right temporal pole in reactive aggressive men compared to low aggressive men suggest a more efficient information transfer in reactive aggressive men between subcortical-prefrontal networks that are important for attention allocation, motivated aggressive behavior, and the execution and control of behavioral responses. These alterations at rest may underlie a more efficient transfer of information regarding provoking events, culminating in a readiness to aggress at the slightest of provocations. Combined with reduced self-control, this readiness could translate to actual aggression at a trait level. The association between higher trait aggression with negative habenula-vlPFC rsFC in the severely reactive aggressive men provide first evidence for an inhibitory control role of the habenula in stable patterns of reactive aggression, in line with findings from preclinical studies. Alternatively, and irrespective of the interpretation of the direction of this association, positive habenula-vlPFC rsFC in the less severely reactive aggressive individuals could also point to a compensatory process that contributes to the ability

to experience less trait aggression. Overall, a better understanding of habenula-vlPFC functioning and connectivity in human reactive aggression is crucial for exploring new therapeutic targets for the treatment of severe reactive aggression.

5. References

- Achard, S., Salvador, R., Whitcher, B., Suckling, J., Bullmore, E., 2006. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J Neurosci* 26, 63-72.
- Adolphs, R., 2009. The social brain: neural basis of social knowledge. *Annu Rev Psychol* 60, 693-716.
- Alexander-Bloch, A. F., Gogtay, N., Meunier, D., Birn, R., Clasen, L., Lalonde, F., Lenroot, R., Giedd, J., Bullmore, E. T., 2010. Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. *Front Syst Neurosci* 4, 147.
- Alia-Klein, N., Goldstein, R. Z., Tomasi, D., Woicik, P. A., Moeller, S. J., Williams, B., Craig, I. W., Telang, F., Biegon, A., Wang, G. J., Fowler, J. S., Volkow, N. D., 2009. Neural mechanisms of anger regulation as a function of genetic risk for violence. *Emotion* 9, 385-396.
- Alia-Klein, N., Preston-Campbell, R. N., Moeller, S. J., Parvaz, M. A., Bachi, K., Gan, G., Zilverstand, A., Konova, A. B., Goldstein, R. Z., 2018. Trait anger modulates neural activity in the fronto-parietal attention network. *PLoS One* 13, e0194444.
- Alia-Klein, N., Wang, G. J., Preston-Campbell, R. N., Moeller, S. J., Parvaz, M. A., Zhu, W., Jayne, M. C., Wong, C., Tomasi, D., Goldstein, R. Z., Fowler, J. S., Volkow, N. D., 2014. Reactions to media violence: it's in the brain of the beholder. *PLoS One* 9, e107260.
- Anders, S., Lotze, M., Erb, M., Grodd, W., Birbaumer, N., 2004. Brain activity underlying emotional valence and arousal: a response-related fMRI study. *Hum Brain Mapp* 23, 200-209.
- Bassett, D. S., Bullmore, E. T., 2009. Human brain networks in health and disease. *Curr Opin Neurol* 22, 340-347.
- Beck, A. T., Steer, R. A., 1984. Internal consistencies of the original and revised Beck Depression Inventory. *J Clin Psychol* 40, 1365-1367.

- Beckmann, C. F., DeLuca, M., Devlin, J. T., Smith, S. M., 2005. Investigations into resting-state connectivity using independent component analysis. *Philos Trans R Soc Lond B Biol Sci* 360, 1001-1013.
- Behzadi, Y., Restom, K., Liau, J., Liu, T. T., 2007. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* 37, 90-101.
- Beyer, F., Munte, T. F., Gottlich, M., Kramer, U. M., 2015. Orbitofrontal Cortex Reactivity to Angry Facial Expression in a Social Interaction Correlates with Aggressive Behavior. *Cereb Cortex* 25, 3057-3063.
- Birn, R. M., Murphy, K., Handwerker, D. A., Bandettini, P. A., 2009. fMRI in the presence of task-correlated breathing variations. *Neuroimage* 47, 1092-1104.
- Blair, R. J., 2012. Considering anger from a cognitive neuroscience perspective. *Wiley Interdiscip Rev Cogn Sci* 3, 65-74.
- Boulos, L. J., Darcq, E., Kieffer, B. L., 2017. Translating the Habenula-From Rodents to Humans. *Biol Psychiatry* 81, 296-305.
- Brunner, H. G., Nelen, M., Breakefield, X. O., Ropers, H. H., van Oost, B. A., 1993. Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. *Science* 262, 578-580.
- Buckholtz, J. W., Meyer-Lindenberg, A., 2008. MAOA and the neurogenetic architecture of human aggression. *Trends Neurosci* 31, 120-129.
- Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 10, 186-198.
- Buss, A. H., Perry, M., 1992. The aggression questionnaire. *J Pers Soc Psychol* 63, 452-459.
- Chang, C., Glover, G. H., 2009. Effects of model-based physiological noise correction on default mode network anti-correlations and correlations. *Neuroimage* 47, 1448-1459.
- Chou, M. Y., Amo, R., Kinoshita, M., Cherng, B. W., Shimazaki, H., Agetsuma, M., Shiraki, T., Aoki, T., Takahoko, M., Yamazaki, M., Higashijima, S., Okamoto, H., 2016. Social conflict resolution regulated by two dorsal habenular subregions in zebrafish. *Science* 352, 87-90.
- Clemens, B., Voss, B., Pawliczek, C., Mingoia, G., Weyer, D., Repple, J., Eggermann, T., Zerres, K., Reetz, K., Habel, U., 2015. Effect of MAOA Genotype on Resting-State Networks in Healthy Participants. *Cereb Cortex* 25, 1771-1781.
- Coccaro, E. F., 2012. Intermittent explosive disorder as a disorder of impulsive aggression for DSM-5. *Am J Psychiatry* 169, 577-588.

- Coccaro, E. F., Fitzgerald, D. A., Lee, R., McCloskey, M., Phan, K. L., 2016. Frontolimbic Morphometric Abnormalities in Intermittent Explosive Disorder and Aggression. *Biol Psychiatry Cogn Neurosci Neuroimaging* 1, 32-38.
- Coccaro, E. F., McCloskey, M. S., Fitzgerald, D. A., Phan, K. L., 2007. Amygdala and orbitofrontal reactivity to social threat in individuals with impulsive aggression. *Biol Psychiatry* 62, 168-178.
- Coccaro, E. F., Schmidt, C. A., Samuels, J. F., Nestadt, G., 2004. Lifetime and 1-month prevalence rates of intermittent explosive disorder in a community sample. *J Clin Psychiatry* 65, 820-824.
- Corbetta, M., Shulman, G. L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3, 201-215.
- Crossley, N. A., Mechelli, A., Scott, J., Carletti, F., Fox, P. T., McGuire, P., Bullmore, E. T., 2014. The hubs of the human connectome are generally implicated in the anatomy of brain disorders. *Brain* 137, 2382-2395.
- Crossley, N. A., Mechelli, A., Vertes, P. E., Winton-Brown, T. T., Patel, A. X., Ginestet, C. E., McGuire, P., Bullmore, E. T., 2013. Cognitive relevance of the community structure of the human brain functional coactivation network. *Proc Natl Acad Sci U S A* 110, 11583-11588.
- Davidson, R. J., Putnam, K. M., Larson, C. L., 2000. Dysfunction in the neural circuitry of emotion regulation--a possible prelude to violence. *Science* 289, 591-594.
- Diedenhofen, B., Musch, J., 2015. cocor: a comprehensive solution for the statistical comparison of correlations. *PLoS One* 10, e0121945.
- Duann, J. R., Ide, J. S., Luo, X., Li, C. S., 2009. Functional connectivity delineates distinct roles of the inferior frontal cortex and presupplementary motor area in stop signal inhibition. *J Neurosci* 29, 10171-10179.
- Eklund, A., Nichols, T., Knutson, H., 2015. Can parametric statistical methods be trusted for fMRI based group studies? ArXiv e-prints <http://adsabs.harvard.edu/abs/2015arXiv151101863E>.
- Ely, B. A., Xu, J., Goodman, W. K., Lapidus, K. A., Gabbay, V., Stern, E. R., 2016. Resting-state functional connectivity of the human habenula in healthy individuals: Associations with subclinical depression. *Hum Brain Mapp* 37, 2369-2384.
- Falkner, A. L., Dollar, P., Perona, P., Anderson, D. J., Lin, D., 2014. Decoding ventromedial hypothalamic neural activity during male mouse aggression. *J Neurosci* 34, 5971-5984.
- First, M., Gibbon M, Spitzer RL, Williams JBW, Benjamin LS, 1997. Structured Clinical Interview for DSM-IV Axis II Personality Disorders, (SCID-II). American Psychiatric Press, Inc., Washington, D.C.

- Fischer, J., Whitney, D., 2012. Attention gates visual coding in the human pulvinar. *Nat Commun* 3, 1051.
- Fisher, R. A., 1925. *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh: Scotland.
- Flanigan, M., Aleyasin, H., Takahashi, A., Golden, S. A., Russo, S. J., 2017. An emerging role for the lateral habenula in aggressive behavior. *Pharmacol Biochem Behav*.
- Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., Raichle, M. E., 2006. Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proc Natl Acad Sci U S A* 103, 10046-10051.
- Fox, M. D., Greicius, M., 2010. Clinical applications of resting state functional connectivity. *Front Syst Neurosci* 4, 19.
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., Raichle, M. E., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci U S A* 102, 9673-9678.
- Frith, C. D., 2007. The social brain? *Philos Trans R Soc Lond B Biol Sci* 362, 671-678.
- Fulwiler, C. E., King, J. A., Zhang, N., 2012. Amygdala-orbitofrontal resting-state functional connectivity is associated with trait anger. *Neuroreport* 23, 606-610.
- Gan, G., Preston-Campbell, R. N., Moeller, S. J., Steinberg, J. L., Lane, S. D., Maloney, T., Parvaz, M. A., Goldstein, R. Z., Alia-Klein, N., 2016. Reward vs. Retaliation-the Role of the Mesocorticolimbic Salience Network in Human Reactive Aggression. *Front Behav Neurosci* 10, 179.
- Gan, G., Sterzer, P., Marxen, M., Zimmermann, U. S., Smolka, M. N., 2015. Neural and Behavioral Correlates of Alcohol-Induced Aggression Under Provocation. *Neuropsychopharmacology* 40, 2886-2896.
- Gilam, G., Lin, T., Raz, G., Azrielant, S., Fruchter, E., Ariely, D., Hendler, T., 2015. Neural substrates underlying the tendency to accept anger-infused ultimatum offers during dynamic social interactions. *Neuroimage*.
- Gilam, G., Maron-Katz, A., Kliper, E., Lin, T., Fruchter, E., Shamir, R., Hendler, T., 2017. Tracing the Neural Carryover Effects of Interpersonal Anger on Resting-State fMRI in Men and Their Relation to Traumatic Stress Symptoms in a Subsample of Soldiers. *Front Behav Neurosci* 11, 252.
- Golden, S. A., Heshmati, M., Flanigan, M., Christoffel, D. J., Guise, K., Pfau, M. L., Aleyasin, H., Menard, C., Zhang, H., Hodes, G. E., Bregman, D., Khibnik, L., Tai, J., Rebusi, N., Krawitz, B.,

- Chaudhury, D., Walsh, J. J., Han, M. H., Shapiro, M. L., Russo, S. J., 2016. Basal forebrain projections to the lateral habenula modulate aggression reward. *Nature* 534, 688-692.
- Greicius, M. D., Krasnow, B., Reiss, A. L., Menon, V., 2003. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc Natl Acad Sci U S A* 100, 253-258.
- Gusnard, D. A., Akbudak, E., Shulman, G. L., Raichle, M. E., 2001. Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc Natl Acad Sci U S A* 98, 4259-4264.
- Haller, J., 2013. The neurobiology of abnormal manifestations of aggression--a review of hypothalamic mechanisms in cats, rodents, and humans. *Brain Res Bull* 93, 97-109.
- Hallquist, M. N., Hwang, K., Luna, B., 2013. The nuisance of nuisance regression: spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures functional connectivity. *Neuroimage* 82, 208-225.
- Harmon-Jones, E., Harmon-Jones, C., Price, T. F., 2013. What is Approach Motivation? *Emotion Review* 5, 291-295.
- Hayasaka, S., Nichols, T. E., 2003. Validating cluster size inference: random field and permutation methods. *Neuroimage* 20, 2343-2356.
- Huang, A. S., Mitchell, J. A., Haber, S. N., Alia-Klein, N., Goldstein, R. Z., 2018. The thalamus in drug addiction: from rodents to humans. *Philos Trans R Soc Lond B Biol Sci* 373.
- Jacob, Y., Gilam, G., Lin, T., Raz, G., Hendler, T., 2018. Anger Modulates Influence Hierarchies Within and Between Emotional Reactivity and Regulation Networks. *Front Behav Neurosci* 12, 60.
- Jo, H. J., Saad, Z. S., Simmons, W. K., Milbury, L. A., Cox, R. W., 2010. Mapping sources of correlation in resting state FMRI, with artifact detection and removal. *Neuroimage* 52, 571-582.
- Kanai, R., Komura, Y., Shipp, S., Friston, K., 2015. Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philos Trans R Soc Lond B Biol Sci* 370.
- Kessler, R. C., Coccaro, E. F., Fava, M., Jaeger, S., Jin, R., Walters, E., 2006. The prevalence and correlates of DSM-IV intermittent explosive disorder in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 63, 669-678.
- Kim, J. W., Naidich, T. P., Ely, B. A., Yacoub, E., De Martino, F., Fowkes, M. E., Goodman, W. K., Xu, J., 2016. Human habenula segmentation using myelin content. *Neuroimage* 130, 145-156.
- Kose, S., Steinberg, J. L., Moeller, F. G., Gowin, J. L., Zuniga, E., Kamdar, Z. N., Schmitz, J. M., Lane, S. D., 2015. Neural correlates of impulsive aggressive behavior in subjects with a history of alcohol dependence. *Behav Neurosci* 129, 183-196.

- Kramer, U. M., Jansma, H., Tempelmann, C., Munte, T. F., 2007. Tit-for-tat: the neural basis of reactive aggression. *Neuroimage* 38, 203-211.
- Kroemer, N. B., Guevara, A., Vollstadt-Klein, S., Smolka, M. N., 2013. Nicotine alters food-cue reactivity via networks extending from the hypothalamus. *Neuropsychopharmacology* 38, 2307-2314.
- Laird, A. R., Fox, P. M., Eickhoff, S. B., Turner, J. A., Ray, K. L., McKay, D. R., Glahn, D. C., Beckmann, C. F., Smith, S. M., Fox, P. T., 2011. Behavioral interpretations of intrinsic connectivity networks. *J Cogn Neurosci* 23, 4022-4037.
- Latora, V., Marchiori, M., 2001. Efficient behavior of small-world networks. *Phys Rev Lett* 87, 198701.
- Lawson, R. P., Drevets, W. C., Roiser, J. P., 2013. Defining the habenula in human neuroimaging studies. *Neuroimage* 64, 722-727.
- Lawson, R. P., Seymour, B., Loh, E., Lutti, A., Dolan, R. J., Dayan, P., Weiskopf, N., Roiser, J. P., 2014. The habenula encodes negative motivational value associated with primary punishment in humans. *Proc Natl Acad Sci U S A* 111, 11858-11863.
- Lee, J. H., Garwood, M., Menon, R., Adrian, G., Andersen, P., Truwit, C. L., Ugurbil, K., 1995. High contrast and fast three-dimensional magnetic resonance imaging at high fields. *Magn Reson Med* 34, 308-312.
- Lee, R., Arfanakis, K., Evia, A. M., Fanning, J., Keedy, S., Coccaro, E. F., 2016. White Matter Integrity Reductions in Intermittent Explosive Disorder. *Neuropsychopharmacology* 41, 2697-2703.
- Lenzenweger, M. F., Lane, M. C., Loranger, A. W., Kessler, R. C., 2007. DSM-IV personality disorders in the National Comorbidity Survey Replication. *Biol Psychiatry* 62, 553-564.
- Linnman, C., Moulton, E. A., Barmettler, G., Becerra, L., Borsook, D., 2012. Neuroimaging of the periaqueductal gray: state of the field. *Neuroimage* 60, 505-522.
- Lynall, M. E., Bassett, D. S., Kerwin, R., McKenna, P. J., Kitzbichler, M., Muller, U., Bullmore, E., 2010. Functional connectivity and brain networks in schizophrenia. *J Neurosci* 30, 9477-9487.
- Matsumoto, M., Hikosaka, O., 2007. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* 447, 1111-1115.
- McCloskey, M. S., Phan, K. L., Angstadt, M., Fettich, K. C., Keedy, S., Coccaro, E. F., 2016. Amygdala hyperactivation to angry faces in intermittent explosive disorder. *J Psychiatr Res* 79, 34-41.

- Miller, E. K., Cohen, J. D., 2001. An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24, 167-202.
- Moeller, S. J., Frobose, M. I., Konova, A. B., Misyrilis, M., Parvaz, M. A., Goldstein, R. Z., Alia-Klein, N., 2014. Common and distinct neural correlates of inhibitory dysregulation: stroop fMRI study of cocaine addiction and intermittent explosive disorder. *J Psychiatr Res* 58, 55-62.
- Morawetz, C., Kellermann, T., Kogler, L., Radke, S., Blechert, J., Derntl, B., 2016. Intrinsic functional connectivity underlying successful emotion regulation of angry faces. *Soc Cogn Affect Neurosci* 11, 1980-1991.
- Morris, S. E., Cuthbert, B. N., 2012. Research Domain Criteria: cognitive systems, neural circuits, and dimensions of behavior. *Dialogues Clin Neurosci* 14, 29-37.
- Murphy, K., Fox, M. D., 2017. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *Neuroimage* 154, 169-173.
- Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D., Gross, J. J., 2004. For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. *Neuroimage* 23, 483-499.
- Olson, I. R., McCoy, D., Klobusicky, E., Ross, L. A., 2013. Social cognition and the anterior temporal lobes: a review and theoretical framework. *Soc Cogn Affect Neurosci* 8, 123-133.
- Portas, C. M., Rees, G., Howseman, A. M., Josephs, O., Turner, R., Frith, C. D., 1998. A specific role for the thalamus in mediating the interaction of attention and arousal in humans. *J Neurosci* 18, 8979-8989.
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., Shulman, G. L., 2001. A default mode of brain function. *Proc Natl Acad Sci U S A* 98, 676-682.
- Redcay, E., Moran, J. M., Mavros, P. L., Tager-Flusberg, H., Gabrieli, J. D., Whitfield-Gabrieli, S., 2013. Intrinsic functional network organization in high-functioning adolescents with autism spectrum disorder. *Front Hum Neurosci* 7, 573.
- Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059-1069.
- Sanz-Arigita, E. J., Schoonheim, M. M., Damoiseaux, J. S., Rombouts, S. A., Maris, E., Barkhof, F., Scheltens, P., Stam, C. J., 2010. Loss of 'small-world' networks in Alzheimer's disease: graph analysis of FMRI resting-state functional connectivity. *PLoS One* 5, e13788.
- Scott, K. M., Lim, C. C., Hwang, I., Adamowski, T., Al-Hamzawi, A., Bromet, E., Bunting, B., Ferrand, M. P., Florescu, S., Gureje, O., Hinkov, H., Hu, C., Karam, E., Lee, S., Posada-Villa, J.,

- Stein, D., Tachimori, H., Viana, M. C., Xavier, M., Kessler, R. C., 2016. The cross-national epidemiology of DSM-IV intermittent explosive disorder. *Psychol Med* 46, 3161-3172.
- Shannon, B. J., Raichle, M. E., Snyder, A. Z., Fair, D. A., Mills, K. L., Zhang, D., Bache, K., Calhoun, V. D., Nigg, J. T., Nagel, B. J., Stevens, A. A., Kiehl, K. A., 2011. Premotor functional connectivity predicts impulsivity in juvenile offenders. *Proc Natl Acad Sci U S A* 108, 11241-11245.
- Sharp, D. J., Bonnelle, V., De Boissezon, X., Beckmann, C. F., James, S. G., Patel, M. C., Mehta, M. A., 2010. Distinct frontal systems for response inhibition, attentional capture, and error processing. *Proc Natl Acad Sci U S A* 107, 6106-6111.
- Shelton, L., Pendse, G., Maleki, N., Moulton, E. A., Lebel, A., Becerra, L., Borsook, D., 2012. Mapping pain activation and connectivity of the human habenula. *J Neurophysiol* 107, 2633-2648.
- Siegel, A., Victoroff, J., 2009. Understanding human aggression: New insights from neuroscience. *Int J Law Psychiatry* 32, 209-215.
- Siever, L. J., 2008. Neurobiology of aggression and violence. *Am J Psychiatry* 165, 429-442.
- Strang, A., Haynes, O., Cahill, N. D., Narayan, D. A., 2018. Generalized relationships between characteristic path length, efficiency, clustering coefficients, and density. *Social Network Analysis and Mining* 8.
- Swick, D., Ashley, V., Turken, U., 2011. Are the neural correlates of stopping and not going identical? Quantitative meta-analysis of two response inhibition tasks. *Neuroimage* 56, 1655-1665.
- Tomasi, D., Caparelli, E. C., Chang, L., Ernst, T., 2005. fMRI-acoustic noise alters brain activation during working memory tasks. *Neuroimage* 27, 377-386.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273-289.
- Ventura, J., Liberman, R. P., Green, M. F., Shaner, A., Mintz, J., 1998. Training and quality assurance with the Structured Clinical Interview for DSM-IV (SCID-I/P). *Psychiatry Res* 79, 163-173.
- Wang, L., Zhu, C., He, Y., Zang, Y., Cao, Q., Zhang, H., Zhong, Q., Wang, Y., 2009. Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder. *Hum Brain Mapp* 30, 638-649.

Watts, D. J., Strogatz, S. H., 1998. Collective dynamics of 'small-world' networks. *Nature* 393, 440-442.

Whitfield-Gabrieli, S., Nieto-Castanon, A., 2012. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect* 2, 125-141.

Zalesky, A., Fornito, A., Harding, I. H., Cocchi, L., Yucel, M., Pantelis, C., Bullmore, E. T., 2010. Whole-brain anatomical networks: does the choice of nodes matter? *Neuroimage* 50, 970-983.

6. Acknowledgements

We thank Dr. Sam Golden (NIDA) for comments on the manuscript. This research was supported by the following funding agencies: National Institute of Mental health (NIMH; R01MH090134 to NAK), National Institute on Drug Abuse (NIDA; 1R01DA041528 to RZG; 1K01DA043615 to MAP; 1K01DA037452 to SJM), University of Heidelberg, Germany (the Olympia-Morata Program to GG), NWO Netherlands (Rubicon 446-14-015 to AZ).

7. Conflicts of Interest

The authors report no conflicts of interest.

Figure legends

Figure 1: Whole-brain graph theory analyses. A: Reactive aggressive individuals showed higher global efficiency in the left habenula (Hb), the left thalamus (thal), the left dorso-lateral prefrontal cortex (dlPFC, Brodmann area [BA] 9), the right temporal pole (temp. pole, BA 38), and the superior temporal gyrus (STG, BA 22) relative to the control group. B: Reactive aggressive individuals showed lower clustering coefficient in the left precuneus (BA 7), the left dorso-medial prefrontal cortex (dmPFC, BA 10), and left occipital regions relative to the control group. Areas surviving significance at an FDR-corrected threshold are highlighted in boldface and in bar plots on the right. Error bars indicate ± 2 standard errors. Network-level (summary measure averaged across all nodes): * $p < .05$; whole-brain node-level: ** $p < .05$ (FDR-corrected), # $p < .001$ uncorrected. *Abbreviations:* L = left, R = right.

Figure 2: Correlation analyses between trait aggression (BPQ total score) and global efficiency in the left habenula (Hb) (A-B), and whole-brain seed-to-voxel regression analysis with the left habenula as a seed (C-E). A: Left habenula (Hb) node; for the correlation analyses, global efficiency values have been extracted from this node showing significant group differences in the main graph theory analysis (Figure 1). B: Correlation between trait aggression and global efficiency in the left habenula in the reactive aggressive group (full and subclinical IED), and the controls. Correlation coefficients between the reactive aggressive and control group were significantly different. C: In the reactive aggressive group, higher levels of trait aggression (BPQ aggression) were linked to more negative rsFC between the left habenula and a left-lateralized cluster encompassing the ventro-lateral prefrontal cortex (vlPFC), and the temporal pole compared to less reactive aggressive individuals. D: Scatter plot depicting the association

between Hb-vlPFC rsFC and trait aggression in the reactive aggressive group based on the whole-brain regression analysis is presented for demonstration purposes only. Seed-to-voxel results were significant at a corrected cluster-threshold of $p < .05$ (FDR) at a primary voxel-wise threshold of $p < .001$ uncorrected. *Abbreviations:* BPQ = Buss-Perry aggression questionnaire, L = left, R = right.

Table 1. Comparison between the reactive aggressive group and controls on demographics, depressive symptoms, substance use, aggression, anger, and hostility.

Participants (all male)	RA (7 full IED, 4 sub-clinical IED)	Controls (n=12)	Test statistic	p-value
Age (years)	35.36 ± 7.4	30.7 ± 5.2	$t(21) = -1.78$	$p=.09$
Race (Black/Hispanic/Caucasian)	6/4/1	6/4/2	$Pearson \chi^2=.290$	$p=.865$
Education (years)	13.5 ± 1.7	13.9 ± 1.4	$t(21) = .66$	$p=.514$
Depressive Symptoms (BDI-II)***	8.3 ± 5.0	1.8 ± 2.0	$t(21) = -4.81$	$p<.001$
<i>Substance use</i>				
Cigarette use (none/current/past/occasional)§	5/4/1/0	10/1/0/0	$Pearson \chi^2=4.43$	$p=.109$
Alcohol use (none/current/past/occasional)§	3/6/0/1	1/4/0/6	$Pearson \chi^2=4.94$	$p=.085$
- # of alcohol use days (past 30 days)°	4.6 ± 5.3	4.7 ± 4.4	$t(20) = .088$	$p=.931$
- Years of alcohol use (lifetime)°	10.7 ± 8.0	8.7 ± 5.9	$t(20) = -.669$	$p=.511$
Cannabis use (none/current/past/occasional)§	5/1/1/2	7/0/3/1	$Pearson \chi^2=2.50$	$p=.477$
- # of cannabis use days (past 30 days)§	none	none	-	-
- Years of cannabis use (lifetime)	2.4 ± 4.2	0.8 ± 2.1	$t(21) = -.669$	$p=.511$
BPQ total***	150.7 ± 18.0	53.8 ± 13.9	$t(21) = -14.57$	$p<.001$
Physical aggression***	49.8 ± 4.5	20.1 ± 6.1	$t(21) = -13.19$	$p<.001$
Verbal aggression***	28.0 ± 4.0	10.6 ± 3.9	$t(21) = -10.48$	$p<.001$
Anger***	35.2 ± 4.0	10.1 ± 3.8	$t(21) = -15.69$	$p<.001$
Hostility***	37.7 ± 11.5	13.1 ± 6.4	$t(21) = -6.43$	$p<.001$

*** = $p<.001$ (p-value Bonferroni corrected for 15 multiple comparisons, $p<.0033$); §, missing data for alcohol, cigarette use (n=1 RA, n=1 control), and for cannabis use (n=2 RA, n=1 control); °, missing data n=1 control; \$, cannot be computed because standard deviations of both groups are zero. *Abbreviations:* BDI-II = Beck depression inventory, BPQ = Buss-Perry aggression questionnaire, IED = intermittent explosive disorder, RA = reactive aggressive.

Table 2: Group differences between reactive aggressive individuals and controls in resting-state functional connectivity at the network-level (summary measures averaged across all nodes) and across the whole brain at the node-level as measured by graph theory.

Brain areas	BA	MNI coordinates			Test-statistic	p (un-corrected)	p (FDR-corrected)
		x	y	z			

Global efficiency: reactive aggressive > controls							
Network-level					t(21) = 2.56	0.009	
Whole-brain node-level							
L thalamus – pulvinar		-14	-27	4	t(21) = 4.33	0.0001	0.033
L dlPFC	9	-30	28	44	t(21) = 4.31	0.0002	0.033
R temporal pole	38	52	8	-13	t(21) = 4.31	0.0002	0.033
L habenula		-3	-24	2	t(21) = 4.13	0.0002	0.038
R superior temporal gyrus	22	50	-1	-10	t(21) = 3.55	0.0009	0.122

Clustering coefficient: reactive aggressive < controls							
Network-level					t(21) = -2.28	0.016	
Whole-brain node-level							
L superior occipital gyrus	19	-24	-82	36	t(21) = -4.10	0.0003	0.109
L precuneus	7	-8	-71	48	t(21) = -3.83	0.0005	0.109
L calcarine	18	-18	-97	-1	t(21) = -3.81	0.0005	0.109
L dmPFC	10	-10	66	7	t(21) = -3.57	0.0009	0.132

Significance surviving Bonferroni-correction at the network-level and FDR-correction for multiple comparisons across the whole brain is indicated in bold font. *Abbreviations:* BA=Brodman Area; dlPFC=dorso-lateral prefrontal cortex; dmPFC=dorso-medial prefrontal cortex, L = left, R = right.

Table 3: Correlation between trait aggression (BPQ total score) and global efficiency in the left habenula, the left thalamus, the left dorso-lateral prefrontal cortex, and the right temporal pole separately for the reactive aggressive group and control group.

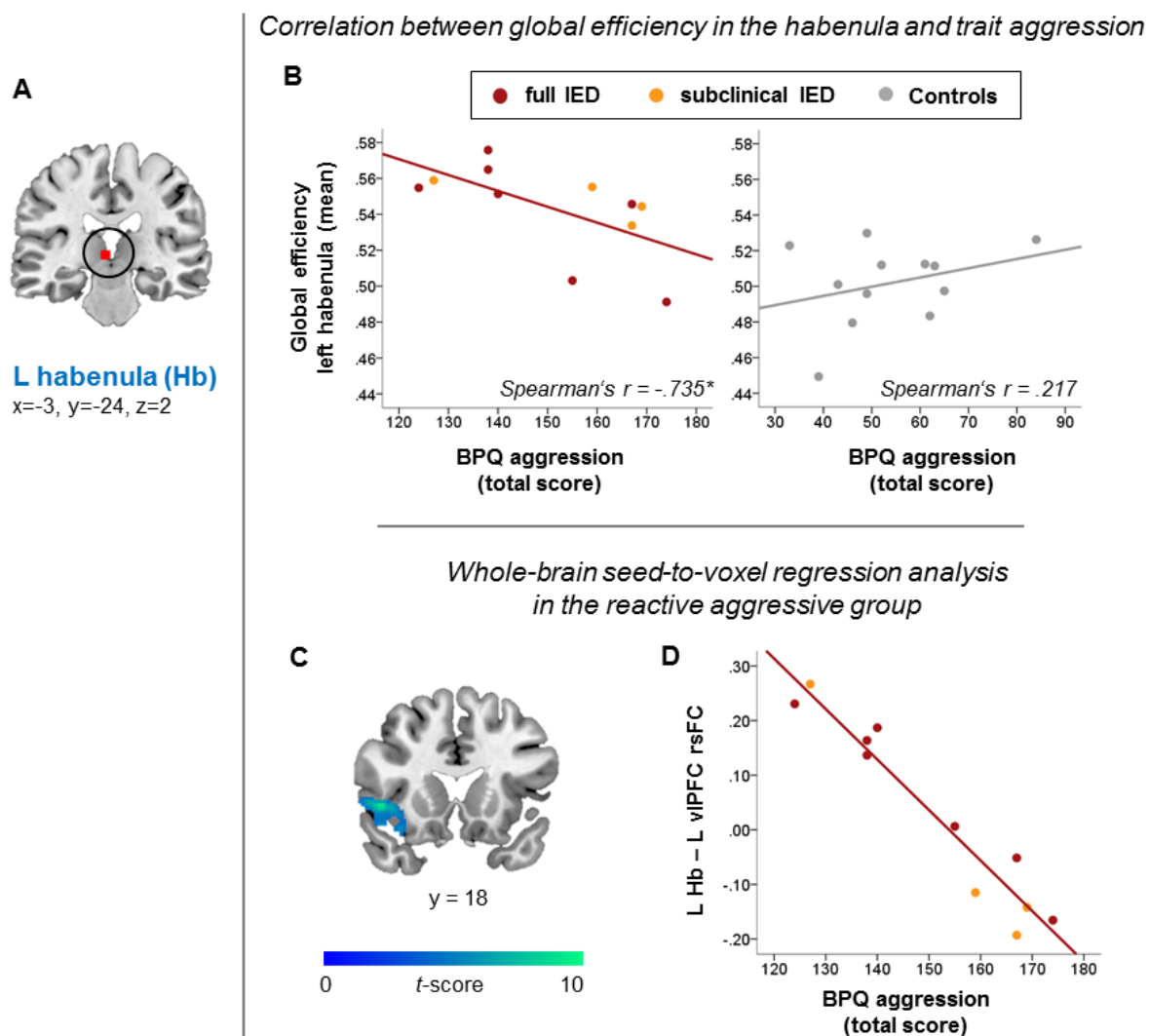
		Global efficiency			
		L habenula	L thalamus	L dIPFC	R temporal pole
Reactive aggressive (n=11)					
BPQ trait aggression (total score)	<i>Spearman's r</i>	-0.735*	-0.384	-0.039	-0.329
	<i>p-value</i>	0.01	0.244	0.909	0.324
Controls (n=12)					
BPQ trait aggression (total score)	<i>Spearman's r</i>	0.217	0.337	-0.294	-0.168
	<i>p-value</i>	0.498	0.284	0.353	0.601

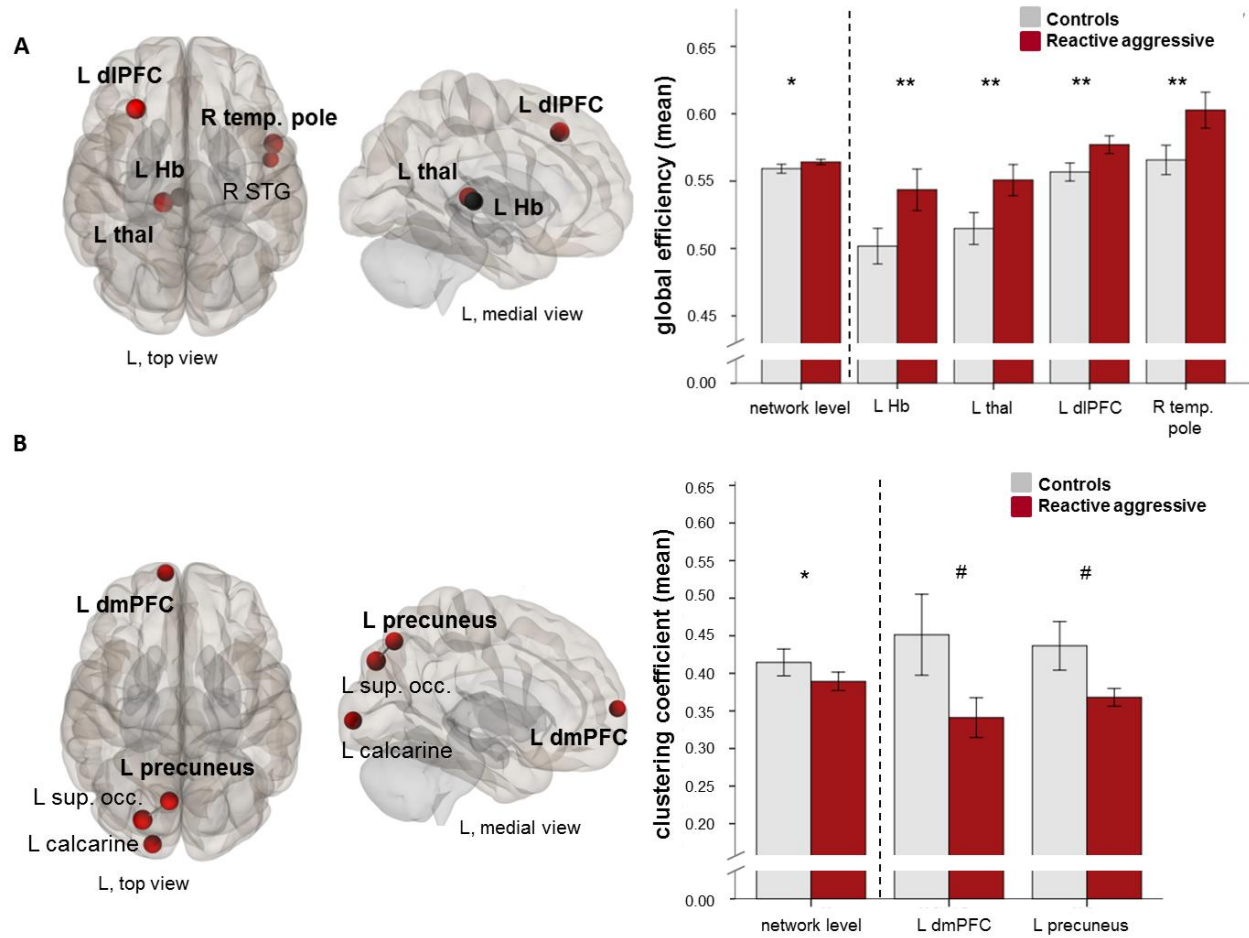
Significant correlations are shown in bold face ($p < .0125$; Bonferroni-corrected for $n=4$ tests in the reactive aggressive group). *, indicates significant differences between correlation coefficients of the reactive aggressive and the control group based on Fisher's r -to- z transformation. *Abbreviations:* dIPFC=dorso-lateral prefrontal cortex, BPQ = Buss-perry aggression questionnaire, L = left, R = right.

Table 4: Post-hoc whole-brain seed-to-voxel regression analysis between trait aggression and the left habenula (i.e., seed) resting-state functional connectivity.

Brain areas	BA	MNI coordinates			k-cluster size	Test-statistic	cluster, <i>p</i> (FDR-corrected)	voxel-wise <i>p</i> (uncorrected)
		x	Y	z				
Negative correlation with the left habenula rsFC in RA individuals								
L vIPFC	47	-46	18	-2	411	<i>t</i> (21) = 9.47	0.001	<.0001
L vIPFC	47	-34	26	-10		<i>t</i> (21) = 7.27		<.0001
L temp. pole	38	-50	16	-14		<i>t</i> (21) = 6.44		<.0001

Significance surviving FDR-correction for multiple comparisons at the cluster-level is indicated in bold face. *Abbreviations:* vIPFC = ventrolateral prefrontal cortex, L=left, temp. pole = temporal pole





Highlights

- Higher global efficiency in subcortical-cortical nodes in reactive aggressive men.
- Aggression severity was linked to habenula global efficiency.
- Severely reactive aggressive men showed lower habenula-vlPFC rsFC.
- Findings suggest an involvement of the habenula in aggression severity.