

# Out of control: Evidence for anterior insula involvement in motor impulsivity and reactive aggression

Franziska Dambacher,<sup>1,2</sup> Alexander T. Sack,<sup>1,2</sup> Jill Lobbstaël,<sup>3</sup> Arnoud Arntz,<sup>3</sup> Suzanne Brugman,<sup>3</sup> and Teresa Schuhmann<sup>1,2</sup>

<sup>1</sup>Department of Cognitive Neuroscience, Maastricht University, <sup>2</sup>Maastricht Brain Imaging Center, and <sup>3</sup>Department of Clinical Psychological Science, Maastricht University, Maastricht, The Netherlands

**Inhibiting impulsive reactions while still defending one's vital resources is paramount to functional self-control and successful development in a social environment. However, this ability of successfully inhibiting, and thus controlling one's impulsivity, often fails, leading to consequences ranging from motor impulsivity to aggressive reactions following provocation. Although inhibitory failure represents the underlying mechanism, the neurocognition of social aggression and motor response inhibition have traditionally been investigated in separation. Here, we aimed to directly investigate and compare the neural mechanisms underlying the failure of inhibition across those different modalities of self-control. We used functional imaging to reveal the overlap in neural correlates between failed motor response inhibition (measured by a go/no-go task) and reactive aggression (measured by the Taylor aggression paradigm) in healthy males. The core overlap of neural correlates was located in the anterior insula, suggesting common anterior insula involvement in motor impulsivity as well as reactive aggression. This evidence regarding an overarching role of the anterior insula across different modalities of self-control enables an integrative perspective on insula function and a better integration of cognitive, social and emotional factors into a comprehensive model of impulsivity. Furthermore, it can eventually lead to a better understanding of clinical syndromes involving inhibitory deficits.**

**Keywords:** response inhibition; reactive aggression; impulsivity; anterior insula cortex; fMRI

## INTRODUCTION

To ensure survival and sufficient supply of necessary resources in a competitive world, impulsive reactions reverting to rather archaic mechanisms such as retaliation might sometimes be the right choice. At the same time, the ability to use one's cognitive capacities to inhibit overly impulsive actions seems equally important for the progressing of individuals within human societies. The right balance between impulsivity and inhibition is paramount to functional self-control and successful development in a social environment. However, this ability of successfully inhibiting, and thus controlling one's impulsivity, can also fail, leading to a variety of behavioural consequences ranging from poor motor response inhibition performance (motor impulsivity) to overly aggressive reactions to provocation in social interaction. Although for different behavioural outcomes the underlying mechanism might be a general failure of inhibition, the neurocognition of social aggression and motor response inhibition have traditionally been investigated in separation. The current study aims to directly investigate the overlap in neural components involved in motor impulsivity and reactive aggression.

Response inhibition—defined as the cognitive ability to deliberately withhold any planned or automatic reaction (Logan *et al.*, 1997)—as well as impulsive behaviour in the context of self-control, retaliation and aggression, has attracted interest in cognitive and social neuroscience. Neuroimaging studies investigating the involvement of neural components in response inhibition mainly used simple motor response

inhibition paradigms such as stop signal and go/no-go tasks (GNGTs). Such imaging studies consistently reported task-related activity within inferior frontal gyrus, anterior insula, pre-supplementary motor area and subcortical circuitries involving thalamic regions and the striatum (Chambers *et al.*, 2009; Simmonds *et al.*, 2008; Swick *et al.*, 2011). In most studies on inhibitory processing, an asymmetric distribution of neuronal activity towards the right hemisphere was found (reviewed in Aron, 2004).

The display of reactive aggression—defined as an aggressive reaction to provocation—has empirically been associated with impaired executive functioning (Hoaken *et al.*, 2003). It might lead to interesting insights to look at how response inhibition-related neural activity might be mirrored in trait impulsiveness and self-reported reactive aggression. Horn and colleagues (2003) showed that impulsive individuals (measured with the Eysenck Impulsivity Scale) reverted stronger right orbitofrontal activity to maintain inhibitory capacities in a GNGT. Furthermore, GNGT-inhibition-related activity in the right dorsolateral prefrontal cortex was found to be negatively correlated with motor impulsiveness measured by the Barratt Impulsiveness Scale (Asahi *et al.*, 2004). Recently, Pawliczek and colleagues (2013) showed that high trait aggressiveness (measured with a questionnaire) is attended not only by inhibition deficits in an emotional stop signal task on a behavioural level but also by lower inhibition-related brain activity in the pre-supplementary motor area and the primary motor cortex. Despite their valuable achievement of explicitly demonstrating a link between aggression and different inhibition-related brain regions that has mostly been built on implicit assumptions, the mentioned studies hold one constraint. The measures they used to determine impulsivity and aggression are self-report questionnaires. It has been demonstrated that such self-report measures correlate rather weakly with behavioural measures of similar concepts (Giancola and Parrot, 2008). This is not surprising, as both ways of measuring aggression refer to slightly different constructs. Therefore, it is rather interesting to apply a neuroscientific methodology to both self-report measures and actual behavioural measures. The latter has rarely been done. It is necessary to investigate common neural mechanisms of motor

Received 27 November 2013; Revised 28 April 2014; Accepted 14 May 2014

Advance Access publication 16 May 2014

We thank Ulrike Krämer and Frederike Beyer for their massive support regarding the implementation of the TAP. Furthermore, we thank Armin Heinecke, Martijn Stroom, Martijn van Teffelen, Clara Bernier, Oana Gurau and Thomas C. Emmerling for their help.

This work was supported by a grant from the Netherlands Organization for Scientific Research (NWO; HCMI 10-19).

Correspondence should be addressed to Franziska Dambacher, Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Oxfordlaan 55, 6229 EV Maastricht, The Netherlands. E-mail: franziska.dambacher@maastrichtuniversity.nl

impulsivity and aggression by looking at how behavioural inhibition measures such as the GNGT relate to other behavioural rather than self-report-based impulsivity or aggression measures.

When it comes to behavioural measurements of reactive aggression, the Taylor aggression paradigm (TAP; Taylor, 1967) proved itself the most feasible option in an imaging environment (Krämer *et al.*, 2007; Lotze *et al.*, 2007). This task demonstrated high construct and internal, discriminant as well as external validity (Anderson *et al.*, 1999; Bernstein *et al.*, 1987; Giancola and Zeichner, 1995). During the task, participants are provoked by one or more virtual opponent, and their aggressive behaviour in reaction to provocation is measured by recording the severity level of the feedback or 'punishment' they administer to their opponents. Punishment can be operationalized as administration of a highly aversive stimulus (e.g. an aversive noise or pneumatic pressure stimulus) to the opponent.

So far, surprisingly few imaging studies have investigated the neural correlates of reactive aggression in the TAP or a similar social interaction task in healthy adults (Krämer *et al.*, 2007; Lotze *et al.*, 2007; Krämer *et al.*, 2011), adolescents (White, 2013) and psychopaths (Veit, 2010). Most prominently, the anterior insula, anterior cingulate cortex, ventral and dorsal mediodorsal cortex, hypothalamic areas and striatum were found to be involved in standard reactions to provocative situations and retaliation. Considering especially the insula and the cingulate cortex activation, neural networks involved in response inhibition and reaction to provocation should substantially overlap. However, such an overlap has never been directly investigated with adequate methodology, and, thus, claims on the specific and precise involvement of brain areas and their functional structure potentially contributing to several modalities of self-control remain to be tested.

Until today, no systematic inquiry evaluated the two concepts within the same psychological modality comparing a behavioural response inhibition paradigm such as the GNGT with a behavioural paradigm measuring aggression such as the TAP. The present study combines the advantages of using both of these tasks with the high spatial resolution of functional magnetic resonance imaging (fMRI) in a within-subject, within-session design. This enabled exploratory investigations to identify and compare network neural components involved in successful and unsuccessful response inhibition as well as aggressive reactions within provocative social interactions.

## MATERIALS AND METHODS

### Participants

Eighteen male participants volunteered, gave their written informed consent and were paid for participating in this study. An extensive screening on neurological and psychological syndromes ensured that all participants were healthy. Only males were included to avoid gender-related confounding effects. Two participants had to be excluded from the analyses, as they did not perform the tasks according to the instructions of the experimenter. One participant was excluded, as he did not show any reaction to provocation (he did not show any variance in his behaviour and chose only equally low reactions). Data of 15 participants were included in further analyses (mean age = 22.33; s.d. = 2.35). Most participants were university students and were recruited via advertisements on university-related websites and flyers.

### Procedure

Prior to the experiment, participants were told that they would take part in an experiment investigating the impact of human feedback on reaction time performance, together with two other participants. They were instructed that one task was performed (GNGT) independently and one task (TAP) would be a reaction time game in which all three

participants would play together. Before entering the scanner, the participant and the two opponents (collaborators of the experimenters) were introduced, had a chance for casual conversation and followed the experimenters' instruction together. The experimenters' collaborators were trained beforehand and acted according to a script to keep interactions constant with each participant. Throughout the experiment, the real participant would hear the experimenter talk with the 'fake' opponents via intercom during the breaks. Participants performed GNGT and TAP. Task-order was counterbalanced. Immediately after completion of the experiment, participants underwent an oral manipulation check to make sure that they were fully deceived by the experimental setup. After all the measurements had been finalized, they were provided with a written debriefing to explain the deception and disclose the experiments' real purpose and motivation. The study was approved by the local Ethical Committee of the Faculty of Psychology and Neuroscience at Maastricht University and conformed to the Code of Ethics of the World Medical Association (Declaration of Helsinki).

### Go/no-go task

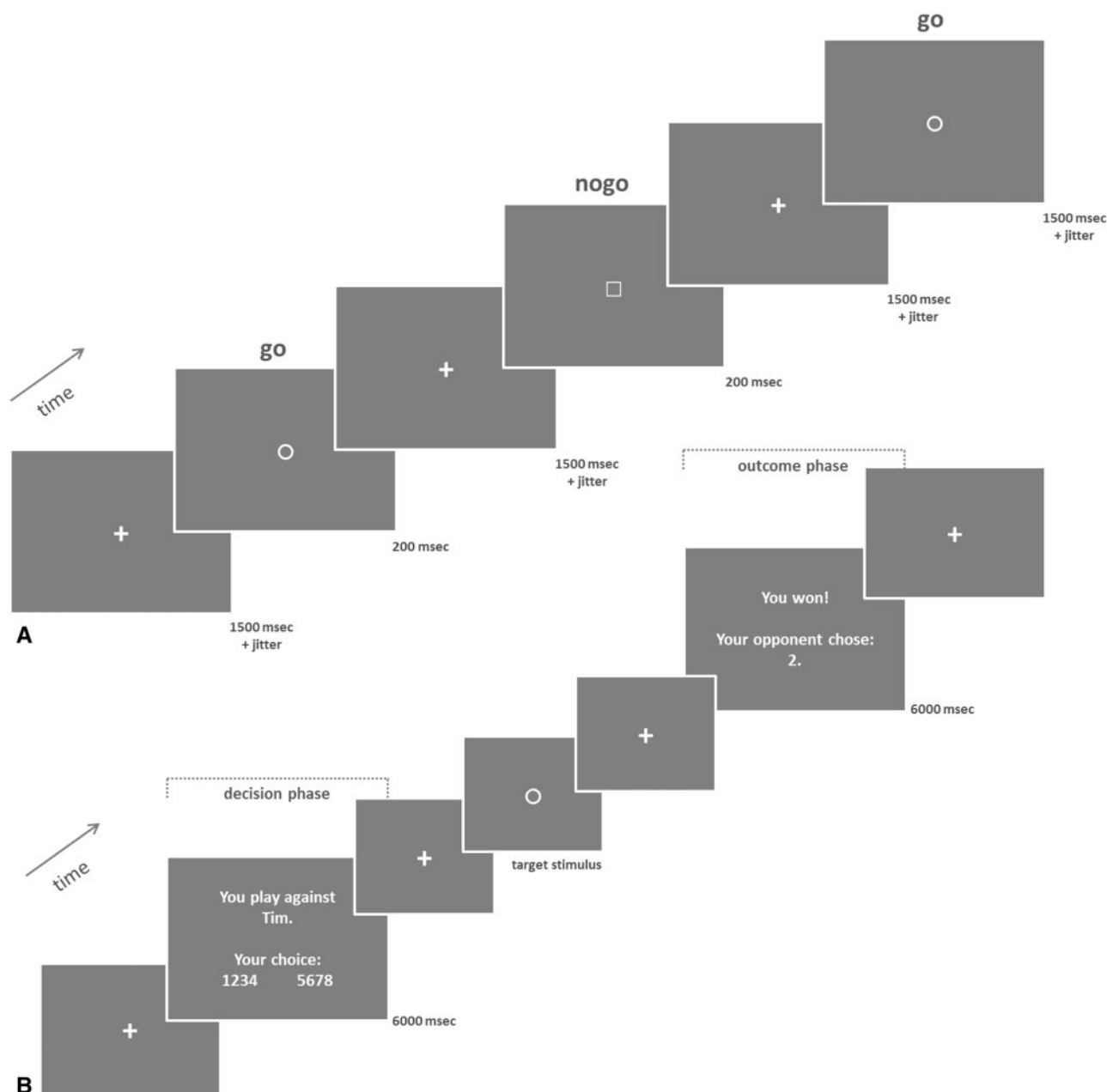
To elicit action restraint, a simple go/no-go motor-response task was used (Figure 1A). Participants were instructed to respond as fast and accurately as possible to a go stimulus via button press with the right index finger, while they should not respond to a rare no-go stimulus. Go as well as no-go stimuli were presented for 200 ms followed by a fixation cross for a randomized interval of 1300, 2800 or 4300 ms. No-go events occurred in 25% of all trials. The trials were pseudo-randomized to avoid unreasonably long concatenations of only inhibition events.

Stimuli and fixation crosses were presented in white (RGB 255/255/255; Arial pt 24) on a grey background (RGB 125/125/125). Participants performed two runs of 200 trials each for the GNGT, leading to a total of 400 trials (100 no-go trials). Stimuli were presented using Presentation software (Neurobehavioral Systems, Inc., Albany, USA). Behavioural statistical analyses were performed using SPSS19 (IBM Statistics, USA).

A square and a circle were used as stimuli. Which stimulus was assigned to the go respectively the no-go condition was randomized between participants. After every 40 trials there was a resting period of 19 500 ms integrated into the task, where only a fixation cross was shown.

### Taylor aggression paradigm

To elicit reactive aggression, an adaptation of the TAP (Taylor, 1967; Krämer *et al.*, 2007) was used (Figure 1B). Participants were told that they would play a competitive reaction time game against two opponents. Whoever would be slower in reacting to a target stimulus by button press with the right index finger would lose the trial and be administered an aversive feedback noise chosen by the winning opponent. The intensity of the feedback noises could be set by the participant for each trial on a scale from 1 (lowest noise) to 8 (loudest noise). For analyses, the feedback given by the participants was grouped into low (level 1–3), middle (4 and 5) and high (level 6–8) punishment. All task parameters were controlled by the experimenters as follows: participants randomly played against each of the putative opponents in 50% of the trials. One of the opponents always selected relatively low feedback noises (from 1 to 4; non-provoking opponent) and the other relatively loud feedback noises (from 4 to 8; provoking opponent). Participants randomly won (and lost) in 50% of trials for each opponent with the constraint that if they needed longer than 500 ms to respond, they always lost. This ensured the credibility of competing against a human opponent.



**Fig. 1** GNGT (A) and TAP (B). During the decision phase, participants were presented with a screen that informed them against whom they were playing in this trial (in this case 'Tim') and asked to choose the feedback noise level that should be administered to this opponent in case the opponent would lose ('1234 5678'). During the outcome phase, participants were informed on whether they won and what feedback noise level the particular opponent chose for this trial.

Each trial of 2700 ms consisted of a decision phase (6000 ms), the actual reaction time game (jittered between 4500 and 7500 ms), an outcome phase (6000 ms) and a jittered resting period. During the decision phase, participants were presented with a screen that informed them against whom they were playing in this particular trial ('Rob' or 'Tim') and were asked to choose the feedback noise level that should be administered to this opponent in case he would lose. As the actual reactive aggressive behaviour is measured during the decision phase of the TAP, the analyses focused exclusively on this phase. During the outcome phase, participants were informed about whether they won or lost in this particular trial and what feedback noise level the particular opponent chose for this trial. In case of losing, the feedback noise was administered right after the outcome information was presented.

Stimuli and fixation crosses were presented in white (RGB 255/255/255; Arial pt 24) on a grey background (RGB 125/125/125). Participants performed three runs of the TAP including 28 trials (14 trials per opponent) each, leading to a total of 84 trials (42 trials per opponent). Stimuli were presented using Presentation software (Neurobehavioral Systems, Inc., Albany, USA). Behavioural statistical analyses were performed using SPSS19 (IBM Statistics, USA).

Feedback noises were adjusted to a subjective threshold of endurance while running a functional sequence before the experiment started. This way, it was possible to control for scanner noise. No noises > 100 decibels were administered to ensure that the hearing of the participants could be in no way compromised.

### Technical details, fMRI acquisition and analysis

Stimulus material was displayed onto a frosted screen, positioned at the rear of the scanner bore, using a liquid-crystal display projector (Panasonic, No. PT-EZ57OEL). Responses were collected using a standard fMRI-compatible button box (Current Designs, 8-button response device, HHSC-2x4-C, Philadelphia, USA).

Using a 3-Tesla Siemens Prisma fMRI scanner, structural (high resolution T1-weighted MPRAGE; isotropic voxel resolution  $1 \times 1 \times 1 \text{ mm}^3$ ; 192 sagittal slices) and functional whole-brain (Gradient-Echo-EPI-sequence; repetition time = 1500 ms; echo time = 26 ms; field of view = 224 mm; flip angle =  $73^\circ$ ; matrix =  $64 \times 64$ ; distance factor = 20%; 478 volumes per run for the GNGT, 512 volumes per run for the TAP) images were acquired. Twenty-eight oblique transversal slices of  $3.5 \times 3.5 \times 3.5 \text{ mm}$  voxel tilted  $30^\circ$  relative to the anterior–posterior commissure plane were obtained to avoid signal drop-out in frontal areas (Deichmann *et al.*, 2003).

Data analyses were performed using Brain Voyager QX (Brain Innovation BV, Maastricht, The Netherlands). Prep scans at the beginning of each functional run ensured equilibrium effects for signal saturation. Preprocessing included 3D motion correction (as implemented in Brain Voyager QX with trilinear/sinc interpolation and intra-session alignment to the first functional volume recorded after the individual anatomical scan), cubic spline slice scan time correction and the application of a temporal high pass filter [general linear model (GLM) with Fourier basis set of three cycles sine/cosine per run including linear trend removal]. Images were co-registered to the individual anatomical scans and normalization to Talairach stereotaxic space (Talairach and Tournoux, 1988). Volume time courses were spatially smoothed using a 6-mm full-width half maximum Gaussian kernel.

To establish activation patterns for the two tasks, random effects group analyses were performed. A GLM was defined to analyse specific activation patterns for no-go trials and false alarms in the GNGT and specific activation patterns with respect to the behaviour displayed during the decision phase in the TAP. In the TAP, the first three trials per opponent were excluded from the analyses to ensure that analyses were based on the trials during which the participant could already estimate the behavioural pattern of the opponents. The following conditions were included as predictors (i) for the GNGT: successful go trials, successful no-go trials, false alarms (commission errors on the no-go trials) and (ii) for the TAP: number of trials that the participant chooses high punishment for the provoking opponent, the participant chooses low punishment for the non-provoking opponent. Most participants rarely or never chose a low or middle punishment for the provoking opponent. Furthermore, a middle or high punishment for the non-provoking opponent was almost never chosen. Therefore, these conditions could not be taken into account on the level of group analyses. The following specific contrasts were examined: for the GNGT, neural correlates of successful inhibition were investigated by contrasting successful no-go trials with successful go trials (no-go trials > go trials) and neural correlates of unsuccessful inhibition were investigated by contrasting unsuccessful no-go trials with successful go trials (false alarms > go trials). For the TAP neural mechanisms, underlying aggressive reactions to the provoking opponent were examined by contrasting trials in which the participant gave a high punishment to the provoking opponent with trials in which the participant gave a low punishment to the non-provoking opponent (provocation > no provocation). Furthermore, the overlap of neural activation during unsuccessful inhibition and aggressive reaction to the provoking opponent was revealed through a conjunction analysis (false alarms > go trials in conjunction with provocation > no provocation).

To reduce error variance, one noise regressor consisting of the first eigenvariate time series from cerebrospinal fluid regions and motion artefacts were included into the analyses as covariates. Statistical maps were created using a threshold of  $P < 0.001$  corrected for multiple comparisons by means of cluster threshold level estimation (1000 Monte Carlo simulation iterations; Forman *et al.*, 1995).

## RESULTS

### Behavioural data

#### GNGT

On average, in the GNGT, participants showed reaction times on go trials of 404.42 ms (s.d. = 35.29) and committed 1.58% misses (omission errors on go trials, s.d. = 2.19) as well as 24.4% false alarms (commission errors on no-go trials, s.d. = 12.36).

#### TAP

The average feedback (i.e. punishment by mean of auditory feedback noise) selected by the participants for the opponents was of medium intensity (mean = 3.54; s.d. = .04). However, for the provoking opponent, significantly higher feedback was chosen than for the non-provoking opponent (provoking opponent: mean = 4.52, s.d. = .64; non-provoking opponent: mean = 2.56, s.d. = 1.17;  $t = 4.59$ ,  $P < 0.001$ ). Within the 42 trials playing against the provoking opponent, participants selected on average 22 times a high feedback (6, 7 or 8; s.d. = 7.73), 9 times a middle feedback (4 or 5; s.d. = 5.80) and 11 times a low feedback (1, 2 or 3; s.d. = 6.62). In contrast, for the non-provoking opponent, a reversed behavioural pattern was observed: participants now selected on average 6 times a high feedback (6, 7 or 8; s.d. = 6.98), 6 times a middle feedback (4 or 5; s.d. = 5.84) and 30 times a low feedback (1, 2 or 3; s.d. = 11.48). The mean reaction time to the target stimulus was 204.53 ms (s.d. = 31.83). During the oral manipulation check, no participant doubted the cover story for the experiment and all 15 participants reported that they perceived one opponent as more provocative than the other. Twelve participants reported explicitly that they adapted their reaction to that perception. Similarly, an analysis of variance with repeated measures revealed a significant interaction effect of high, middle and low feedback level chosen by the participants and the type of opponent ( $F = 19.5$ ,  $P < 0.001$ ). Thereby, high punishment was chosen more frequently for the provoking than the non-provoking opponent, and low punishment was chosen more often for the non-provoking than for the provoking opponent.

### fMRI data

Talairach coordinates of the brain regions showing increased activation associated with the investigated contrasts are reported in Table 1. Statistical maps of random effects group analyses are depicted in Figures 2 (for successful inhibition) and 3 (for unsuccessful inhibition, aggressive reaction to provocation and the corresponding activation overlap).

For successful inhibition (no-go trials > go trials), the examination of blood oxygenation level-dependent signal change in 15 participants revealed a significant increase of activation in the right anterior insula cortex (AIC), right superior frontal gyrus and left fusiform gyrus (Figure 2, Table1).

During unsuccessful inhibition (false alarms > go trials) associated with commission errors in the inhibition trials, increased activation in the bilateral insula cortex, bilateral middle and superior frontal gyri, anterior cingulate cortex (ACC) and thalamus was found. Furthermore, the supramarginal gyri and occipital regions showed significantly increased activation for unsuccessful inhibition in the GNGT



Table 1 Talairach coordinates of activated brain regions

Region	Talairach coordinates			Cluster size Voxel	t	
		x	y			z
No-go trials > go trials						
Frontal cortex						
Anterior insula cortex	R	31	17	1	596	6.12
Superior frontal gyrus	R	35	38	26	688	7.34
Temporal/occipital cortex						
Fusiform gyrus	L	−42	−52	−12	645	5.47
False alarms > go trials						
Frontal cortex						
Insula cortex	R	38	9	6	14 153	9.11
Insula cortex	L	−43	5	6	30 260	11.27
Middle frontal gyrus	R	23	38	36	2343	6.44
Middle frontal gyrus	L	−25	38	36	5068	6.74
Superior frontal gyrus	R	8	−7	65	2471	6.33
Superior frontal gyrus	L	−14	−12	65	4018	6.61
Cingulate cortex						
Anterior cingulate cortex		0	11	37	22 392	8.58
Parietal cortex						
Supermarginal gyrus	R	51	−46	32	1411	5.67
Supermarginal gyrus	L	−54	−46	32	8080	7.12
Occipital gyrus						
Lingual gyrus	R	11	−65	3	27 854	6.88
Lingual gyrus	L	−18	−64	0	(R & L)	(R & L)
Subcortical regions						
Thalamus	R	10	−12	8	1016	5.84
Thalamus	L	−15	−14	8	7972	5.66
Provocation > no provocation						
Frontal cortex						
Anterior insula cortex	R	27	17	6	718	6.11
Anterior insula cortex	L	−33	18	16	1104	5.43
Insula cortex	L	−41	−7	18	769	6.46
Parietal cortex						
Inferior parietal lobe	L	−41	−35	50	19 896	8.58
Subcortical regions						
Putamen/globus pallidus	R	13	3	5	1307	6.58
Putamen/globus pallidus	L	−15	0	5	1098	6.06
Thalamus	L	−19	−19	8	1104	5.37
Caudate	L	−17	−8	24	601	5.82
Cerebellum						
Anterior lobe	R	13	−48	−18	800	5.24
Posterior lobe	R	15	−74	−16	2277	7.15
False alarms > go in conjunction with provocation > no provocation						
Frontal cortex						
Anterior insula cortex	R	29	24	9	358	5.59
Anterior insula cortex	L	−35	16	12	1075	5.43
Insula cortex	L	−45	−4	14	416	5.29
Subcortical regions						
Thalamus	L	−17	−18	6	395	5.06
Putamen/globus pallidus	L	−15	−2	7	124	4.98

(Figure 3, Table 1). When contrasting trials in which the participant gave a high punishment to the provoking opponent with trials in which the participant gave a low punishment to the non-provoking opponent (provocation > no provocation), increased activation in the bilateral insula cortex, inferior parietal lobe and cerebellum was observed. Furthermore, several subcortical regions (i.e. the right and left putamen/globus pallidus, left-lateralized thalamic regions and caudate) showed significant activation change associated to this contrast (Figure 3, Table1). No significant activation change associated with giving a low punishment to the non-provoking opponent (no provocation > provocation) was observed.

When examining the significant overlap of neural activation associated with unsuccessful inhibition and aggressive reaction to provocation (false alarms > go trials provocation > no provocation), bilateral AIC, extended parts of the insula cortex in the left hemisphere as

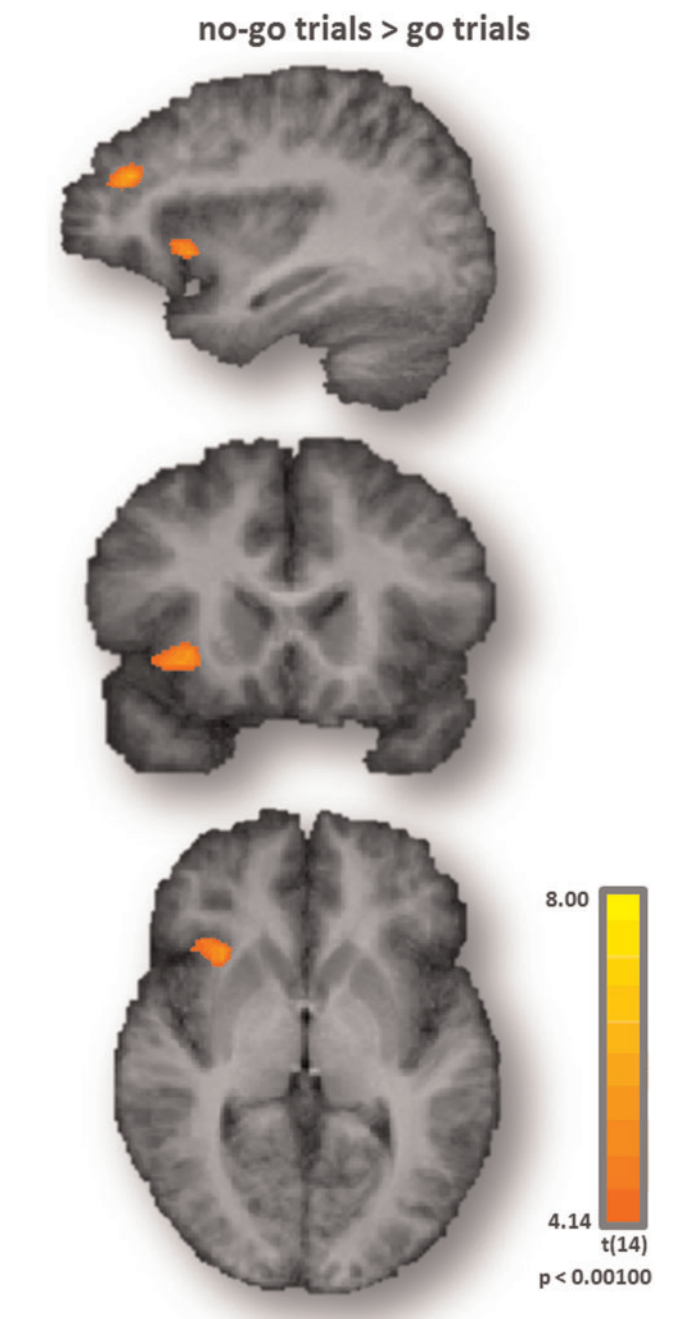
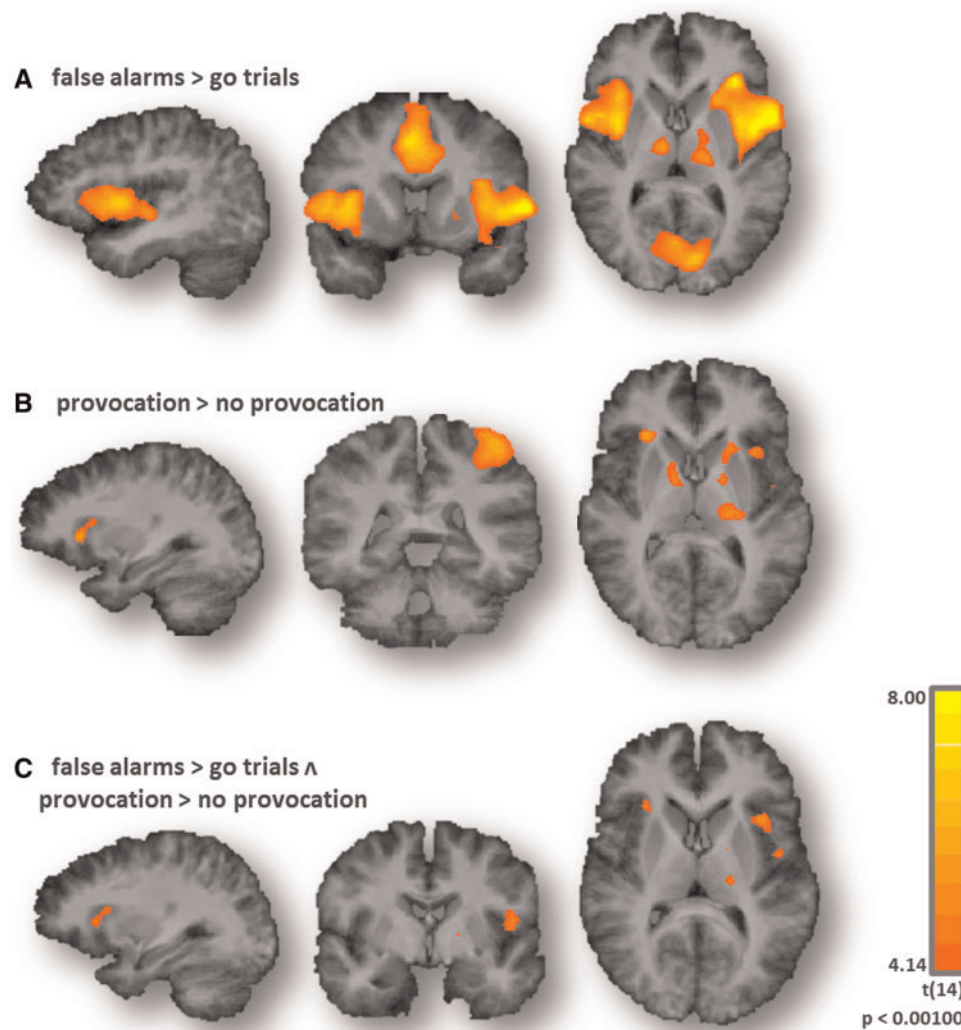


Fig. 2 Successful inhibition. Increased cortical activation associated with successful inhibition during no-go trials vs go trials [no-go trials > go trials; plane x(33) y(18) z(3)]. Statistical map  $P \leq .001$ ; cluster level threshold (CLT) corrected; Random Effects (RFX) group Analysis;  $N = 15$ . Maps projected onto averaged anatomical scan of all 15 participants; radiological convention.

well as left-lateralized subcortical structures (i.e. thalamus and putamen/globus pallidus) showed to be activated for both aspects of disinhibition likewise (Figure 3, Table1). Furthermore, activation in the right AIC during unsuccessful inhibition positively correlated with activation in the right AIC during the condition in which a high punishment was assigned to the provoking opponent (right AIC:  $r = 0.642$ ,  $P = 0.01$ ; correlations were obtained by calculating bivariate Pearson correlations between beta-weights for both conditions across participants extracted from a region of interest defined by activation in right AIC). No correlations were found for the other regions showing maximal overlap of activation in the conjunction analysis (left AIC:  $r = 0.386$ ,  $P = 0.156$ ; subcortical structures:  $r = 0.408$ ,  $P = 0.131$ ;



**Fig. 3** GLM maps. (A) Increased cortical activation associated with unsuccessful inhibition during no-go trials vs go trials [false alarms > go trials; plane  $x(37) y(9) z(8)$ ]. (B) Increased cortical activation associated with aggressive reaction to provoking opponent vs non-aggressive reaction to non-provoking opponent [provocation > no provocation; plane  $x(28) y(-37) z(8)$ ]. (C) Increased cortical activation overlapping for aggressive reaction to provocation and unsuccessful inhibition during no-go trials [false alarms > go trials & provocation > no provocation; plane  $x(28) y(-4) z(10)$ ]. Statistical maps  $P \leq .001$ ; CLT corrected; RFX group Analysis;  $N = 15$ . Maps projected onto averaged anatomical scan of all 15 participants; radiological convention.

correlations were obtained by calculating bivariate Pearson correlations between beta-weights for both conditions across participants extracted from a region of interest defined by activation in the left AIC or a common region of interest including all relevant subcortical structures).

## DISCUSSION

We used two behavioural paradigms (GNGT, TAP) to investigate the neural components involved in successful and unsuccessful response inhibition as well as aggressive reactions in provocative social interactions. We also examined how far neural activation associated with unsuccessful inhibition and aggressive reaction to provocation overlaps.

### Response inhibition

We found the right AIC and right superior frontal gyrus to be activated during successful response inhibition as measured by a simple go/no-go motor response task. These findings are in accordance with previous literature on motor action inhibition (for review, see Simmonds *et al.*, 2008 & Swick *et al.*, 2011). We found a right lateralization of

inhibition-related activation, which has been reported previously to be associated to various inhibition paradigms (Aron *et al.*, 2004; Chambers *et al.*, 2009; Swick *et al.*, 2011). Evidence is growing that besides the right inferior frontal gyrus (Aron *et al.*, 2004; Chikazoe *et al.*, 2007), AIC plays a crucial role in response inhibition across different aspects of inhibition (Sharp *et al.*, 2010; Swick *et al.*, 2011). In their activation likelihood estimation over 11 studies investigating the GNGT, Simmonds and colleagues (2008) report AIC to be involved specifically in complex go/no-go paradigms. However, in our paradigm, which would be defined as a simple GNGT according to Simmonds' criteria, we find the same activation. Repeatedly, superior or middle frontal regions similar to the one we find in our analyses have been reported to be crucially involved in GNGT performance (Simmonds *et al.*, 2008; Swick *et al.*, 2011).

During unsuccessful inhibition, we found a significant activation pattern across many cortical and subcortical areas. It should be emphasized that the right AIC and right superior frontal gyrus—the two regions that were specifically involved in successful inhibition—are part of the activation pattern of unsuccessful inhibition. Menon and colleagues (2001) found a pattern similar to ours and acknowledged the substantial overlap of error-related brain activation with activation

found during successful inhibition in a GNGT. They also found a more distributed network for unsuccessful than for successful inhibition and emphasized that this is in accordance with ideas resulting from previous electroencephalography (EEG) studies. The involvement of AIC in successful as well as unsuccessful inhibition might also reflect dynamic behavioural control across different modalities of cognition rather than response inhibition specifically.

Furthermore, the current finding might be explained by a post-error overshoot of the neural network crucial for the execution of a certain task. This would correspond to the subjective feeling of error awareness or mental overcompensation—expressed by elevated alertness, alarm and intense visualization of what should have been done—following an inaccurate reaction (Klein *et al.*, 2007). Furthermore, especially regions such as ACC and AIC that have repeatedly been associated with the successful execution of cognitive tasks have also been reported to be crucially involved in momentary lapses of attention (Weissman *et al.*, 2006) and maladaptive changes of event-related brain networks (Eichele *et al.*, 2008). Along these lines it should be emphasized that, when errors occur, the participant has to deal with a subjective feeling of failing or blundering, which is not necessarily present in an aggressive reaction to provocation. This might be an explanation of our finding that the ACC is specifically involved in motor impulsivity but not the TAP.

Generally, it is debatable in how far specificity of functional involvement with respect to certain brain regions can be assumed in this context. The mentioned brain areas might rather be involved in general cognitive control and monitoring of actions than in the specific cognitive aspect a task is supposed to tap into. The activations could not only be related to the inability to withhold a response but also to error monitoring and reactions to committing an error and subsequently implementing a cognitive control process. Overall activations might be due to more than just greater impulsivity.

Additionally, it is crucial to emphasize that we cannot rule out emotional processes that came along with error processing. False alarms, thus not succeeding in the specific task, could be accompanied by a sense of failure and subsequently negative emotions such as irritation and anger. Therefore, related neural activity might also reflect such emotional processes.

### Aggressive reaction to provocation

To elicit aggressive reactions in a provocative social interaction we used the TAP. Participants consistently chose significantly higher feedback, which means more uncomfortable noise stimuli, for the provoking compared to the non-provoking opponent and afterwards reported to deliberately have done so. Not just qualitatively but also quantitatively we could confirm that the behavioural pattern varied with respect to the characteristics of the opponent. This implies that our experimental manipulation was successful and that we managed to simulate interactions as similar to natural behaviour as possible in a laboratory situation. However, at the same time, this very pronounced and natural behavioural pattern restricted our analyses to the two most obvious reactions: aggressive reactions to provocation and no aggressive behaviour following no provocation. Therefore, in our data, the intensity of provocation and the level of the chosen punishment cannot be disentangled. Ultimately though, this is the case for any retaliating behaviour observable in a natural environment, and, thus, provocation and the resulting retaliation is understood as a unity in the presented analyses. When directly contrasting these two conditions, no significant change of neural activation specifically associated with a non-aggressive reaction to the non-provoking opponent was observed.

During periods in which participants reacted aggressively towards the provoking opponent, increased neural activation was found in the

bilateral insula cortex, left inferior parietal lobe, a variety of mostly left lateralized subcortical structures (i.e., bilateral putamen/globus pallidus as well as thalamus and caudate in the left hemisphere) and the cerebellum. Bearing in mind that specific contrasts were chosen slightly differently, these results overlap substantially with the findings of Krämer and colleagues (2007). Activation of AIC in the TAP or similar fairness games has been associated with the processing of negative emotions such as anger and disgust (Sanfey *et al.*, 2003; Krämer *et al.*, 2007; White *et al.*, 2013). This follows a tradition of associating AIC with experience of mostly, but not exclusively, negative salient emotions (Damasio *et al.*, 2000; Jabbi *et al.*, 2007; Craig, 2009). Furthermore, increased activation of subcortical structures such as the putamen/globus pallidus and the caudate as part of the striatum has been observed. This has previously been reported for punishment of unfair offers in a social interaction game investigating altruistic punishment (de Quervain *et al.*, 2004). Krämer and colleagues (2007) argue that by punishing the provoking opponent, participants might aim to modify the opponent's behaviour. In case the provoking opponent could be educated, his punishment towards the participant might decrease after consequent retaliation, which in turn would be rewarding for the participant. Besides AIC and subcortical activation, the inferior parietal lobe was significantly involved in retaliation. This could reflect putting oneself in the opponent's position; in particular, the temporoparietal junction was repeatedly associated with the theory of mind and social cognition (for review, see Corbetta and Schulman, 2008). To understand the other's position and motivation is a prerequisite for the successful and dynamic adaption of behaviour within social interactions. In the context of the TAP, as we implemented it in a healthy population, this is especially crucial when interacting with the provoking opponent. To adapt one's default reaction (choosing relatively low feedback as done for the non-provoking opponent) and retaliate, one has to quickly identify the opponent, extract his putative motivation from his previous behaviour and choose a reasonable response. The relatively strong involvement of the left inferior parietal lobe found in our data might reflect such a process amongst others. This explanation might be supported by the fact that the inferior parietal lobe is the only activation hotspot within the provocation-related neural network that was not at the same time involved in unsuccessful inhibition (see conjunction analysis). As a simple motor response task does not require social cognition of any kind, this finding might imply that the specific social component present in the TAP and absent in the GNGT is mirrored in the inferior parietal activation.

Not just with respect to parietal involvement does our data show greater left- than right-hemispheric neural involvement during aggressive reactions to provocation. It might be interesting to consider this result in light of theoretical frameworks based on inter-hemispheric balance and motivational direction. Approach motivation and anger on state as well as trait levels are consistently reported to be accompanied by a shift in inter-hemispheric balance in favour of the left hemisphere (Harmon-Jones and Sigelman, 2001; Harmon-Jones, 2004; van Honk & Schutter, 2006; Carver and Harmon-Jones, 2009; Hortensius *et al.*, 2012). This is well mirrored in our results. On the other hand, in line with this perspective, successful inhibition, a cognitive process conceptually based on avoidance motivation, shows exclusively right-hemispheric neural correlates and, thus, the exact opposite pattern.

### Anterior insula involvement across modalities of (self-)control

When investigating conjoint activity associated to disinhibition on the level of motor action and social interaction, the bilateral insula cortex and left-lateralized subcortical structures (i.e. the thalamus and putamen/globus pallidus) revealed the strongest functional overlap.



The strong and constant involvement of AIC in motor impulsivity as well as reactive aggression is striking. Our data not only show an activation overlap within this region but also that the more a participant recruits the right AIC during unsuccessful inhibition, the more it is recruited during aggressive behaviour. AIC has been the target of numerous investigations across almost countless domains, including interoception, awareness of body movement, self-recognition, vocalization and music, emotional awareness, uncertainty and anticipation, visual and auditory awareness, time perception, attention, perceptual decision-making, cognitive control and performance monitoring (following the review of [Craig, 2009](#)). In the context of the present study, the AIC has been repeatedly associated with successful inhibition ([Sharp et al., 2010](#); [Swick et al., 2011](#)). Especially voluntary inhibition, so to speak the 'free won't' as opposed to the free will ([Brass and Haggard, 2007](#)), has been the function attributed to AIC activation. As mentioned previously, AIC involvement in the context of social interaction paradigms related to retaliation and the punishment of unfairness has mostly been interpreted as reflecting the processing of negative emotions ([Sanfey et al., 2003](#); [de Quervain et al., 2004](#); [Krämer et al., 2007](#); [White et al., 2013](#)). These interpretations are highly reasonable, although seemingly conflicting with results that reveal AIC to be equally involved in unsuccessful inhibition ([Menon et al., 2001](#)) and the processing of positive emotions ([Hennenlotter et al., 2005](#); [Jabbi et al., 2007](#)).

Generally, taking a broader perspective on insula function might resolve these potential conflicts. All contexts for which significant anterior insula activation has been reported (reviewed in [Craig, 2009](#)) require the ability to focus on the immediate presence and the involvement of rather unusual, potentially threatening or change-demanding stimuli. This involves an instantaneous need to monitor whether the current or planned behaviour is still adequate and if necessary to quickly adapt responses. Thus, saliency monitoring, response switching, attention and control are key aspects in this process ([Menon and Uddin, 2010](#)). In their review on anterior insula activation in perceptual paradigms, [Sterzer and Kleinschmidt \(2010\)](#) emphasize that the anterior insula is recruited as soon as any perceptual input poses a challenge to the given *modus operandi*. [Craig \(2009\)](#) goes a step further by suggesting AIC to be the crucial node in a human awareness network whose main responsibility is the subjective regulation of psychological and physiological reactions to cognitively challenging situations. Following this reasoning, he proposes that AIC might be a potential candidate for the neural correlate of consciousness. Our data clearly support the interpretation that AIC might be crucial for dynamic behavioural control across different modalities such as cognition and social interaction. Our findings provide evidence on the neural level for the speculation [Hoaken and colleagues \(2003\)](#) derived from their behavioural data: more than inhibition as such, it might be rather general problems with the processing of social information that mediate the link between executive functioning and aggression. Both motor response inhibition and naturalistic aggressive reaction to provocation constitute cognitive situations requiring high mental involvement and flexible adaption to rare, respectively unusual and, thus, salient stimuli. Therefore, we interpret our findings as evidence for the involvement of AIC not in specifically inhibitory, but general cognitive self-control across different modalities.

This does not imply that AIC does not equally incorporate functions such as voluntary motor inhibition or the processing of negative emotions. It has to be taken into account that not only an aggressive reaction to provocation but also dysfunctional motor inhibition, and thus failure in a specific task, could be accompanied by negative emotions. AIC has frequently been considered a crucial node of interoception in the cognitive system (reviewed in [Craig, 2009](#)) and as such could merely reflect the processing of anger.

Reactive aggression does not only implicitly comprise a highly impulsive component. Aggression—despite its immediate, possibly functional consequences for the executor—is mostly considered unwanted in society and is highly stigmatized. For this reason, it often imposes negative long-term consequences on the aggressor. Therefore, it could be argued that to avoid these negative consequences, aggressive reactions to provocation should generally be inhibited. In the same line, a motor response should be inhibited in a response inhibition paradigm. In this respect, reactive aggression is closely related to motor impulsivity, and a better grasping of this overlap can eventually lead to a better understanding of clinical syndromes involving aggression and other inhibitory deficits.

### Limitations

The current findings and conclusions have to be considered bearing in mind the limitations of our study. We examined only 15 healthy, young males. Thus, our findings cannot be generalized to bigger populations, to a gender-unspecific context or any clinical population.

This study was designed to research the specific question in how far neural correlates of motor impulsivity and reactive aggression overlap. Nevertheless, it has to be emphasized that the interpretations of our results are of a rather exploratory nature and comprise several alternative explanations. This has to be kept in mind, when evaluating the common neural ground of impulsivity and aggression.

Furthermore, it has to be emphasized that despite their experimental advantages, the chosen behavioural paradigms measure rather specific expressions of the underlying concepts. The GNGT is one among various paradigms used to measure response inhibition. [Eagle and colleagues \(2008\)](#) differentiated several aspects of action inhibition, attributing action restraint to the GNGT as opposed to, for instance, action cancellation measured in stop signal paradigms. This should be kept in mind when interpreting our findings, and only further research including different inhibition paradigms can allow for generalizations.

Despite providing us with a well-established tool to investigate provocative social interactions and retaliation in an imaging environment, it has to be considered that the TAP proves to be an unusually complex paradigm for an fMRI setting. For the presented analyses, we focused merely on the decision phase of the task in which actual behaviour can be observed. Although our behavioural results show that the experimental manipulation was successful and a rather natural pattern of social interaction could be simulated, the limited variety of observed behaviour poses a challenge regarding the choice of maximally specific GLM contrasts. To ensure sufficient power of statistical analyses and maximal convergence towards real-life social behaviour, we restricted our analyses to the contrast of aggressive behaviour in a provocative social interaction and not aggressive behaviour in a social interaction involving no provocation. Due to this circumstance, we cannot fully dissociate pure provocation and pure reactive aggression as, for instance, in the study by [Krämer and colleagues \(2007\)](#), in which punishment could be selected from only four levels instead of eight as in our study. It might be more likely that the entire scale is at least once administered for each opponent if there are fewer options to choose from. All conclusions we draw are based on perceiving retaliation and the provocative situation causing it as a unity. Therefore, the interpretations resulting from the presented findings definitely lack in specificity. It might, though, be questionable, whether neural components exclusively involved in pure inhibition or pure reactive aggression exist and whether identifying them will ultimately enable us to better understand such complex real-life concepts as self-control and retaliation. For future studies, it would be rather interesting to look at participants administering relatively greater punishments than their peers and, at the same time, displaying a weaker ability to inhibit motor responses.



While enabling a more precise description of the specific overlap between the two domains, this would require a significantly larger sample than our study features.

## CONCLUSION AND IMPLICATIONS

In this study, we investigated the neural correlates involved in successful and unsuccessful response inhibition as well as retaliation towards a provoking social counterpart. Furthermore, we focused on the overlap of neural components involved in disinhibition in the context of motor action and provocation. Our findings provide evidence of anterior insula involvement in general self-control across different domains including motor action and social interaction. This might enable a broader perspective on insula function in terms of the awareness model proposed by Craig (2009).

## REFERENCES

- Anderson, C.A., Lindsay, J.J., Bushman, B.J. (1999). Research in the psychological laboratory: truth or triviality? *Current Directions in Psychological Science*, 8(1), 3–9.
- Aron, A.R., Robbins, T.W., Poldrack, R.A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, 8(4), 170–7.
- Asahi, S., Okamoto, Y., Okada, G., Yamawaki, S., Yokota, N. (2004). Negative correlation between right prefrontal activity during response inhibition and impulsiveness: a fMRI study. *European Archives of Psychiatry and Clinical Neuroscience*, 254(4), 245–51.
- Bernstein, S., Richardson, D., Hammock, G. (1987). Convergent and discriminant validity of the Taylor and Buss measures of physical aggression. *Aggressive Behavior*, 13(1), 15–24.
- Brass, M., Haggard, P. (2007). To do or not to do: the neural signature of self-control. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 27(34), 9141–5.
- Carver, C.S., Harmon-Jones, E. (2009). Anger is an approach-related affect: evidence and implications. *Psychological bulletin*, 135(2), 183–204.
- Chambers, C.D., Garavan, H., Bellgrove, M.A. (2009). Insights into the neural basis of response inhibition from cognitive and clinical neuroscience. *Neuroscience and Biobehavioral Reviews*, 33(5), 631–46.
- Chikazoe, J., Konishi, S., Asari, T., Jimura, K., Miyashita, Y. (2007). Activation of right inferior frontal gyrus during response inhibition across response modalities. *Journal of Cognitive Neuroscience*, 19(1), 69–80.
- Corbetta, M., Patel, G., Shulman, G.L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron*, 58(3), 306–24.
- Craig, A.D. (2009). How do you feel – now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, 10, 59–70.
- Damasio, A.R., Grabowski, T.J., Bechara, A., et al. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience*, 3(10), 1049–56.
- Deichmann, R., Gottfried, J.A., Hutton, C., Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage*, 19, 430–41.
- Eagle, D.M., Bari, A., Robbins, T.W. (2008). The neuropsychopharmacology of action inhibition: cross-species translation of the stop-signal and go/no-go tasks. *Psychopharmacology*, 199(3), 439–56.
- Eichele, T., Debener, S., Calhoun, V.D., et al. (2008). Prediction of human errors by maladaptive changes in event-related brain networks. *Proceedings of the National Academy of Sciences of the United States of America*, 105(16), 6173–8.
- Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI), use of a cluster-size threshold. *Magnetic Resonance in Medicine*, 33, 636–47.
- Giancola, P.R., Parrott, D.J. (2008). Further evidence for the validity of the Taylor Aggression Paradigm. *Aggressive Behavior*, 34, 214–29.
- Giancola, P.R., Zeichner, A. (1995). Construct validity of a competitive reaction-time aggression paradigm. *Aggressive Behavior*, 21(3), 199–204.
- Harmon-Jones, E. (2004). Contributions from research on anger and cognitive dissonance to understanding the motivational functions of asymmetrical frontal brain activity. *Biological Psychology*, 67(1–2), 51–76.
- Harmon-Jones, E., Sigelman, J. (2001). State anger and prefrontal brain activity: evidence that insult-related relative left-prefrontal activation is associated with experienced anger and aggression. *Journal of Personality and Social Psychology*, 80(5), 797–803.
- Hennenlotter, A., Schroeder, U., Erhard, P., et al. (2005). A common neural basis for receptive and expressive communication of pleasant facial affect. *Neuroimage*, 26(2), 581–91.
- Hoaken, P.N.S., Shaughnessy, V.K., Pihl, R.O. (2003). Executive cognitive functioning and aggression: is it an issue of impulsivity? *Aggressive Behavior*, 29(1), 15–30.
- Horn, N.R., Dolan, M., Elliott, R., Deakin, J.F.W., Woodruff, P.W.R. (2003). Response inhibition and impulsivity: an fMRI study. *Neuropsychologia*, 41(14), 1959–66.
- Hortensius, R., Schutter, D.J., Harmon-Jones, E. (2012). When anger leads to aggression: induction of relative left frontal cortical activity with transcranial direct current stimulation increases the anger-aggression relationship. *Social Cognitive and Affective Neuroscience*, 7(3), 342–7.
- Jabbi, M., Swart, M., Keysers, C. (2007). Empathy for positive and negative emotions in the gustatory cortex. *Neuroimage*, 34(4), 1744–53.
- Klein, T.A., Endrass, J., Kathmann, N., Neumann, J., von Cramon, D.Y., Ullsperger, M. (2007). Neural correlates of error awareness. *Neuroimage*, 34(4), 1774–81.
- Krämer, U.M., Jansma, H., Tempelmann, C., Münte, T.F. (2007). Tit-for-tat: the neural basis of reactive aggression. *Neuroimage*, 38(1), 203–11.
- Krämer, U.M., Riba, J., Richter, S., Münte, T.F. (2011). An fMRI study on the role of serotonin in reactive aggression. *PLoS one*, 6(11), e27668.
- Logan, G.D., Schachar, R.J., Tannock, R. (1997). Impulsivity and inhibitory control. *Psychological Science*, 8, 60–4.
- Lotze, M., Veit, R., Anders, S., Birbaumer, N. (2007). Evidence for a different role of the ventral and dorsal medial prefrontal cortex for social reactive aggression: an interactive fMRI study. *Neuroimage*, 34(1), 470–8.
- Menon, V., Adelman, N.E., White, C.D., Glover, G.H., Reiss, A.L. (2001). Error-related brain activation during a Go/NoGo response inhibition task. *Human brain mapping*, 12(3), 131–43.
- Menon, V., Uddin, L.Q. (2010). Saliency, switching, attention, and control: a network model of insula function. *Brain Structure and Function*, 214, 655–67.
- Pawliczek, C.M., Derntl, B., Kellermann, T., Kohn, N., Gur, R.C., Habel, U. (2013). Inhibitory control and trait aggression: neural and behavioral insights using the emotional stop signal task. *Neuroimage*, 79, 264–74.
- de Quervain, D.J., Fischbacher, U., Treyer, V., et al. (2004). The neural basis of altruistic punishment. *Science*, 305(5688), 1254–8.
- Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science*, 300(5626), 1755–8.
- Sharp, D.J., Bonnelle, V., De Boissezon, X., et al. (2010). Distinct frontal systems for response inhibition, attentional capture, and error processing. *Proceedings of the National Academy of Sciences of the United States of America*, 107(13), 6106–11.
- Simmonds, D.J., Pekar, J.J., Mostofsky, S.H. (2008). Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia*, 46(1), 224–32.
- Sterzer, P., Kleinschmidt, A. (2010). Anterior insula activations in perceptual paradigms: often observed but barely understood. *Brain Structure and Function*, 214(5–6), 611–22.
- 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(3), 1655–65.
- Talairach, J., Tournoux, P. (1988). *Co-planar Stereotaxic Atlas of the Human Brain: An approach to Medical Cerebral Imaging*. New York: Thieme Medical Publishers.
- Taylor, S.P. (1967). Aggressive behavior and physiological arousal as a function of provocation and the tendency to inhibit aggression. *Journal of personality*, 35(2), 297–310.
- van Honk, J., Schutter, D.J. (2006). From affective valence to motivational direction: the frontal asymmetry of emotion revised. *Psychological Science*, 17(11), 963–5.
- Veit, R., Lotze, M., Sewing, S., Missenhardt, H., Gaber, T., Birbaumer, N. (2010). Aberrant social and cerebral responding in a competitive reaction time paradigm in criminal psychopaths. *Neuroimage*, 49(4), 3365–72.
- Weissman, D.H., Roberts, K.C., Visscher, K.M., Woldorff, M.G. (2006). The neural bases of momentary lapses in attention. *Nature Neuroscience*, 9(7), 971–8.
- White, S.F., Brislin, S.J., Meffert, H., Sinclair, S., Blair, R.J.R. (2013). Callous-unemotional traits modulate the neural response associated with punishing another individual during social exchange: a preliminary investigation. *Journal of Personality Disorders*, 27(1), 99–112.