UNIVERSITÄT REGENSBURG

Inaugural-Dissertation zur Erlangung der Doktorwürde

der Fakultät für Humanwissenschaften der Universität Regensburg

Towards a Biobehavioral Understanding of Trait Externalization:

Neural and Endocrine Responses during Acute Stress and Reactive Aggression within the Non-clinical Range

vorgelegt von

Dipl.-Psych. Julian Konzok

geboren am 07.07.1987 in Memmingen

Regensburg

Abgabedatum: 10.05.2021

Erstgutachterin: Prof. Dr. Brigitte M. Kudielka

Zweitgutachter: Prof. Dr. Boris Schiffer

This dissertation thesis and the presented research were performed at the

Department of Medical Psychology, Psychological Diagnostics and Research Methodology

Institute of Psychology – University of Regensburg

Affiliation of the Supervisors:

Prof. Dr. Brigitte M. Kudielka
Chair of Medical Psychology, Psychological Diagnostics and Research Methodology
Institute of Psychology
University of Regensburg

Prof. Dr. Boris Schiffer
Division of Forensic Psychiatry
LWL-University Hospital
Ruhr University of Bochum

Funded by

the German Research Foundation
(Deutsche Forschungsgemeinschaft: DFG)
KU 1401/6-1

"Psychobiological stress regulation and aggression: Investigations on the dimensionality of externalization"

Acknowledgments

I would like to thank many people. Without them this work would not have been possible. Foremost, I am deeply grateful to Prof. Dr. Brigitte M. Kudielka for the helpful discussions, permanently supporting me and taking always time to share her expertise with me. Also, I am thankful for giving me the chance to visit scientific conferences and gain enriching experiences. Thank you for the great supervision!

Moreover, I would like to thank Prof. Dr. Boris Schiffer and Dr. Christian Kärgel for their experimental support and helpful ideas.

Furthermore, I am very grateful to my colleague M.Sc. Gina-Isabelle Henze for her great support, suggestions and fruitful discussions as well as her honesty. Additionally, I would like to thank Prof. Dr. Stefan Wüst for the support, guidance as well as enriching ideas.

Also, I am thankful to my great colleagues in the laboratory, Dr. Ludwig Kreuzpointner for sharing his knowledge and statistical expertise with me, M.Sc. Hannah Peter, M.Sc. Marina Giglberger, and M.Sc. Tanja-Julia Hiltl for the fantastic cooperation, Heike Hochmuth for solving the crucial problems and having always a sympathetic ear, and all of the current and former research as well as Master students at the lab: Maximilian Artinger, Alexander Barbatsalos, Patricia Bohmann, Vreni Brunnthaler, Thomas Buchmin, Zora-Pia Gerbers, Miriam Heide, Susanne Kargel, Elias Landfried, Sabrina Meier, Christina Neumann, Amelie Propstmeier, Sabrina Seipp, Daniel Schleicher, Katharina Spannruft, Marina Weber, Philipp Vogel, and Florian Zimmer. Thanks for your great work and commitment.

Next, I would particularly like to thank my colleague M.Sc. Christoph Bärtl for the great time at the office and sharing creative ideas with me at the white board. :-)

Last but not least, I am very, very grateful to my wife Susanne Konzok as well as my parents Irmgard and Norbert Konzok for backing me up as well as the mental and emotional support.

General abstract

The externalizing spectrum is characterized by disinhibition, impulsivity, antisocial-aggressive behavior as well as substance (mis)use. Empirical work has shown that the externalizing spectrum is dimensionally distributed characteristic that can be found also in the non-pathological range of variation. Studies in forensic samples and mentally impaired children suggested that higher rates of externalization are linked to lower cortisol stress responses, to abnormalities in neural threat and reward responses as well as to deficits in the inhibitory control system. Furthermore, aggression is a main criterion when diagnosing externalizing disorders such as antisocial personality disorder (ASPD) and conduct disorder (CD). In aggression research, the Taylor Aggression Paradigm (TAP; Taylor, 1967) is widely used to measure reactive aggression in laboratory settings. While modified versions (mTAPs) with various stimulus characteristics (shocks, noise, pressure, heat) have already been established, a modified version with monetary stimuli has only been introduced very recently. Coming back to the externalizing spectrum, mechanisms within the non-clinical range remain still unclear, in particular with respect to stress responses and reactive aggression.

The major aims of this thesis were to investigate whether interindividual differences in non-clinical externalizing behavior was associated with aberration in neural, emotional, and psychoendocrine stress responses. Differences in the extent of externalizing behavior might predict reactive aggression and differential neual activation patterns during a beforehand validated monetary variant of the mTAP.

For the validation of a monetary variant of the mTAP (study 1), 209 young healthy participants (104 men, 105 women) completed a mock Competitive Reaction Time Task (CRTT), so called mTAP, with a fictional opponent with preprogrammed 40 win and 60 lose trials. In lose trials, participants were provoked by subtracting a low (0 – 20 euro cents), medium (30 – 60 cents) or high (70 – 90 cents) amount of money from their fictitious account. To check for sequence effects, provocation stimuli were either presented randomly or in a fixed sequence (experimental conditions). In contrast to a random sequence, the fixed sequence was generated by repeating trials from the same

provocation category in series of three. Study 2 was a functional magnetic resonance imaging study (fMRI). Sixty-one healthy participants (31 men, 30 women) from the higher versus lower range of the non-clinical variation in externalization (31 participants with high externalization), as assessed by the subscales disinhibition and meanness of the Triarchic Psychopathy Measure (TriPM), were exposed to Scan*STRESS*, a standardized psychosocial stress paradigm for the scanner environment. In study 3, participants performed the beforehand validated monetary modified Taylor Aggression Paradigm (study 1) in the scanner environment. During experiment 2 and 3, cortisol and testosterone samples were collected repeatedly and neural correlates were assessed by the blood-oxygen-level-dependent (BOLD) response.

The present findings provide new evidence supporting the view that the monetary mTAP is able to induce and capture reactive aggression in the laboratory. Additionally, no advantage of a fixed sequence was found, as the level of reactive aggression in a given trial appeared to be mainly predicted by the preceding provocation trial. In the high and low externalization group, Scan*STRESS* (study 2) induced a significant rise in cortisol levels with the high externalization group showing significantly lower cortisol stress responses than the low externalization group. *Post-hoc* analysis revealed that this effect was mainly observed in men. Individual increases in cortisol predicted neural response differences between externalization groups, indicating more activation in the dorsal striatum in low externalizing participants. In contrast to the lower cortisol stress responses, this effect was primarily driven by women. In study 3, high externalization predicted reduced anterior cingulate cortex (ACC) activation under high provocation during the monetary mTAP. However, high externalizing participants did not behave more aggressively than the low externalization group. Follow-up analyses revealed that aggression levels and salivary testosterone correlated positively in women and negatively in men. At the neural level, aggression-related activity in the supplementary motor cortex (SMA) was associated with gender-specific testosterone levels.

In conclusion, individuals from the general population with high-level externalization show reduced cortisol stress responses and aberrations in fronto-striatal functioning (ACC, dorsal striatum), thus confirming earlier findings in clinical samples. However, the observed physiological

changes did not lead to higher behavioral aggression levels. Provocation during monetary mTAP elicits testosterone-related aggression rather in women than in men. Moreover, SMA seems to play a relevant role in the neural network mediating this gender-specific testosterone-aggression relationship. Thus, the results support the view that the association between aggression and testosterone depends strongly on moderating factors (e.g., contextual, personal and biological influences).

Table of contents

Index of figures

Index of tables

Index of abbreviations

ACC ..... *Anterior cingulate cortex*
ACTH ..... *Adrenocorticotropic hormone*
ADHD ..... *Attention-deficit/hyperactivity disorder*
ANOVA ..... *Analysis of variance, Analysis of variance*
ASPD ..... *Antisocial personality disorder*
BGHA ..... *Brown-Goodwin Lifetime History of Aggression Scale*
BIS-15 ..... *Barratt Impulsiveness Scale – Short Version*
BOLD ..... *Blood-oxygen-level-dependent*
BPAQ ..... *Buss Perry Aggression Questionnaire*
bpm ..... *Beats per minute*
C1 – C10 ..... *Cortisol samples*
CAR ..... *Cortisol awaking rise*
CBG ..... *Corticosteroid binding globulin*
CD ..... *Conduct disorder*
CRH ..... *Corticotropin-releasing hormone*
CRTT ..... *Competitive Reaction Time Task*
DELFIA ..... *Time-resolved immunoassay with fluorometric detection*
DMN ..... *Default mode network*
DSM-5 ..... *Diagnostic and Statistical Manual of Mental Disorders*
DTI ..... *Diffusion tensor imaging*
EGA ..... *Extreme groups approach*
EPI ..... *Echo-planar imaging*
fMRI ..... *Functional magnetic resonance imaging*
FMTD ..... *Fitness Model of Testosterone Dynamics*
FOV ..... *Field of view*
FWE ..... *Family-wise error rate*
g ..... *Gram*
GAS ..... *General adaption syndrome*
GLM ..... *General linear model*
HDL ..... *High-density lipoprotein*
HPA ..... *Hypothalamic-pituitary-adrenal*
HR ..... *Heart rate*
K-FAF ..... *"Kurzfragebogen zur Erfassung von Aggressivitätsfaktoren"*
l ..... *Liter*
L ..... *Left*
LC-NA ..... *Locus coeruleus-noradrenaline/autonomic nervous system*
LR ..... *Likelihood ratio*
MCC ..... *Middle cingulate cortex*
min ..... *Minute*
mm ..... *Millimeter*
mm^2 ..... *Square millimeter*
mm^3 ..... *Cubic millimeter*
MNI ..... *Montreal Neurological Institute*
mPFC ..... *Medial prefrontal cortex*
MP-RAGE ..... *Magnetization-prepared rapid gradient-echo*
mR^2 ..... *Marginal R2*
ms ..... *Millisecond*
mTAP ..... *Modified Taylor Aggression Paradigm*
N/A ..... *Not found in any probability map*

NIMH ..... *U.S. National Institute of Mental Health*
nmol/l ..... *Nanomol per liter*
OFC ..... *Orbitofrontal cortex*
PANAS ..... *Positive and Negative Affect Schedule*
pg/ml ..... *Picogram per milliliter*
POMC ..... *Proopiomelanocortin*
PSAP ..... *Point Subtraction Aggression Paradigm*
R ..... *Right*
RDoC ..... *Research Domain Criteria*
ROI ..... *Region of interest*
RPQ ..... *Reactive Proactive Questionnaire*
s ..... *Second*
SAM ..... *Sympathetic-adrenal-medulla*
SCID-I ..... *Structured Clinical Interview for DSM-IV Axis I Disorders*
SCID-II ..... *Structured Clinical Interview for DSM-IV Axis II Disorders*
SEM ..... *Standard error of mean*
SMA ..... *Supplementary motor area*
SUD ..... *Substance abuse disorders*
T1 - T3 ..... *Testosterone samples*
TAP ..... *Taylor Aggression Paradigm*
TE ..... *Echo time*
TR ..... *Repetition time*
TriPM ..... *Triarchic Psychopathy Measure*
TSST ..... *Trier Social Stress Test*
WST ..... *"Wortschatztest"*
yrs ..... *Years*

Chapter I:

General introduction and outline of the thesis

1.1 Introduction

> "The crucial point is that psychopathology is inextricably tied in with mechanisms of normal development." (Hinshaw, 2016, p. 94)

According to the statement above, the Research Domain Criteria (RDoC) project of the U.S. National Institute of Mental Health (NIMH) was established to connect normal and atypical development as well as to meet findings from cognitive neuroscience ("cross-domain measurement approach"). In contrast to the categorial Statistical Manual of Mental Disorders (DSM-5), RDoC is a multi-dimensional system considering mental disorders as extreme extents of different traits. This approach is guided by biobehavioral substrates of these trait dimensions (Insel et al., 2010). The constructs of negative valence (e.g., fear, anxiety, stress) and arousal/regulation specified in RDoC are linked to externalizing psychopathologies such as attention-deficit/hyperactivity disorder (ADHD), CD, ASPD and substance abuse disorders (SUD) (Devika, Tuvblad, & Baker, 2016). Especially with regard to externalizing disorders, former findings indicate common etiological mechanisms (Krueger et al., 2002; Markon & Krueger, 2005). Referring to the statement above, little is known about externalization within the non-clinical range. Filling this gap can lead to a better understanding of those processes derailing during the development of externalizing psychopathologies.

From a developmental perspective, it is assumed that environmental adversity, such as parenting stress, parenting strategies (e.g., inconsistency and hostility) or child maltreatment, increases the likelihood of externalizing behavior and elicits changes in responsivity to stress, including hypothalamic-pituitary-adrenal (HPA) axis functioning (VanZomeren-Dohm, Xu, Thibodeau, & Cicchetti, 2016). This allows the hypothesis that the externalizing spectrum is accompanied with altered psychobiological stress responses, to a different extent in both non-clinical and pathological manifestations.

Interpersonal aggression is one of the main diagnostic criteria of externalizing disorders, which is empirically supported by studies using self-report questionnaires (White, Jarrett, & Ollendick,

2013). Aside from aggression questionnaires, laboratory aggression paradigms such as the Taylor Aggression Paradigm (Taylor, 1967) and the Point Subtraction Aggression Paradigm (PSAP; Cherek, 1981) were developed as indirect measures to avoid disadvantages (e.g., social desirability) associated with self-reports (McCloskey & Coccaro, 2003). Studies using these paradigms revealed that patients with CD, ADHD, SUD, and ASPD behaved more aggressively than the control group (Moeller, Dougherty, Lane, Steinberg, & Cherek, 1998; Waschbusch et al., 2002). Empirical evidence on aggression assessed with aggression induction paradigms and non-clinical externalization is rare. Regarding this, there is only one study, which revealed a positive association between laboratory aggression and subclinical externalization under high provocation (Subramani, Parrott, Latzman, & Washburn, 2019).

Therefore, this thesis aimed (1) to develop a validated standard procedure for a variant of the mTAP, also called CRTT, as well as to investigate (2) stress reactivity and (3) aggression processing in participants exhibiting externalizing behavior within the non-clinical range.

The following chapter provides an overview of the externalizing spectrum, stress, aggression, and the interconnections between these topics. In particular, the theoretical framework, the neurobiological background as well as assessment strategies are described. The next chapters focus on a validation study of a modified TAP with monetary stimuli (Chapter II) and two fMRI studies (Chapter III & IV), investigating moderating effects of externalization within a non-clinical range on psychobiological stress response as well as reactive aggression. Finally, chapter V provides a general discussion and integration of the three studies, including limitations and implications for future research.

1.2 Externalization

Several lines of research showed high comorbidity rates among externalizing pathologies as well as significant heritability of a general externalizing factor indicating common etiological mechanisms (Krueger et al., 2002; Young, Stallings, Corley, Krauter, & Hewitt, 2000). To model these observations, different model types of disinhibitory personality and psychopathology were

developed and empirically evaluated, following a transactional perspective (Krueger et al., 2002; Walton, Ormel, & Krueger, 2011). In comparative investigations, dimensional models (e.g., latent trait model) showed the best fit as opposed to categorical (e.g., latent class model) models (Krueger, Markon, Patrick, & Iacono, 2005; Markon & Krueger, 2005). It is assumed that the externalizing spectrum is continuously distributed throughout the different samples (e.g., general population, forensic samples). Krueger, Markon, Patrick, Benning, and Kramer (2007) empirically developed the externalizing spectrum model by using item response theory and semiparametric factor analysis. Over three waves of iterative data collection, in total 1,787 participants fulfilled a pool of items associated with various constructs of externalizing pathologies to clarify the structure of the externalizing spectrum. The resulting inventory is named the Externalizing Spectrum Inventory (ESI; Krueger, Markon, Patrick, Benning, & Kramer, 2007) comprising 415 items and 23 subfactors. The bifactor structure of the ESI consists of a general factor ("externalizing" or "disinhibition") and two specific factors ("substance use" and "callous-aggression").

Based on the ESI, the TriPM (Patrick, 2010) was developed with the subscales disinhibition, meanness, and boldness for economical assessment of externalization and psychopathy. The subscale disinhibition exclusively loads on the general externalizing factor of the above mentioned ESI, whereas items of the meanness scale exhibits appreciable loadings on the callous-aggression factor (Venables & Patrick, 2012). Additionally, high scores on boldness, representing fearlessness in interpersonal, affective-experiential and behavioral domains (Venables & Patrick, 2012), matches the closely-related construct of psychopathy.

Considering connections to other models structuring personality and pathology, the DSM-5 provides a dimensional diagnostic approach including externalizing behavior problems flanked by internalizing pathologies (e.g., depression, anxiety). Additionally, externalization correlates negatively with the dimensions agreeableness and conscientiousness as well as positively with extraversion of the five-factor personality model (DeYoung, Peterson, Séguin, & Tremblay, 2008; John, Caspi, Robins, Moffitt, & Stouthamer-Loeber, 1994).

1.3 Stress

1.3.1 Definition & theoretical background

One of the first stress researchers, Selye (1936), defined stress as unspecific reaction of the body to environmental strains involving physiological (e.g. heartbeat, respiration, blood pressure) and endocrinologic (e.g., glucocorticoids) changes. Selye's experiments with mice revealed that the physiological stress reaction is associated with a swelling of the adrenal cortex, an atrophy of the thymus as well as gastric and duodenal ulcers. Based on this, Selye developed the concept of the general adaption syndrome (GAS). Within this concept, it is hypothesized that the stress reaction starts with an alarm response (e.g., sympathetic nervous system), followed by a stage of resistance, during which the body can compensate the stressor, however remains alert. After a prolonged phase of stress, the body is set in a stage of exhaustion also characterized by an enhanced susceptibility to diseases (Selye, 1946).

Subsequently, Mason (1968) discovered that emotional and appraisal processes affect the endocrine stress response. It could be shown that endocrine reactions can be triggered by situations, which are interpreted as novel, uncontrollable, unpredictable or connected with negative consequences.

Over the decades of the last century, the term stress was described in various ways resulting in a lack of a consistent definition as basis for experimental designs and conclusions. In 1991, Levine and Ursin aimed at delivering a widespread definition for the term stress. The authors distinguished between stress stimuli (input), individual processing as well as the resulting stress reaction. Referring to Mason's findings, within this concept (Levine & Ursin, 1991), appraisal processes (e.g., threatening stimuli) play a key role in experiencing a stimulus as stressor. In turn, these processes depend strongly on predisposition (e.g., genetics) and environmental influences (e.g., former experiences).

In the context of stress and cognitive appraisal, the transactional model of stress developed by Lazarus and Folkman (1984) has to be reported. The transactional model proposes that stress is a

product of three appraisal processes, the interpretation of the situation (e.g., relevant, dangerous), evaluation of the abilities and resources to cope with the stressor as well as the reappraisal (e.g., whether coping efforts were successful or not) (Folkman & Lazarus, 1984).

After this introduction to stress research, the following section will focus the biological stress response.

1.3.2 Biological stress responses

In the biological stress response, the secretion of catecholamines and glucocorticoids is regulated by two major hormonal signaling pathways, the locus coeruleus-noradrenaline/autonomic nervous system (LC-NA), also referred to as sympathetic-adrenal-medulla (SAM) axis, and the HPA axis (see Figure 1; Baritaki, de Bree, Chatzaki, & Pothoulakis, 2019, p. 3) (Kudielka & Kirschbaum, 2007).

1.3.2.1 Locus coeruleus-noradrenaline/autonomic nervous system

The highest hierarchical level of the LC-NA system consists of the locus coeruleus and the hypothalamus. When activated, the neurotransmitter noradrenaline is released by a dense network of neurons into the brain (Chrousos & Gold, 1992), which is also controlled by activation of the amygdala, the hippocampus as well as the mesocortical and mesolimbic dopamine systems (Chrousos, 1998). These processes trigger physiological accommodations (e.g., enhanced heart rate and blood pressure) and thereby increase arousal, vigilance, alertness, and attention for preparing adaptive behavior (De Kloet, Joëls, & Holsboer, 2005; Kudielka & Kirschbaum, 2005). These accommodations are mediated by the sympathetic division of the autonomic system, via its effectors, the sympathetic nerves and the adrenal medulla (Chrousos & Gold, 1992). The sympathetic nerves send the signal over the preganglionic and postganglionic neurons to the target peripheral organs (e.g., heart, skeletal muscles). To activate receptors on the postganglionic neurons the preganglionic neurons release the neurotransmitter acetylcholine, in contrast, the endings of the postganglionic neurons discharge noradrenaline, which stimulates adrenergic receptors on the peripheral target tissues. However, catecholamine (mainly noradrenaline) released by the

sympathetic nerve endings acts mainly locally and barely reaches the blood stream. This fact should be kept in mind when interpreting plasma noradrenalin levels in response to stress (Kudielka & Kirschbaum, 2007; Stalder & Kirschbaum, 2013).

Additionally, the preganglionic fibers stimulate chromaffin mark cells with cholinergic synapses. These cells reside in the adrenal medulla and can be seen as modified postganglionic neurons. Under resting condition, the adrenal medulla secretes a small proportion of catecholamines (80% adrenaline and 20% noradrenaline). During stress, this level rises up to 35% of the total circulating catecholamines into the blood stream influencing, for instance, heart rate, and basal metabolic rate (for review see Kudielka & Kirschbaum, 2007).

Catecholamine secretion can be stimulated by extreme heat or cold, pain, physical effort as well as psychological stress (e.g., exams, free speech). In contrast to the HPA axis, the SAM system shows no comparable habitation effect after repeated exposure to psychosocial stress (Schommer, Hellhammer, & Kirschbaum, 2003).

In short, stress-mediated catecholamine release by the SAM system leads to a rapid mobilization of stored energy depots (e.g., via glucogenolysis, lipolysis) and a downregulation of less relevant organ functions (e.g., the gastrointestinal tract) as well as an increase in heart rate, cardiac output, and blood pressure. Furthermore, SAM activation enhances respiratory extraction of oxygen, platelet aggregation and reduces time to blood clotting (for review see Kudielka & Kirschbaum, 2007).

1.3.2.2 Hypothalamus-pituitary-adrenal axis

The HPA axis is the second major stress system comprising the hypothalamus, the pituitary gland, and the adrenal cortex. This system regulates the stress response and is vital for maintaining many other physiological processes (e.g., immune system) (for review see Kudielka & Kirschbaum, 2007).

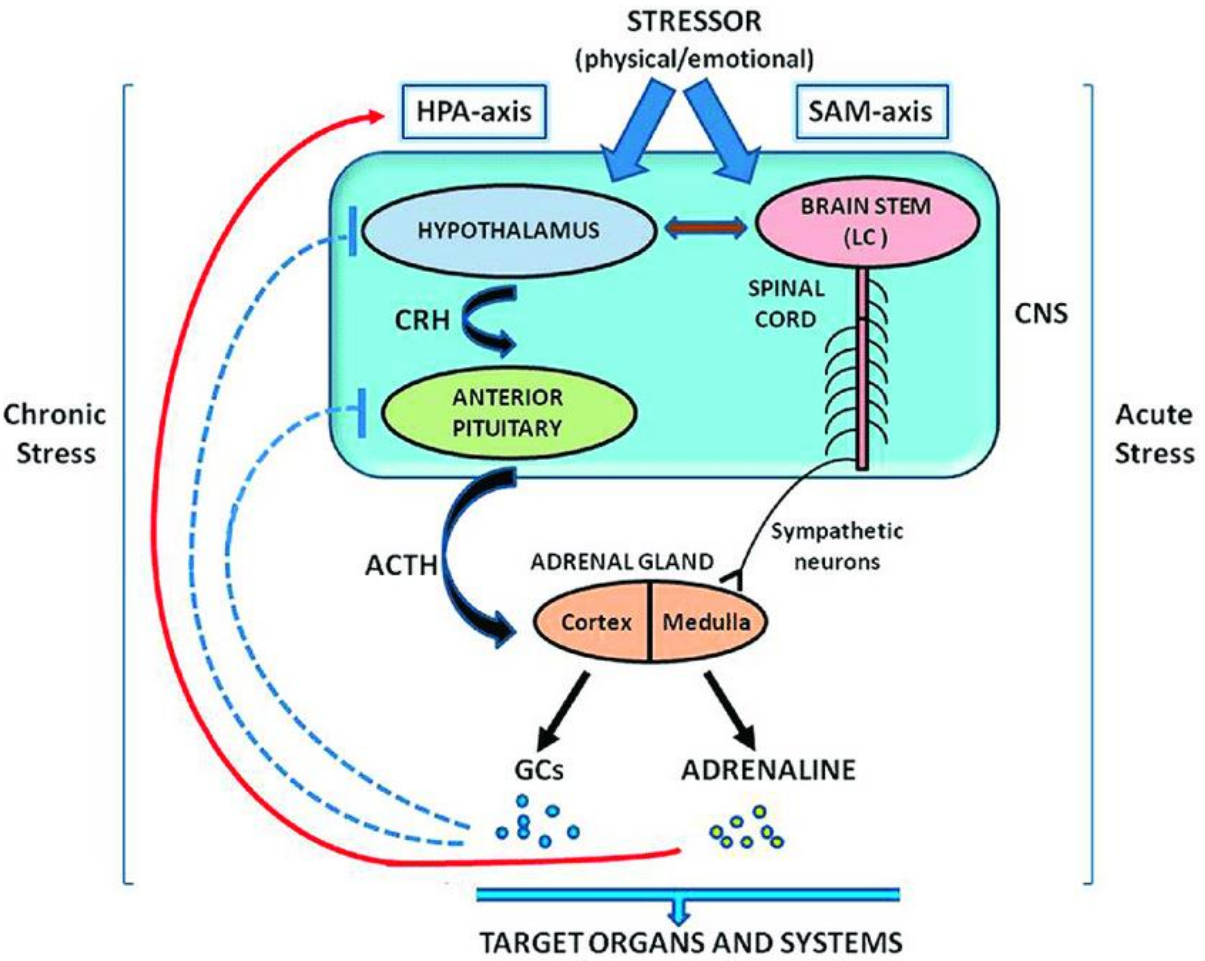

Figure 1. Overview of HPA and SAM axes regulating stress response. Notes: ACTH = adrenocorticotropic hormone, CNS = central nervous system, CRH = corticotropin releasing hormone, GCs = glucocorticoids, HPA axis = hypothalamic-pituitary-adrenal axis, LC = locus coeruleus, SAM axis = sympatho-adrenal-medullary axis. Reprinted from "Chronic stress, inflammation, and colon cancer: a CRH system-driven molecular crosstalk," by S. Baritaki, E. de Bree, E Chatzaki, & C. Pothoulakis, 2019, *Journal of Clinical Medicine*, 8, 1669, p. 3.

When the organism is confronted with internal and external challenges (e.g., psychosocial stress), the hypothalamus (paraventricular nucleus) releases vasopressin and corticotropin-releasing hormone (CRH), which, in the pituitary gland, leads to a cleavage of proopiomelanocortin (POMC) into adrenocorticotropic hormone (ACTH), beta-endorphin, and other peptides. Release of CRH can be facilitated by neural pathways such as the limbic system which is responsive to psychosocial stimuli. Release of ACTH, however, can also be modulated by vasopressin, oxytocin, adrenaline, and noradrenaline. Arriving at adrenal cortex via the blood stream, ACTH stimulates the secretion of glucocorticoids (e.g., cortisol), regulated by several negative feedback loops and mineralocorticoid

as well as glucocorticoid receptors. A large proportion (90 – 95%) of the total plasma cortisol is bound to transport proteins, mostly to corticosteroid-binding globulin (CBG). Only a small portion is biologically active, available to tissues or "free" (10%) (De Kloet et al., 2005; Kudielka & Kirschbaum, 2005). At the target tissue, cortisol diffuses through the cell membrane and binds to intracellular receptors in the cytoplasm. The cortisol-receptor complex translocates into the nucleus and stimulates gene expression (Stalder & Kirschbaum, 2013).

Glucocorticoids drain resource as well as energy from less relevant organ functions (e.g., growth and reproductive systems) and enhance circulating levels of energy substrates (e.g., glucose, free amino as well as free fatty acids) to cope with the demands of the challenging situation (Chrousos & Gold, 1992). Furthermore, cortisol secretion affects the cardiovascular system (e.g., sensitivity for catecholamines) and the immune system (e.g., anti-inflammatory effects, inhibition of pro-inflammatory cytokine synthesis) and controls fluid volume and the response to hemorrhage (Kudielka & Kirschbaum, 2007). Moreover, affective or cognitive processes (e.g., learning, memory) are influenced by cortisol (Buchanan, al'Absi, & Lovallo, 1999; Wolf, 2009).

The cortisol stress response is affected by a variety of factors such as age, gender, substance intake, early life experiences as well as personality (for review see Kudielka, Hellhammer, & Wüst, 2009; Zänkert, Kudielka, & Wüst, 2020). Also, cortisol is secreted along a circadian rhythm. Cortisol levels reach a maximum shortly after awaking (cortisol awaking rise, CAR) and decrease thereafter over the day (Clow, Hucklebridge, & Thorn, 2010; Stalder et al., 2016).

Physical activity, psychological stress, and administration of pharmacological substances can stimulate HPA axis function (Fink, 2007). One example of a psychosocial stressor is the Trier Social Stress Test (TSST; Kirschbaum, Pirke, & Hellhammer, 1993) and its adaption for MRI environments, the Scan*STRESS* paradigm (Streit et al., 2014), which is described in more detail in section 1.3.4.

Various psychosomatic and psychiatric disorders are linked to a dysfunction of the HPA axis. For instance, studies revealed that major depression (Swaab, Bao, & Lucassen, 2005) and cardiovascular problems (McEwen, 1998a; McEwen, 1998b) are associated with enhanced cortisol responses to

acute stress, whereas hyporeactivity of the HPA axis more likely occurs in patients with multiple sclerosis (Adams & Victor, 1989), fatigue syndrome (Cleare, 2003) or post-traumatic stress (Yehuda, 2002). Interestingly, externalizing disorders are also related to alterations in HPA axis function (e.g., Fairchild et al., 2008). For more detailed information, see section 1.3.5.

1.3.3 The concept of allostatic load

The physiological processes described above can be viewed as reactions to adapt to internal or external challenges, maintain homeostasis and protect the body. Accordingly, Sterling and Eyer (1988) introduced the term allostasis as regulation of the internal milieu through dynamic changes of hormonal and physical parameters. Thus, allostasis can be regarded as "stability through change". However, early studies (e.g., from Selye, 1936) already documented that, in the long run, stress and stress-related changes can cause diseases. To describe this trajectory of short-term and long-term effects of stress, McEwen and Stellar (1993) introduced the model of "allostatic load" as a synonym for the biological costs associated with the adaption to stress. Hence, if an allostatic reaction persists over a longer period of time due to chronic or repeated challenges, the organism reaches a state of chronic or repeated overactivity of the physiological systems (e.g., endocrine, neural responses) and develops diseases (McEwen, 1998a). According to the model of allostatic load, the allostasis-adaptation process has three stages. First, a stress response stimulates primary mediators (e.g., cortisol, catecholamines, dehydroepiandrosterone-sulfate, Interleukin-6) which induce primary effects (e.g., anxiety, sleeping problems, mood changes, etc.). In the second stage, prolonged stress results in changes of six secondary outcome parameters, namely blood pressure, waist-hip ratio, high-density lipoprotein (HDL) level, cholesterol level, cholesterol/HDL ratio, and glycosylated hemoglobin level. These changes, in turn, lead to tertiary outcomes representing actual diseases (e.g., diabetes, arterial hypertension, cancer, and cardiovascular disease) (Kudielka, Hellhammer, & Wüst, 2009).

Following the model of McEwen (McEwen, 1998a; McEwen, 1998b), four scenarios are proposed to lead to allostatic load, three of which comprise frequent exposure to stress, inability to

habituate to repeated challenges, and inability to terminate a stress response. In the fourth scenario, an inadequate stress response of one allostatic system (e.g., HPA axis) results in destabilization of interrelated allostatic systems. To explain individual differences in the development of stress-related diseases, the model includes factors such as early life experiences, genetic predispositions, environmental influences, as well as psychological and behavioral variables (Kudielka, Hellhammer, & Wüst, 2009). Additionally, the concept of allostatic load plays a role in biobehaviorial etiology of externalizing disorders (see section 1.3.5).

1.3.4 Standardized psychosocial stress paradigms

One example for a widely used and well validated laboratory paradigm to induce psychosocial stress is the Trier Social Stress Test (TSST). A meta-analysis from Dickerson and Kemeny (2004) including data from 208 laboratory studies applying acute psychological stress paradigms revealed that tasks with uncontrollable and social-evaluative elements evoked enhanced cortisol responses. In addition, a context of forced failure (e.g., unavoidable negative consequences, unsolvable tasks) contributed to increased cortisol levels. These findings provided evidence for the social self-preservation theory postulating that uncontrollable threats to the goal of maintaining the social self are required for enhanced cortisol responses. The authors concluded that, in particular, the TSST with its components meets these empirical and theoretical criteria.

In the standardized procedure of the TSST, participants are instructed to perform a mock job interview with the company's staff managers. The TSST comprises three phases: a brief preparation period (5 min), a free speech (5 min) and a mental arithmetic task (5 min). After a resting period of at least 30 to 45 min, participants are introduced to the committee (dressed in white coats), told to be trained in monitoring non-verbal behavior. Additionally, it is pointed out that the interview will be recorded by a video camera. The participants receive standardized instructions from the committee before each task. The communicative behavior of the committee is designed by an unresponsive neutral manner (no facial or verbal feedback). After the preparation period for the free speech, participants have to report on their personal characteristics that qualifies them for the job

on offer. When participants finish their speech before the end of phase 2 or start to report on other topics than their personal strength, the committee responds in a standardized way and replies "You still have some time left. Please continue!" or "Please come back to your personal strength". In case of a second interruption of the speech before the end of phase 2, the committee waits for 20 sec and subsequently asks prepared questions ("What are your personal strengths?", "What are your major shortcomings?", "What do you think about team work?", "What do your boss/family/colleagues think about you?" etc.).

In the third phase, participants are instructed to serially subtract the number 17 from 2023 as quickly as possible. In case of a mistake, participants are stopped by the committee and requested to restart again with 2023. After 5 min, one member of the committee announces the end of the task and directs the participant to leave the test room (Kirschbaum et al., 1993; Kudielka, Hellhammer, & Kirschbaum, 2007). It is recommended that samples (e.g., saliva, blood) are collected repeatedly over the test session to assess pre-stress levels, initial stress responses, peak levels and recovery (Kudielka et al., 2007). In the past decades, research has shown that the TSST provokes cortisol responses with a 2- to 4-fold elevation above the pre-stress level and exhibits responder rates > 70% (Kirschbaum et al., 1993; Schommer et al., 2003).

An adaption of the Trier Social Stress Test (TSST) for MRI scanner environments is provided with the Scan*STRESS* paradigm, implementing the two essential situational components: uncontrollability and social-evaluative threat (Dickerson & Kemeny, 2004). Participants have to fulfill a mental arithmetic and spatial mental rotation task under two different conditions. While performing the tasks under the stress condition (performance trials), participants are confronted with an observation panel giving feedback regarding their working speed ("work faster") and accuracy ("error") via live stream. During the control condition, the live stream of the committee is interrupted as indicated by a black cross (see Figure 2) and participants receive no feedback (Henze et al., 2020; Streit et al., 2014).

The paradigm consists of two runs lasting 11:20 min each. In each run, trials are presented alternating between the stress and control condition. Each includes a block of mental arithmetic tasks for 60 s and spatial mental rotation tasks for 60 s. Between the task-specific blocks, a break (22 s) and an instruction for the following phase (4 s) are displayed. The first run starts with the control condition, the second run with the stress condition. The order of the blocks and runs is identical for every participant. During the break between the two runs, the MRI scanner is paused, and subjects receive additional feedback that the committee is unsatisfied with their performance and that they need to increase their efforts.

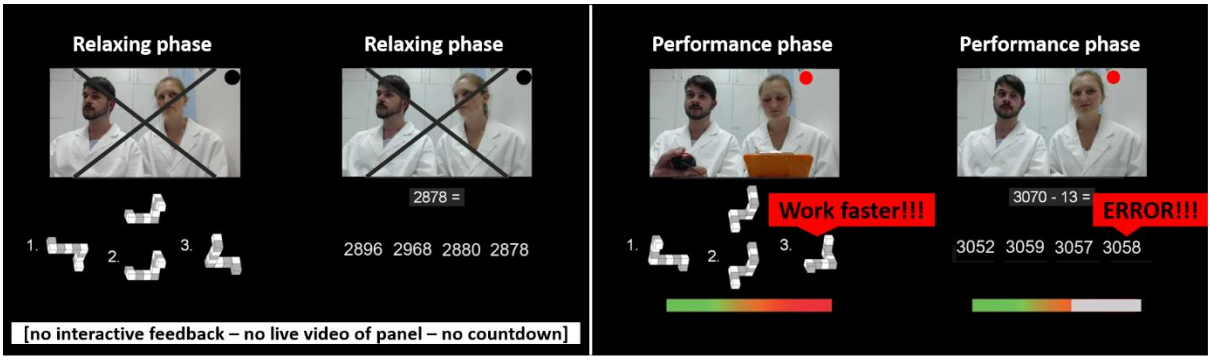

Figure 2. Example of a mental arithmetic and spatial mental rotation task during the control and stress condition of the Scan*STRESS* paradigm.

According to the TSST, samples (e.g., saliva, blood) are collected to assess pre-stress levels, initial stress responses, peak levels, and recovery (Konzok et al., under review; Henze et al., 2020). Responder rates calculated for five studies ranged from 52% to 67% indicate that the Scan*STRESS* is a reliable and valid stress induction paradigm (Dahm et al., 2017; Henze et al., 2020; Konzok et al., under review; Lederbogen et al., 2011; Streit et al., 2014).

1.3.5 Neurobiology of stress and externalization

Animal studies in mammals and rodents revealed that the HPA axis is regulated by networks ranging from brainstem nuclei to limbic system structures (e.g., hippocampus) (Kudielka & Kirschbaum, 2007). These networks lead on to the paraventricular nucleus of the hypothalamus (Herman et al.,

2003) from which CRH is released initiating HPA signaling. Activation in the hippocampus and anterior cingulate cortex (ACC) appear to inhibit glucocorticoid secretion, whereas amygdala activity facilitates cortisol stress responses (for review see Herman, Ostrander, Mueller, & Figueiredo, 2005). Herman et al. (2003) suggest that reactive stressors (e.g., sensory stimulation such as cold) elicit activity in the brainstem and the hypothalamus, while limbic regions (e.g., hippocampus, amygdala, medial prefrontal areas) are associated with the neural response to anticipatory stressors such as social challenges or unfamiliar situations.

Neuroimaging studies could confirm a contribution of prefrontal areas (Dahm et al., 2017; Pruessner et al., 2008), the amygdala (Akdeniz et al., 2014; Pruessner et al., 2008; Streit et al., 2014), hypothalamus (Lederbogen et al., 2011; Pruessner et al., 2008), and the hippocampus (Dahm et al., 2017; Pruessner et al., 2008; Streit et al., 2014) to human neural stress responses during psychological stress. However, both the specific pattern of activation and deactivation and the direction of cortisol-related associations appear to depend strongly on the type of paradigm that is applied (for review see Noack, Nolte, Nieratschker, Habel, & Derntl, 2019).

Concerning the neuroendocrine stress responses in externalization, previous studies regarding the association between externalization and the regulation of the HPA axis as primary biomarker for stress presented conflicting results. This may be due to the heterogeneity of the samples (e.g., psychiatric and forensic patients) with a wide age spectrum (e.g., ranging from children to adolescents and adults) as well as methodological differences between studies. For example, some stress induction paradigms appeared to fail to reliably induce HPA axis stress responses (Alink et al., 2008; Kobak, Zajac, & Levine, 2009). Nevertheless, the majority of studies showed that children and adolescents with conduct disorder and disruptive behavior disorder (Fairchild et al., 2008; McBurnett et al., 2005; Popma et al., 2006; Snoek, Van Goozen, Matthys, Buitelaar, & van Engeland, 2004; Van Goozen et al., 2000), violent adult offenders (Virkkunen, 1985) and substance misusers (Couture et al., 2008) displayed reduced HPA axis responses. To explain this neuroendocrine dysfunction one prime candidate is the concept of allostatic load (see section 1.3.3). Studies show that patients with

externalizing disorders are more frequently exposed to early life stress (e.g., child maltreatment, maladjustment to environmental demands) (Deater-Deckard, Dodge, Bates, & Pettit, 1998) and show heightened emotional reactivity (Calkins, Gill, Johnson, & Smith, 1999). These higher levels of stress lead to allostatic load and, consequently, to chronic downregulation of HPA axis signaling (Alink, Cicchetti, Kim, & Rogosch, 2012; Alink et al., 2008). The question, which of the above-mentioned stress-responding neural networks are responsible for HPA axis dysfunction in externalization, can not be answered plausibly, since current knowledge, especially with regard to stress paradigms, is scarce. Studies with other paradigms (e.g., reward processing) revealed that patient with externalizing disorders (e.g., CD) increasingly respond to acute threat, which is most notable in regions such as amygdala and prefrontal areas (Coccaro, McCloskey, Fitzgerald, & Phan, 2007). In turn, acute threat is strongly connected with the cortisol stress response and, on a psychological level, with the subjective experiences of stress (for review see Brosschot, 2010). Additionally, results from other domains such as deficits in emotion processing and regulation in externalization suggest also neurobiological abnormalities in response to psychological stress (Fairchild et al., 2019; Lee, Kavoussi, & Coccaro, 2008).

Given the idea of biobehavioral substrates of continuously distributed traits such as externalization (e.g., RDoc), some alterations in neural stress response should be observable in participants exhibiting high externalizing behavior within the non-clinical range. However, there is still a lack of data on psychobiological stress responses in healthy adults exhibiting externalizing behaviors within a subclinical range.

1.4 Aggression

Aggression and its extreme forms are an omnipresent topic in societies and politics. Aggressive behavior emerges in many forms and in various environments (e.g., bullying in schools or at work, violence in families). The World Health Organization estimates that 1.43 million people per year die from either self-inflicted or interpersonal violence, all over the world (Word Health Organization,

2007). These data emphasize the relevance of this topic and the need for extensive research activities in this field.

Due to the complexity of affective, cognitive and physiological processes related to aggressive behavior, to formulate a comprehensive definition of aggression is challenging. Nevertheless, most researchers agree that two components are mandatory to classify a behavior as aggressive. While the actor has to deliver an aversive stimulus and intend to harm the victim, the victim needs to be motivated to avoid the attack (Baron & Richardson, 1994; Geen, 2001). Based on this definition and the extensive body of research identifying aggression subtypes (Crick & Dodge, 1996; Geen, 2001), two different types of aggression can be distinguished, namely reactive versus proactive aggression. Reactive aggression, the affective form of aggression, occurs in response to provocation or frustration while proactive aggression, the instrumental type of aggression, is driven by obtaining desired goals (e.g., goods, coercive power, etc.) (Crick & Dodge, 1996; Geen, 2001).

1.4.1 Theoretical background

Several lines of research attempted to answer the questions why some people become aggressive (and others not) or why some humans show aggressive behavior sooner than others do. One of the pioneers in this field was a group of psychologists at the Yale University that published the frustration-aggression hypothesis (Dollard, Miller, Doob, Mowrer, & Sears, 1939). Dollard and colleagues (1939) stated that frustration induced by blocking goal-directed behavior can lead to diverse reactions, however, predominantly to aggression. Nevertheless, it turned out that this theory could not explain several other manifestations of aggression (e.g., aggressive behavior without antecedent frustration). Bandura and Walters (1977) emphasized the influence of external factors eliciting aggression and raised the question to what extent human behavior is a result of social learning processes. Establishing the social learning theory, Bandura could show that watching violence enhanced aggressive behavior in three to five years old children within the "Bobo Doll" experiment. However, the learning "success" depended on the consequences of the aggressive

behavior for the model, similarities between model and observer as well as the level of development of the observer.

Subsequently, cognitive models emerged explaining effects of attitudes, perceptual bias, response bias or scripts on (aggressive) behavior (e.g., Crick & Dodge, 1994; Eron, Gentry, & Schlegel, 1994). For example, Huesmann (1998) postulated that (aggressive) behavior is guided by mental programs (scripts) for social situations, which were learned during childhood and adolescence. Features of a given situation (e.g., words, objects) elicit these scripts guiding a person's behavior, expectations and intentions. It was postulated that multiple rehearsals of a script leads to enhanced accessibility by creating new connections to other concepts in memory (Anderson & Bushman, 2002). An enhanced accessibility of aggressive scripts increases the likelihood that the person behaves more aggressively.

Integrating existing aggression theories (e.g., the above-mentioned models), the general model of aggression (GAM; Anderson & Bushman, 2002) postulates that aggressive behavior is influenced by personal and situational variables ("inputs") through their impacts on the current cognitive, affective and arousal state ("routes"). These routes, in turn, affect the resulting behavioral actions ("outcomes") (see Figure 3; Anderson & Bushman, 2002, p. 34).

Situational factors comprise important features of the specific situation. For instance, aggressive cues (e.g., picture of a gun) prime aggression-related concepts in memory (Anderson & Bushman, 2002) and enhance aggressive behavior, on this way (Berkowitz & LePage, 1967). Additionally, studies with laboratory paradigms such as the TAP or PSAP could show that interpersonal provocation promotes aggressive behavior (Cherek, 1981; Taylor, 1967). Furthermore, pain and discomfort as situational factors (e.g., hot temperatures, loud noises, unpleasant odors) can influence aggression (Berkowitz, 1993).

Moreover, a further component of the GAM consists of personal factors including traits, attitudes, beliefs, gender or genetic predispositions affecting the individual's preparedness to aggress (Anderson & Bushman, 2002). For example, self-esteem or narcissism were identified as

aggression-related personality traits (Baumeister, Smart, & Boden, 1996; Donnellan, Trzesniewski, Robins, Moffitt, & Caspi, 2005). In laboratory paradigms, studies yielded a significant correlation between self-reported trait reactive aggression and behavioral reactive aggression (Beyer, Münte, & Krämer, 2014; Chester & Lasko, 2018; Konzok et al., 2020). Furthermore, the connection between gender and aggression seems to depend strongly on environmental factors. For instance, a meta-analysis from Bettencourt and Miller (1996) demonstrated a significant moderating role of the level of provocation in the gender-aggression relationship. Other studies also revealed qualitative differences between women and men regarding the type of aggression. While men rather tend to physical aggression, women are prone to more indirect forms (e.g., social exclusion, gossip) (Björkqvist, Österman, & Lagerspetz, 1994; Taylor & Epstein, 1967). In terms of genetic predispositions, for instance, a polymorphism in the promoter region of monoamine oxidase-A modulates the neural networks mediating the aggressive response (Chester et al., 2015).

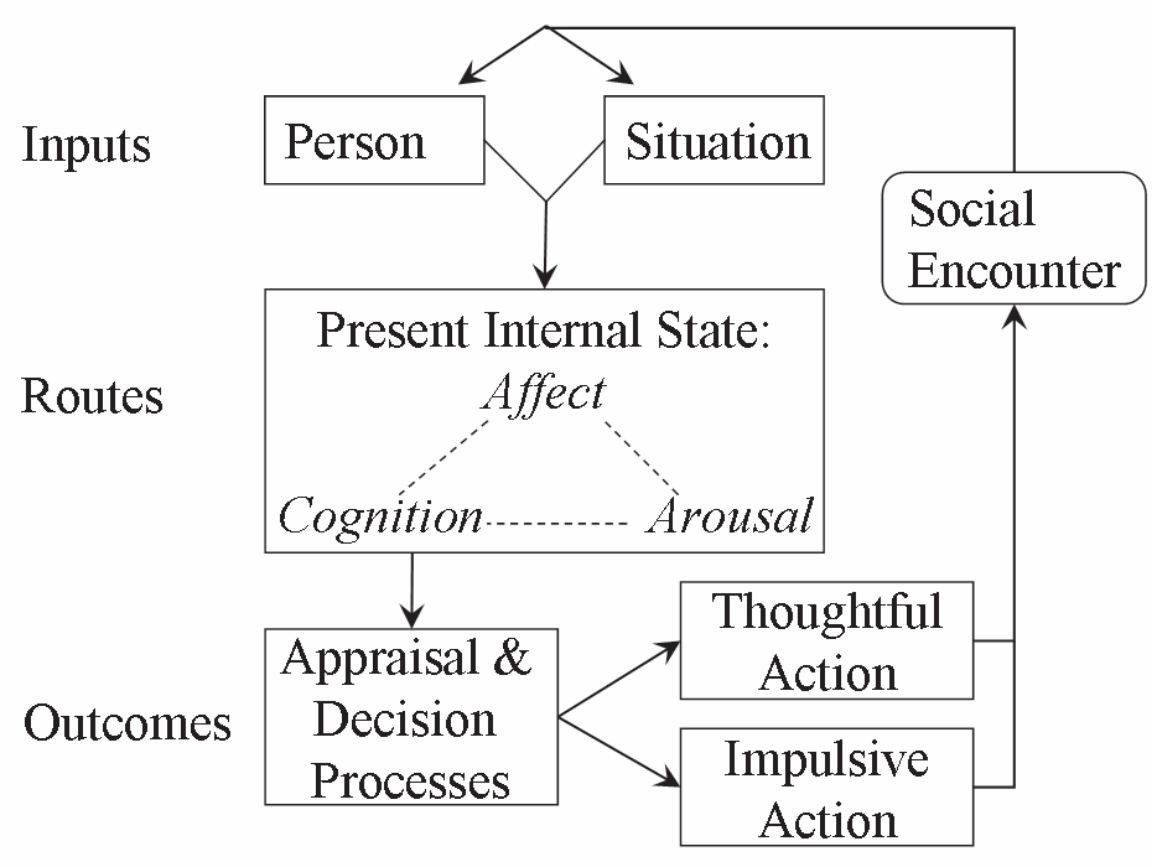

Figure 3. The main components of the General Model of Aggression. Reprinted from "Human Aggression," by C. A. Anderson & B. J. Bushman, 2002, *Annual Review of Psychology,* 53, p. 34. Copyright 2002 by Annual Reviews.

Taken together, the GAM postulates that input variables (situation and person factors) exert influence on aggressive behavior ("outcome") through the present internal state ("routes"). These "routes" comprise cognitive, affective and arousal components. Studies demonstrated that aggressive personality traits are associated with an increased susceptibility to hostile attribution, perception, and expectation biases (Wilkowski & Robinson, 2008) and aggressive affect (Anderson, 1997). Additionally, it is known that there is an effect of an arousing experience (e.g., sports) on subsequent emotional reactions (e.g., aggressive behavior) (Bunce, Larsen, & Cruz, 1993). Finally, the internal present state affects appraisal and decision processes resulting in thoughtful or impulsive behavioral actions.

1.4.2 Measuring aggression – methodological challenges

Traditionally, aggressive behavior is assessed by self-report measures as, for example, the Buss Perry Aggression Questionnaire (BPAQ; Buss & Perry, 1992) or the German "Kurzfragebogen zur Erfassung von Aggressivitätsfaktoren" (K-FAF; Heubrock & Petermann, 2008). These psychometric procedures are associated with potential methodological problems such as response tendencies, poor motivational conditions or insufficient linguistic competence. Therefore, behavioral laboratory aggression paradigms were developed to observe aggressive behavior directly, under strict experimental control conditions (McCloskey & Coccaro, 2003). One of these paradigms comprises the TAP which is widely used to measure aggressive behavior in response to interpersonal provocation stimuli (Elson, Mohseni, Breuer, Scharkow, & Quandt, 2014; Taylor, 1967). Subjects are told that they play a competitive reaction time game against a mock opponent. The winner of each round of this preprogrammed task sets the intensity of a punishment (e.g., electric shock, unpleasant noise) for the opponent. The participants' choice of punishment in response to the preceding provocation by the mock opponent serves as behavioral measure for reactive aggression. Beside other laboratory paradigms, the TAP, also called CRTT, was used to clarify the open questions in terms of aggression, in the last decades. Central advantages of laboratory aggression paradigms over non-experimental approaches are the increased experimental control conditions as well as the potential to examine causal relations between variables.

However, these paradigms have been criticized in a number of ways. For instance, since the development of the original version, researchers have implemented several modifications, resulting in a lack of standardization. At qualitative level, mTAPs with noise stimuli, pneumatic pressure and heat have emerged (Beyer, Münte, Göttlich, & Krämer, 2015; Krämer, Riba, Richter, & Münte, 2011; Lotze, Veit, Anders, & Birbaumer, 2007). Additionally, there are many ways extracting various aggression parameters over one or more trials (e.g., average intensity selection of 25 trials vs. 50 trials). Besides, participants were instructed differently in previous studies (e.g., two opponents with different interactional strategies vs. one mock opponent). Another example of the lack of

standardization in using the mTAP is the order of the presented punishment levels of the mock opponent (provocation). In some studies, participants were confronted with a random order of provocation levels, other procedures consisted of different blocks of the same provocation level positioned in an ascending order. Thus, it can be speculated that the order of provocation levels influences reactivity of the participants (for review see Elson et al., 2014), which has not been empirically tested, yet. Analyzing these problems by comparing different publications with the mTAP, Elson et al. (2014) provided some promising recommendations related to the mTAP with noise stimuli (e.g., inclusion of a non-aggressive option, calibration and scaling of the noise stimuli, preregistration of the analysis plan). Furthermore, the authors suggested to agree on a standard version of the mTAP (e.g., procedure used by Ferguson, Smith, Miller-Stratton, Fritz, and Heinrich, 2008) to minimize the above-mentioned problems.

Recently, Schneider et al. (2015) provided a variant of the mTAP with monetary stimuli (0 – 100 euro cents) with 150 trials. Subsequent studies used this version investigating borderline patients (Kogan-Goloborodko, Brügmann, Repple, Habel, & Clemens, 2016) and healthy controls (Repple et al., 2018; Repple et al., 2017). These authors analyzed outcome measures in the same way by aggregating aggression levels in response to the different provocation levels. These scores entered subsequently a repeated analysis of variance (ANOVA). This statistical analysis has the major disadvantage that information gets lost in the aggregation process (e.g., intraindividual variation). Therefore, researchers from the same work group focused on trial by trial analyses avoiding information loss (Wagels et al., 2018; Wagels et al., 2019; Weidler et al., 2018). Due to the advantages of the monetary mTAP (e.g., usage in special settings, reduced ethical problems), it is worth to investigate this paradigm more in detail on the basis of a validated standard procedure. More importantly, the monetary mTAP is particularly useful when studying aggression within the non-clinical normal range. Because, more aversive stimuli (e.g., electric shocks) are presumably associated with a higher threshold to exhibit aggression and could lead to reduced aggressive responding in a non-clinical sample.

1.4.3 Neurobiology of reactive aggression and externalization

Beside the methodological challenges in measuring aggression, the neurobiology underlying repetitive acts of aggression is of special interest (Siever, 2008). A majority of studies focus on psychiatric disorders associated with aggression (e.g., externalizing disorders such as ASPD and CD). These studies suggest that aggression is linked to abnormalities in three neural systems: experience of threat and aggression, decision making (e.g., cognitive control) and regulation of emotions (for review see Coccaro, Sripada, Yanowitch, & Phan, 2011). Functional neuroimaging studies with laboratory paradigms (e.g., TAP, PSAP) investigating aggression in the non-pathological range showed responses in the same neural systems during retaliation (for review see Fanning, Keedy, Berman, Lee, & Coccaro, 2017), underpinning a continuity of normal and pathological neurobiology of aggression.

In terms of externalizing disorders, trait impulsivity as common etiological factor is associated with variations in mesolimbic dopamine responding. These aberrations in reward-related areas could possibly explain the altered reward sensitivity and sensation seeking in participant exhibiting externalizing problems (Beauchaine, Zisner, & Sauder, 2017; Buckholtz et al., 2010). Furthermore, neuroimaging studies examining externalizing disorders revealed abnormalities in neural systems mediating decision making as well as emotion processing including acute threat responses (Blair, Leibenluft, & Pine, 2014; Fairchild et al., 2019). Especially, increased acute threat responses are thought to mediate impulsive aggression (Blair, 2016). Following the above-mentioned associations between transdiagnostic traits and their biobehavioral underlayers (e.g., RDoC), it can be assumed that these neural abnormalities in externalizing disorders should also be observable, to a lesser extent, in a non-clinical sample during reactive aggression.

Both reactive aggression and externalization are known to be linked to testosterone levels. Multiple lines of evidence suggest that testosterone (as well as cortisol) affects the striatal dopaminergic reward network and neural threat system and thereby enhances aggressive behavior (for review see Dekkers et al., 2019). This indicates that the neural basis of testosterone effects,

externalization, and aggression are strongly connected to each other. While most studies investigating externalizing disorders showed higher testosterone concentrations in externalization (Pajer et al., 2006), testosterone correlates merely modestly with behavioral aggression, according to recently conducted meta-analyses (Archer, Graham-Kevan, & Davies, 2005; Dekkers et al., 2019). It is assumed that the relationship between testosterone and aggression is highly moderated by other factors (e.g., gender, age, type of population, context, traits) (Dekkers et al., 2019). Especially, the context of competitive interactions plays a crucial role in this multifactorial concept. For detailed introducing information, see section 4.1.

In summary, the expression of multifactorial traits such as externalization is subjected to complex interactions at neural, physiological and psychological level und the underlying mechanisms of these interactions still remain unclear.

1.5 Aims of the thesis

The findings described above illustrate that the externalizing pathologies (e.g., CD, ADHD, ASPD, SUD) might be associated with an aberrant psychobiological stress responsivity, linked to an altered acute neural threat and reward processing (Beauchaine et al., 2017; Fairchild et al., 2019). Besides, abnormalities in neural networks mediating decision making and emotion processing seem to contribute to a higher rate of reactive aggressive behavior which is one of the common diagnosis criteria of externalizing disorders. The idea of biobehavioral substrates of trait dimensions (e.g., RDoc) supports the conclusion that such changes in stress and aggression processing should manifest also, to a lesser extent, in participants exhibiting externalizing behavior within a non-clinical range. However, the current knowledge on mechanisms underlying non-clinical externalization is scarce.

Therefore, a first main research question was whether interindividual differences in externalizing behavior within the non-clinical range are associated with different affective, psychoendocrine and neural stress responses. A further aim was to investigate the moderating role of non-clinical externalization on behavioral, affective and endocrine (cortisol, testosterone) responses as well as

neural activation patterns during experimental induction of aggressive behavior by the monetary mTAP. In this context, the developed version of the mTAP should be validated in a pilot study before the application in scanner environments.

To address the aims of this thesis, one behavioral study and two neuroimaging studies were conducted. Prior to the main studies of the thesis, a new version of the monetary mTAP was developed, based on the work of Kogan-Goloborodko et al. (2016) and Schneider et al. (2015), and subsequently validated (study 1, chapter II). Data was analyzed with a linear mixed model (LMM) approach considering individual aggression trajectories with respect to external and convergent validity. Based on this, in study 2 (Chapter III) healthy subjects from the low versus high range of the normal variation in externalizing behavior were exposed to Scan*STRESS,* a psychosocial stress paradigm for scanner environments. In study 3 (Chapter IV), these participants performed the monetary mTAP inside the MRI scanner.

The three studies are derived from original research articles, either published (study 1, chapter II) or currently under peer-review (study 2, chapter III; study 3, chapter IV), listed below. For improved readability, contents, tables, and figures were numbered continuously.

Chapter II

Konzok, J., Kreuzpointner, L., Henze, G.-I., Wagels, L., Kärgel, C., Weidacker, K., Schiffer, B., Eisenbarth, H., Wüst, S., & Kudielka, B. M. (2020). Validation of a monetary Taylor Aggression Paradigm: Associations with trait aggression and role of provocation sequence. *Journal of Experimental Social Psychology, 88*, 103960.

Julian Konzok, Ludwig Kreuzpointner, Brigitte M. Kudielka, and Stefan Wüst designed the study. Julian Konzok, Kathrin Weidacker, and Christian Kärgel worked out the technical details. Julian Konzok and Alexander Barbatsalos performed the experiments. Julian Konzok, and Ludwig Kreuzpointner performed the analysis. Julian Konzok drafted the manuscript. Ludwig Kreuzpointner,

Gina-Isabelle Henze, Lisa Wagels, Christian Kärgel, Kathrin Weidacker, Boris Schiffer, Hedwig Eisenbarth, Stefan Wüst, and Brigitte M. Kudielka provided editorial revisions and suggestions.

Chapter III

Konzok, J., Henze, G.-I., Peter, H., Giglberger, M., Bärtl, C., Massau, C., Kärgel, C., Schiffer, B., Eisenbarth, H., Wüst, S., & Kudielka, B. M. (2020). Externalizing behavior in healthy young adults is associated with lower cortisol responses to acute stress and altered neural activation in the dorsal striatum. Manuscript under review in *Psychophysiology.*

Brigitte M. Kudielka, Stefan Wüst, Hedwig Eisenbarth, Boris Schiffer, Gina-Isabelle Henze, and Julian Konzok designed the study. Julian Konzok, Gina-Isabelle Henze, Hannah Peter, and Marina Giglberger performed the experiments. Julian Konzok and Gina-Isabelle Henze performed the analysis. Julian Konzok drafted the manuscript. Gina-Isabelle Henze, Hannah Peter, Marina Giglberger, Christoph Bärtl, Claudia Massau, Christian Kärgel, Boris Schiffer, Hedwig Eisenbarth, Stefan Wüst, and Brigitte M. Kudielka provided editorial revisions and suggestions.

Chapter IV

Konzok, J., Henze, G.-I., Kreuzpointner, L., Peter, H., Giglberger, M., Bärtl, C., Massau, C., Kärgel, C., Schiffer, B., Eisenbarth, H., Wüst, S., & Kudielka, B. M. (2021). High externalization in healthy young adults predicts reduced neural activation within the inhibitory control network during reactive aggression. Manuscript under review in *Cognitive, Affective, & Behavioral Neuroscience.*

Brigitte M. Kudielka, Stefan Wüst, Hedwig Eisenbarth, Boris Schiffer, and Julian Konzok designed the study. Julian Konzok, Gina-Isabelle Henze, Hannah Peter, and Marina Giglberger performed the experiments. Julian Konzok performed the analysis and drafted the manuscript. Gina-Isabelle Henze, Ludwig Kreuzpointner, Hannah Peter, Marina Giglberger, Christoph Bärtl, Claudia Massau, Christian

Kärgel, Boris Schiffer, Hedwig Eisenbarth, Stefan Wüst, and Brigitte M. Kudielka provided editorial revisions and suggestions.

Chapter II:

Validation of a monetary modified Taylor Aggression Paradigm

2.1 Introduction

Laboratory aggression paradigms allow the observation of aggressive behavior under strict experimental control and, therefore, the investigation of causal relations (McCloskey & Coccaro, 2003). Furthermore, laboratory approaches offer the opportunity to experimentally induce events occurring infrequently in natural settings. One of these experimental methods is the TAP, also known as CRTT. This paradigm is widely used for the measurement of aggressive behavior in response to interpersonal provocation under laboratory conditions (Taylor, 1967; Warburton & Bushman, 2019). In the original version, participants were instructed to play a competitive reaction time game with another individual. The winner of each round of this preprogrammed game received the opportunity to administer a mild electric shock to the mock opponent. The behavioral response of the participant (intensity of administered shock) to the provocation by the mock opponent served as measure for reactive aggression, and was found to vary depending on the preceding provocation. Thus, the TAP captures reactive aggression, the affective form of aggression, occurring in response to provocation or frustration while proactive aggression, the instrumental type of aggression, is driven by obtaining desired goals (e.g., goods, coercive power, etc.; Crick & Dodge, 1996; Geen, 2001).

2.1.1 Variants of the Taylor Aggression Paradigm

Over the years, the TAP underwent various modifications regarding intensity (e.g., Giancola & Parrott, 2008; Gustafson, 1985) and duration of administered electric shocks (e.g., Giancola & Zeichner, 1995; Hoaken & Pihl, 2000), the number of trials, or quantification strategies (see Elson et al., 2014; Hyatt, Chester, Zeichner, & Miller, 2019). Moreover, studies revealed that mTAPs replacing electric shocks with noise blasts (e.g., Beyer et al., 2015; Böhnke, Bertsch, Kruk, Richter, & Naumann, 2010; Chester & Lasko, 2018; Ferguson et al., 2008; Krämer, Jansma, Tempelmann, & Münte, 2007), pneumatic pressure (e.g., Lotze et al., 2007; Veit et al., 2010), and heat (e.g., Krämer et al., 2011) effectively provoked reactive aggression as well. Of particular interest are recent

findings, showing that higher punishment levels after provocation did also occur when withdrawing money (Kogan-Goloborodko et al., 2016; Repple et al., 2017; Schneider et al., 2015; Wagels et al., 2018; Weidler et al., 2018) or points (Beyer, Buades-Rotger, Claes, & Krämer, 2017; Zepf et al., 2008) as provocation.

A key advantage of a monetary mTAP lies in its broad applicability. It can, for instance, be applied when electric shocks or noise stimuli are not feasible due to ethical concerns, for example in studies with children or certain patient groups. Furthermore, in specific laboratory settings (e.g., magnetic resonance imaging) the implementation of shock or noise stimuli might be difficult or even impossible. Beside these more operational aspects, the monetary mTAP matches the long tradition in psychological (aggression) research using monetary paradigms to investigate reward or punishment processing (see e.g., Point Subtraction Aggression Task; Cherek, 1981). With this, the monetary mTAP builds on a psychological, non-physical expression of reactive aggression. While research on different variants of mTAPs has accumulated until today, there is still a paucity of data on the mTAP with monetary stimuli.

Previous research using mTAPs implemented different types of sequences of provocation. For example, block designs of 6 to 60 trials with partly ascending provocation (Bertsch, Böhnke, Kruk, Richter, & Naumann, 2011; Beyer et al., 2015; Böhnke et al., 2010; Gustafson, 1985; Taylor & Epstein, 1967) as well as random order sequences (Anderson & Carnagey, 2009; Beyer, Münte, Erdmann, & Krämer, 2014; Beyer et al., 2014; Krämer et al., 2011) were applied. For the noise mTAP, a random ("ambiguous") versus increasing provocation sequence resulted in different aggression levels (Anderson et al., 2004). For a monetary version of the mTAP, only a random (but for all participants identical) order of stimuli (Kogan-Goloborodko et al., 2016; Repple et al., 2018; Repple et al., 2017; Schneider et al., 2015; Wagels et al., 2018) or a block-wise ascending provocation was implemented so far (Weidler et al., 2018). However, to date there is no systematic comparison of sequence effects in monetary mTAPs. This is surprising given the relevance of revenge motivation (Anderson et al., 2004) and expectancies for behavioral responses (e.g., Krigolson,

Hassall, & Handy, 2014; Talmi, Slapkova, & Wieser, 2019). Based on a reanalysis of data by Wagels and colleagues (2018; see below), we assumed that a certain sequence of similar provocation levels (low, medium, high) would model real behavior of a human opponent better and thus contribute to more pronounced experimental effects.

So far, studies on the convergent, discriminant, and criterion validity of mTAPs indicate good psychometric properties. First, significant positive correlations between behavioral aggression levels and self-report measures of aggression suggest high convergent validity (Beyer et al., 2014; Chester & Lasko, 2018; Giancola & Parrott, 2008; Giancola & Zeichner, 1995; Krämer, Büttner, Roth, & Münte, 2008; Schmidt, Zimmermann, Banse, & Imhoff, 2015; Sherrill, Magliano, Rosenbaum, Bell, & Wallace, 2016; Webster et al., 2014). Similarly, two recent studies using, in particular, the monetary mTAP reported a positive relationship between aggression levels and self-reported physical trait aggression (Weidler et al., 2019; Weidler et al., 2018). Second, Bernstein, Richardson, and Hammock (1987) presented evidence for discriminant validity by showing no associations between reactive aggression levels in a mTAP with shock stimuli and measures of competitive behavior. Third, criterion validity has been demonstrated by Kogan-Goloborodko et al. (2016) reporting higher levels of aggression assessed by a monetary mTAP in patients with borderline symptoms compared to healthy controls.

2.1.2 Gender effects in aggression research

Another important issue in the context of aggression induction und measurement are gender differences. To date, an extensive body of literature reports that men are more engaged in aggressive behavior than women in real-world as well as laboratory settings (Archer, 2004; Hyde, 1984). Gender differences seem to strongly depend on various context variables (Hyde, Bigler, Joel, Tate, & van Anders, 2018; Richardson & Hammock, 2007). A meta-analysis based on laboratory studies demonstrated that men show higher aggression levels under unprovoked conditions while this gender effect attenuates under provocation (Bettencourt & Miller, 1996). The authors concluded that provocation might minimize gender differences in aggression. Our current knowledge on

potential gender effects in the monetary mTAP is scarce since the majority of studies tested men or women exclusively. So far, only two studies compared men and women, based on relatively small sample sizes. While Repple et al. (2018) did not find any gender effects, Weidler et al. (2019) observed a provocation by gender interaction (lower gender differences under higher provocation).

2.1.3 Linear Mixed Models in aggression research

LMMs offer key advantages (e.g., the distinction between fixed and random effects, the possibility to account for a higher number of repeated observations, no exclusion of data due to missing time points and modeling of time-varying covariates). However, they are still rarely applied in aggression research. This also holds for research applying mTAPs (Chester, 2019; Webster et al., 2014; Weidler et al., 2019). Thus, in the current paper, we used growth curve models/random coefficient models (Singmann & Kellen, 2017; West, Welch, & Galecki, 2014), in order to capture the variation of the trajectory of the outcome variable over time as well as variation between subjects.

2.1.4 Research questions

Based on the open questions outlined above, the present study had four objectives. First, we aimed at replicating the effects of provocation on reactive aggression in a monetary mTAP using LMMs (reflected by a main effect). Second, we examined gender differences expecting higher aggression levels in men than women and a moderating role of provocation (reflected by a main effect of gender and interaction provocation by gender). Third, in terms of convergent validity, we hypothesized a significant relationship between self-reported trait reactive aggression and behavioral reactive aggression in the laboratory paradigm. Fourth, to explore the role of provocation sequence, the monetary stimuli were presented either randomly or in a fixed sequence. In contrast to the random sequence, the fixed sequence was generated mainly as triplets of the same provocation category (see below and Appendix B). Because of the more homogeneous provocation sequence in the fixed condition, we expected higher aggression levels after higher provocation and

lower aggression levels after lower provocation in this experimental condition compared to the random order.

2.2 Method

2.2.1 Participants

Two hundred and twenty-six participants (114 men, 112 women) volunteered for this experimental study. Sample size was determined before any data analysis. Seven volunteers were not eligible because of mental illness, regular drug use or intake of psychotropic medication. Ten participants had to be excluded from the analysis due to technical problems during the test session (n = 4) or because they did not complete the online questionnaire assessing trait aggression after the experimental session (n = 6). Thus, the final sample consisted of 209 participants comprised of 104 men and 105 women (mean age 21.89 years, SD = 2.86, range: 18 - 41). All participants gave written informed consent and received a monetary compensation of 12 € or course credit. The study was approved by the ethics committee of the University of Regensburg. We report all measures, manipulations, and exclusions in this study as well as method of determining the final sample size.

2.2.2 Materials and procedure

Upon arrival in the lab, two participants of the same gender were introduced to each other. In order to increase the credibility of the cover story, they were informed that they will act as opponents in a competitive reaction time task while being seated in adjacent rooms. Participants received written instructions and completed a questionnaire on demographic variables and health status. After a training session consisting of five trials, the participants performed the monetary mTAP independently in two separate rooms (see description of task below). Directly before and after the mTAP, subjects reported on their state affect using Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988). Subsequently, participants responded to a series of questions on the credibility of the cover story of the (mock) competitive reaction time task (deception check questionnaire). Within this questionnaire, participants were invited to speculate on the study

questions under investigation. Participants were encouraged to indicate if they noticed anything conspicuous about the game, and were asked whether they had any recommendation to improve the game. As a deception check, two independent raters examined the answers to these questions (inter-rater reliability: Cohen's kappa coefficient .86). In case of non-unanimous ratings, a third rater was involved and the disagreement was resolved by discussion. With this questionnaire, we identified participants who expressed some suspicion that they either have had no control over the outcome of the trials (win vs. lose) or might have played against a computer. The resulting binary coded variable "deception check" indicated suspicion vs. no suspicion. Additionally, the deception check questionnaire comprised four items concerning competitiveness (see items in Appendix A) using 4-point scales anchored at 1 (not true at all) and 4 (completely true). After the exclusion of three participants due to missing data, the McDonald's omega of the four items resulted in ω_{total} = .86, indicating a good internal consistency.

One day after the experimental session, participants received a link to the online trait aggression questionnaire (K-FAF) (Heubrock & Petermann, 2008). The K-FAF is a validated and widely used 49 items trait aggression questionnaire comprising five subscales: spontaneous aggression, reactive aggression, excitability, aggression inhibition and self-aggression. The answering format is a 6-point Likert scale ranging from 0 = not true at all to 5 = completely true. Additionally, subjects filled in the short-version of the NEO-Five-Factor Inventory (NEO-FFI-30; Körner et al., 2008), the Hospital Anxiety and Depression Scale (HADS-D; Herrmann-Lingen, Buss, & Snaith, 2011) as well as the Triarchic Psychopathy Measure (TriPM; Patrick, 2010). In total, participants spent about 90 min for their participation.

Monetary mTAP. We applied a monetary mTAP, which was a mock competitive reaction time task with a fictional opponent. Each trial consisted of a decision phase, the reaction time task itself and a feedback phase (see Figure 4). At the beginning of each trial (decision phase), the participant had to actively set a stake between 0 and 90 euro cents (expressed in steps of 10; i.e. 0, 10, 20 cents etc.). Following the recommendations by Tedeschi and Felson (1994), the amount of 0 cents

was included as nonaggressive option (see also Elson et al., 2014). The given preset starting point was always 45 cents and, thus, could not be chosen by the participant as final stake. For the reaction time task, participants were instructed to press a button as quickly as possible when a green circle appears on the screen. It was emphasized that after each trial the amount that had been preselected by the winner would be subtracted from the account of the loser (feedback phase). In case of winning, the participant constantly received 50 cents. While performing the task, participants did not receive any feedback of their current account balance.

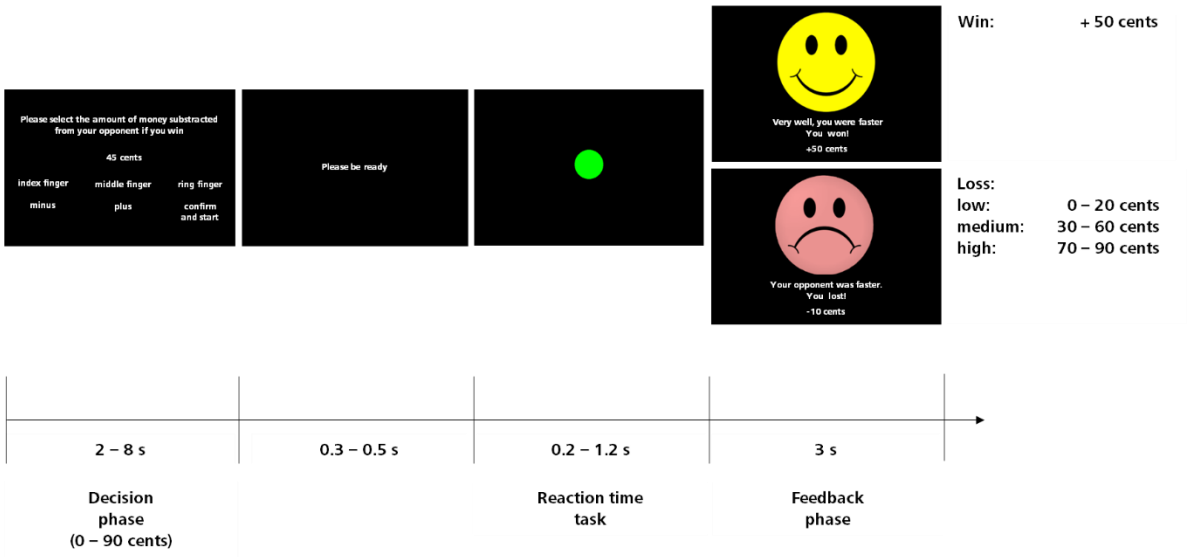

Figure 4. A single trial of the monetary mTAP composed of the decision phase, reaction time task, and feedback phase (screenshots are translated).

In total, the paradigm consisted of 100 trials, with preprogrammed 40 win and 60 lose trials. The entire run time of the task was 25 to 30 minutes depending on the duration of individual decision phases. For data analysis, the level of provocation (amount of money subtracted by the mock opponent in the preceding lose trial) was divided into low (0 - 20 cents), medium (30 - 60 cents) and high (70 - 90 cents). While focusing on reactive aggression, we excluded the first trial from the analysis due to its unprovoked characteristic.

Experimental condition: random versus fixed provocation sequences. The participants were assigned randomly to one of two experimental conditions. In the random order condition, the sequence of win and lose trials was presented in a different random order for each participant. To establish a fixed order condition, we re-analysed data from $N = 144$ participants who had performed a monetary mTAP provided by Wagels and colleagues (partly published in Wagels et al., 2018). We found that participants reacted to provocation of their fictional opponent more often reciprocally when the last three preceding provocation trials were of the same provocation category. Hence, the fixed order condition for the present experiment was generated by repeating trials from the same provocation category in triplets (see Appendix B for the specific provocation sequence in the fixed order condition). Thus, for low and high provocation, the maximum possible number of six triplets and two single trials were realized; consequently, for medium provocation the remaining four triplets, one duplet, and six single trials appeared in the sequence. In both experimental conditions, win trials and the different provocation categories occurred with identical frequency.

2.3 Results

2.3.1 Descriptives

Welch-test comparisons regarding demographics (age) and K-FAF subscales (self-reported trait aggression) did not show any significant differences between experimental groups (all $|t| < 1.30$, $ps > .194$, $Max(|d|) = 0.18$). Out of 209 participants, 50 (24%) expressed suspicion about the cover story (22 in the random and 28 in the fixed order condition). However, a chi-square test of independence for the variable deception check and the experimental condition was not significant ($\chi2(df = 1) = 0.87$, $p = .350$). Table 1 presents the demographic, behavioral, and psychometric variables (mean ± *SD*), separately for men and women.

2.3.2 Aggression levels: trial-by-trial analysis

To test our four study hypotheses, LMMs were performed in R (version 3.5.1; R Core Team, 2018) using lme4 (Bates, Mächler, Bolker, & Walker, 2014) and lmerTest (Kuznetsova, Brockhoff, &

Christensen, 2017). We analyzed participants' aggression levels as a function of the amount of money subtracted by the mock opponent in the previous trial (provocation, low vs. medium vs. high), the condition (random vs. fixed order), gender (man vs. woman), the interaction terms provocation by gender as well as provocation by experimental condition and the z-transformed reactive aggression scale of the K-FAF. Additionally, the binary coded variable "deception check" was added as covariate. The analysis had a power of 81% to detect an effect size of $d \geq .01$ (corresponding $R^2 \geq .01$) for a conditioned fixed effect. Power calculation was based on a simulation approach using simr (Green & MacLeod, 2016) following the recommendations of Arend and Schäfer (2019).

Table 1

Comparison (Mean ± SD) between men and women on demographic, behavioral, and psychometric variables (study 1)

	Man (n = 104)	Woman (n = 105)	*df*	*t*	*p*-value	*d*
Age	22.34 (± 2.68)	21.46 (± 2.98)	205.08	2.24	**.026**	0.31
Reactive aggression overall (in cents)	50.93 (± 24.71)	43.18 (± 22.56)	204.94	2.37	**.019**	0.33
Reactive aggression level (in cents)						
after low provocation	46.96 (± 25.02)	37.61 (± 22.26)	203.77	2.86	**.005**	0.40
after medium provocation	51.31 (± 24.35)	43.30 (± 21.93)	204.35	2.50	**.013**	0.35
after high provocation	55.40 (± 26.49)	48.45 (± 25.86)	206.77	1.92	.056	0.27
K-FAF						
Spontaneous aggression	13.22 (± 8.20)	8.42 (± 6.49)	195.84	4.69	**.001**	0.65
Reactive aggression	22.54 (± 9.93)	17.70 (± 9.24)	205.65	3.65	**.001**	0.51
Excitability	14.80 (± 8.61)	15.29 (± 8.74)	206.99	-0.41	.685	-0.06
Aggression inhibition	18.28 (± 5.45)	20.53 (± 5.40)	206.92	-3.00	**.003**	-0.42
Self-aggression	14.13 (± 7.62)	16.13 (± 8.55)	204.74	-1.79	.075	-0.25

Model building. We calculated four different models to fit a hierarchical linear structure to the data. The first defined null-model included a random intercept by subject accounting for interindividual variability only (model 1). Second, inclusion of provocation as fixed effect (model 2) led to an increase in goodness of fit ($\Delta\chi^2$ (df = 2) = 408.54, $p < .001$, ΔBIC = 426.84). The dummy coded variable provocation consisted of two estimates: Low provocation served as the reference standard to which the effects of medium and high provocation were related to. Third, in order to

determine the best structure for the random effects, model 2 was compared with the 2-level model 3 (provocation levels nested within subjects; level 1 = provocation level, level 2 = subject), to which we added a random slope for provocation by subject. Here, in addition to a general provocation effect for all subjects, the value of the random slope expresses the mean difference of aggression levels between low and medium provocation or low and high provocation across subjects. A likelihood ratio (LR) test resulted in an improved model fit ($\Delta\chi^2$ (df = 5) = 442.60, $p < .001$, ΔBIC = 389.7). Thus, beside the general fixed effect for provocation for all subjects, the influence of provocation on aggression varied significantly from subject to subject.

Considering covariance among grand-mean centered random components under medium provocation (compared to low), a significant negative correlation between random slope and random intercept (r = -.17, p = .012) could be observed. Under high provocation (compared to low), this relation appeared to be almost zero (r = -.02, p = .759). This implies that the subject-specific provocation effect from low to medium provocation on reactive aggression was weaker in subjects with a higher individual aggression level, whereas this was not the case from low to high provocation. However, both results imply some quadratic relation between provocation and individual aggression (see Figure 5), pointing to a higher variance of the provocation effect for subjects with lower and medium individual aggression (r_{x^2}(medium) = .24, $p < .001$; r_{x^2}(high) = .26, $p < .001$).

For the selection of the fixed effects structure, the 2-level final model 4 (provocation levels nested within subjects; level 1 = provocation level, level 2 = subject) was arranged with fixed effects for provocation, conditioned on experimental group (0 = random order, 1 = fixed order) and gender (0 = man, 1 = woman), the z-transformed reactive aggression scale of the K-FAF, the deception check (0 = no suspicion, 1 = suspicion reported) as well as the random intercept for subject and random slope for provocation by subject. A significant improvement of the model fit could be observed ($\Delta\chi^2$(df = 8) = 48.10, p = .002, ΔBIC = -27.33) and the marginal R^2 (mR^2, variance explained by the fixed factors) clearly increased from .01 to .06. When testing the final model exclusively with the

statistically significant effects (provocation, gender, and K-FAF) to diminish the "costs" (the higher number of estimated parameters) the BIC was reduced, reflecting an increased model fit (see Table 2 for detailed model comparison).

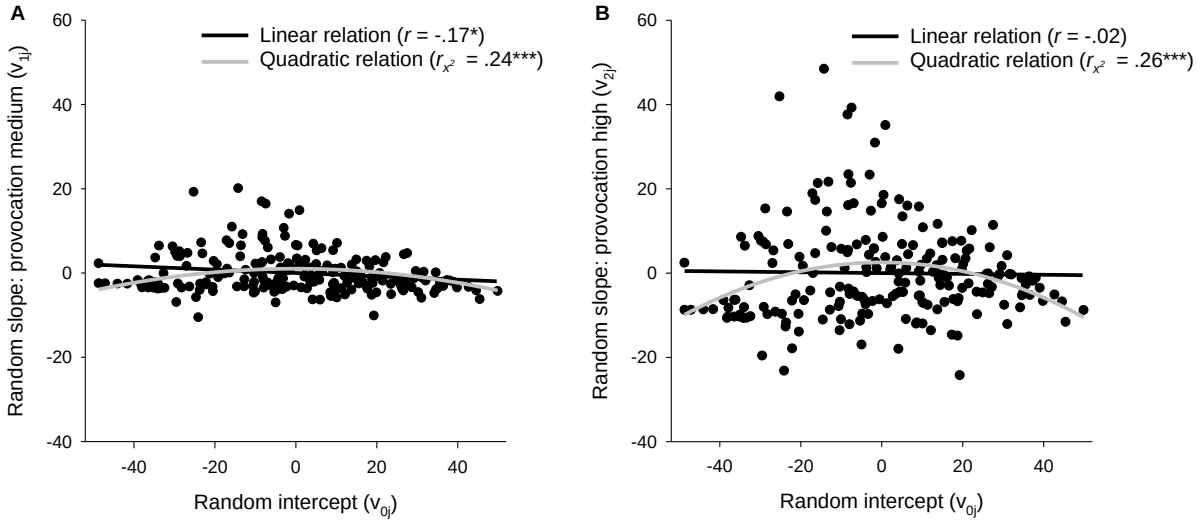

Figure 5. Covariance intercept by slope for provocation medium (v_{1j}) (A) and high (v_{2j}) (B) using grand-mean centering in study 1. *** $p < .001$, ** $p < .01$, * $p < .05$.

Final model. To quantify the sole main effects of final model 4, we performed *F*-tests (Satterthwaite approximation of the degrees of freedom) for overall effects of fixed factors (West et al., 2014), in addition to the *t*-tests for individual fixed-effect estimates (see Table 3). For provocation, we found a significant main effect of provocation ($F(2, 205.76) = 48.78$, $p < .001$, $\eta_p^2 = .32$). Thus, across all subjects and all conditions aggression levels increased with provocation. We further observed a significant main effect of gender ($F(1, 202.88) = 5.16$, $p = .024$, $\eta_p^2 = .02$) but no significant interaction provocation by gender ($F(2, 205.75) = 0.76$, $p = .471$, $\eta_p^2 < .01$). This indicates that men behaved more aggressively than women, however, this effect was not affected by provocation as hypothesized. Furthermore, a significant effect for trait reactive aggression (K-FAF) emerged ($F(1, 203.32) = 12.30$, $p < .001$, $\eta_p^2 = .06$). Hence, trait reactive aggression influenced the behavioral aggression levels during the task significantly. However, there were no

significant differences between the two experimental conditions (F(1, 695.02) = 0.70, p = .404, η_p^2 < .01) and no significant interaction of provocation by condition (F(2, 217.14) = 0.96, p = .383, η_p^2 < .01), suggesting that provocation sequence did not affect participants reactive aggressive behavior (see Figure 6A)

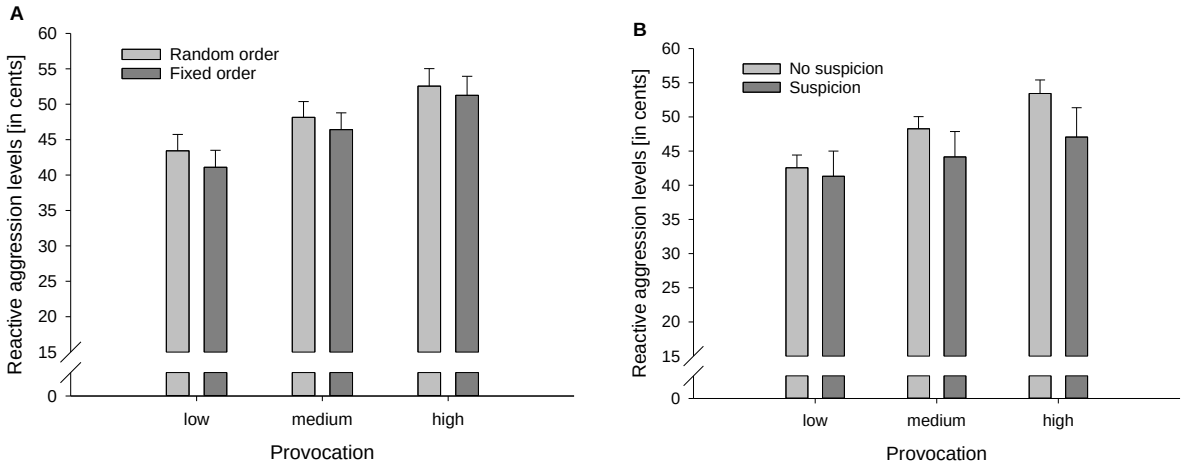

Figure 6. Mean (± SEM) aggression levels in study 1 in response to the provocation of the fictional opponent, separated for the random and fixed order condition (A), and for participants with and without suspicion (B).

Additionally, results demonstrated neither a significant effect for the variable deception check (F(1, 202.79) = 2.35, p = .127, η_p^2 = .01) nor the interaction provocation by deception check (F(2, 204.79) = 2.52, p = .083, η_p^2 = .02). At least, from a statistical point of view, suspicion about our cover story had no significant influence on the behavioral outcome (see Figure 6B). Table 3 illustrates the estimates for the fixed-effects coefficients of the final model, the corresponding standard errors of the mean (*SEM*) and *t*-test results.

Exploratory analysis. *Post-hoc*, we added the *z*-transformed mean competitiveness score as continuous predictor to the final model 4. Three participants had to be excluded due to missing data. Analysis resulted in a main effect of competitiveness (F(1, 198.55) = 26.98, p < .001, η_p^2 = .12). The goodness of the model fit improved significantly ($\Delta\chi^2$ (df = 1) = 27.60, p < .001, ΔBIC = 18.07) with a rise of the mR^2 from .06 to .12. At the same time, cR^2 as well as residual variance

remained unchanged and random intercept variance declined. Thus, the inclusion of competitiveness as predictor transferred a portion of the variance explained by random intercept for subject to the fixed effects.

Table 2

Four linear mixed models with reactive aggression as dependent variable in monetary mTAP (study 1)

	Model 1	Model 2	Model 3	Model 4
Intercept, γ_{00}	**47.15 *****	**42.25 *****	**42.26 *****	**47.25 *****
	(1.64)	(1.67)	(1.66)	(2.73)
Provocation (medium), γ_{10}		**5.04 *****	**5.03 *****	**3.70 *****
		(0.47)	(0.62)	(1.07)
Provocation (high), γ_{20}		**9.65 *****	**9.65 *****	**7.12 *****
		(0.47)	(0.99)	(1.71)
Condition, γ_{01}				0.59
(0=random order, 1=fixed order)				(2.39)
Gender, γ_{02}				**-8.47 ****
(0=man, 1=woman)				(3.22)
Deception check, γ_{03}				-5.63
(0=no suspicion, 1=suspicion)				(3.68)
K-FAF trait reactive aggression, γ_{04}				**3.82 *****
(normalized)				(1.09)
Provocation (medium) x condition, γ_{11}				1.35
				(1.23)
Provocation (high) x condition, γ_{21}				2.70
				(1.96)
Provocation (medium) x gender, γ_{12}				1.31
				(1.24)
Provocation (high) x gender, γ_{22}				2.36
				(1.99)
ICC: participant	.54			
AIC	113 005.37	112 600.81	112 168.24	112 136.14
BIC	113 027.65	112 637.95	112 242.52	112 269.85
Log Likelihood	-56 499.68	-56 295.41	-56 074.12	-56 050.07
Num. obs.	12 438	12 438	12 438	12 438
Num. groups: participant	209	209	209	209
Var: participant (Intercept), Var(v_{0k})	556.06	556.35	557.37	510.45
Var: residual, Var(ε_{ijk})	481.11	465.40	435.97	435.90
Var: provocation (medium), Var(v_{1k})			36.11	36.16
Var: provocation (high), Var(v_{2k})			162.88	162.75
Cov: participant (Intercept) provocation (medium)			-32.59	-28.21
Cov: participant (Intercept) provocation (high)			-24.11	-21.61
Cov: participant (Intercept) provocation (medium and high)			68.43	68.05
Marginal R^2		.01	.01	.06
Conditional R^2	.54	.55	.58	.58

Notes. Model 1 only consists of a random intercept for participant (null-model); model 2 includes a fixed effect for provocation; in model 3 and 4 a random slope for provocation by participant is added; in addition, model 4 (final full model) contains also fixed effects

for experimental condition, gender, deception check, and the trait reactive aggression scale of the K-FAF; *** $p < .001$, ** $p < .01$, * $p < .05$.

Table 3

Parameter estimates for the final model with reactive aggression as dependent variable in the monetary mTAP (study 1)

	beta	SEM	*t*	*p*-value
Intercept	47.25	2.73	17.28	**.001**
Provocation (medium)	3.70	1.07	3.47	**.001**
Provocation (high)	7.12	1.71	4.17	**.001**
Condition (0=random order, 1=fixed order)	0.59	2.39	0.25	.803
Gender (0=man, 1=woman)	-8.47	3.22	-2.63	**.001**
Deception check (0=no suspicion, 1=suspicion)	-5.63	3.68	-1.53	.127
K-FAF trait reactive aggression (normalized)	3.82	1.09	3.51	**.001**
Provocation (medium) x condition	1.35	1.22	1.10	.272
Provocation (high) x condition	2.70	1.96	1.37	.171
Provocation (medium) x gender	1.31	1.23	1.06	.291
Provocation (high) x gender	2.36	1.99	1.19	.236

2.4 Discussion

In line with previous studies in predominantly man samples (Kogan-Goloborodko et al., 2016; Repple et al., 2017; Schneider et al., 2015; Wagels et al., 2018; Weidler et al., 2018), the results of this experiment provide further evidence that the monetary mTAP is able to trigger reactive behavioral responses in a dose-response manner. Furthermore, to the best of our knowledge, the present experiment allowed for the first time the examination of potential gender differences in monetary mTAP responses in a larger sample of participants. Contrary to our hypothesis that provocation would minimize existing gender differences (Bettencourt & Miller, 1996), the significant gender effect and non-significant interaction term provocation by gender rather suggests that women generally behaved less aggressively than men, regardless of the amount of provocation. Only two earlier studies based on relatively small samples sizes investigated potential gender differences using the monetary mTAP and provided inconsistent results (Repple et al., 2018; Weidler et al., 2019). Studies using a mTAP with noise stimuli with small to medium sized samples (up to 80 participants) did primarily not observe gender differences in reactive aggression (Bond & Lader, 1986a; Bond & Lader, 1986b; Dambacher et al., 2015; Riva et al., 2017). Larger studies (samples sizes between 100 to 600 participants) reported reliably higher levels in men (Bushman, 1995; Bushman, 2002; Ferguson et al., 2008; Lawrence & Hutchinson, 2014) with relatively small amounts

of explained variance ranging between .02 and .12. At least, the consistent pattern of gender effects in larger studies with noise stimuli and our present results provide some evidence for the comparability of different versions of the mTAP. This implies that future research on the monetary mTAP needs to acknowledge gender as a factor to avoid incorrect variance allocations, even given the observed small effect size.

So far, evidence on the validity of a modified TAP with monetary stimuli is scarce (Kogan-Goloborodko et al., 2016; Weidler et al., 2019; Weidler et al., 2018). Adding new evidence regarding the convergent validity of the monetary mTAP, our results revealed a significant effect of self-reported trait reactive aggression on aggression levels. Even when considering the individual aggression trajectories during the task, this effect holds. This result suggests that we can use the monetary mTAP as asymptotic measure for reactive aggression acknowledging that there are, of course, other relevant influencing factors. Starting from here, a challenge for future investigations will be to link this self-report-laboratory-connection also with aggression levels in everyday life applying for example methods like ambulatory assessment.

To-date, there is a scarcity of data regarding the impact of different stimulus sequences using varying versions of the mTAP. Anderson and colleagues could show that a random ("ambiguous") versus increasing provocation sequence resulted in different aggression levels (Anderson, Anderson, Dorr, DeNeve, & Flanagan, 2000; Anderson et al., 2004). In our experiment, we compared a typical random order sequence (first experimental group) with a fixed order condition with stimuli presented in triplets from low, medium, or high provocation levels (second experimental group). With the fixed order condition, we aimed to simulate a more "naturalistic" strategic response of the mock opponent facilitating expectancies for the following behavior of the mock opponent (Krigolson et al., 2014; Talmi et al., 2019). However, we did not observe any difference in mean aggression levels between the two experimental conditions. Thus, we did neither find any support for the hypothesis that participants might show higher reactive aggression if confronted with an

unpredictable opponent nor for the assumption that participants would behave more (less) aggressively when provoked by more steady high (low) stimuli sequences.

Furthermore, because of the more "naturalistic" strategic response of the mock opponent in the fixed order condition, we expected a lower number of participants reporting suspicions regarding the (mock) competitive reaction time task. Contrary to this reasoning, we could not support this hypothesis. Furthermore, it can be concluded that participants in our experiment reacted mainly to the directly preceding provocation stimulus in terms of an exposure-response relation. With this, our results did not point to a significant role of the second and third to last provocation stimulus in the participants' response selection. Thus, it can be concluded that the specific provocation sequence has no impact on the behavioral aggression outcome (at least the one we applied). Based on the present data, we would recommend applying the random order version to avoid unfavorable side-effects (e.g., systematic inattention effects in specific paradigm phases).

Regarding a deception check, we developed a low-threshold questionnaire in order to be able to detect any potential suspicion concerning the cover story. Applying this approach, we found a relatively high suspicion rate of 50 participants (24%) in our sample. So far, only few studies applying the monetary mTAP scrutinized the impact of suspicion. Earlier studies reported much lower suspicion rates with the monetary mTAP (Schneider et al., 2015) as well as other variants of the TAP (e.g., shock stimuli; Giancola, Godlaski, & Roth, 2012). It can be speculated that the method of measurement might be essential here since such studies appeared to use only indirect assessments or unstandardized post-task-interviews performed by the experimenter. In contrast, Anderson and colleagues (2000, 2004) performed a structured interview identifying at least about 5 to 11% of highly suspicious participants. Deception is an important issue in aggression research since earlier work has shown that high suspicion potentially decreases behavioral aggression levels in laboratory settings (Anderson et al., 2000; Anderson et al., 2004). In the current study, we thus presented a series of low-threshold questions in order to be able to include this factor to our model. Interestingly, there was no significant difference in aggression outcomes in the mTAP between participants who

did or did not report any suspicion. This finding raises the idea that the provocative impact of the paradigm might remain relatively stable regardless of whether participants develop some suspicion that there might be no real opponent. However, future experimental studies need to investigate whether reactive aggression can also be reliably induced by the monetary mTAP if participants are explicitly instructed that they do not play against a "human" opponent, and if yes to which extend.

LMMs offered the benefit to consider effects at group level as well as at subject level among repeated data (Hesser, 2015). The final full model 4 with the best LMM fit could explain 58% variance considering all factors included in the design. With respect to the fixed factors (provocation level, experimental condition, gender, deception check as well as the interactions provocation by condition and provocation by gender), the explained variance was relatively low (6%). It can further be concluded that even when applying a restrictive analytical approach with a distinct allocation of variances, the expected aggression associated effects can be found in contrast to earlier publications (Chester, 2019; Webster et al., 2014). However, more research has to be done to identify other within-subject mediators according to theoretical implications (see e.g., the general aggression model; Anderson & Bushman, 2002). By accounting for such potential mediators (e.g., attitudes, intentions, motives), the paradigm could be significantly improved and the variance explained by the main aggression associated factors of interest (e.g., provocation, trait aggression) could be increased. On an exploratory level, we could show the potential of this approach with the factor competitiveness. The individual aggression level can partly be attributed to the tendency to act competitively.

Focusing on the random components' variances and their covariances, the main source of explained variance in our final model refers to individual aggression under low provocation. Thus, there is a high portion of variance to be declared as person-specific (44% in the final model 4) that could not be explained by gender or trait aggression. Results from covariance analysis of the random components suggest that participants with a medium individual aggression level were more strongly affected by the provocative character of the paradigm. In contrast, participants with a higher (lower)

individual aggression level selected higher (lower) aggressive responses in general, regardless of the provocation. For future work, it would be very interesting to identify underlying psychological constructs (e.g., game strategies) that correspond to these statistical observations.

Finally, important limitations should be taken into consideration. Since we recruited a sample consisting of young healthy adults, it remains unclear to what extent our results could be generalized to other populations. It also remains unknown if another operationalization of a fixed order condition or block designs with partly ascending/descending provocation would yield different results.

2.4.1 Conclusions

In sum, the present findings provide new evidence supporting the view that the monetary mTAP is a valid paradigm for the induction and measurement of behavioral reactive aggression in the laboratory. This holds true when considering individual trajectories instead of aggregated aggression parameters by using LMMs. First, the monetary mTAP induced reactive aggression in a dose-response manner, with women showing lower aggression levels than men. Further, aggression levels were significantly associated with trait reactive aggression. In contrast to our initial hypothesis, there was no indication that a fixed sequence order would be advantageous. In fact, the current level of reactive aggression appeared to be mostly affected by the immediately preceding provocation trial. Finally, suspicion by the participant regarding the cover story appeared not to invalidate reactive aggression as outcome variable in the monetary mTAP. Thus, the overall picture supports the view that the monetary mTAP can serve as a useful alternate paradigm for the measurement of reactive aggression.

Chapter III:

Externalization and stress reactivity

3.1 Introduction

The externalizing spectrum encompasses a range of heterogeneous personality traits and behavioral patterns, primarily characterized by disinhibition, impulsivity, antisocial-aggressive behavior as well as substance (mis)use (Krueger et al., 2002; Krueger et al., 2007; Patrick et al., 2013). It is closely related to, but conceptually distinct from psychopathy (Patrick, 2010), which is conceptualized by externalization and the additional phenotypic component boldness, expressing an underlying fearless disposition. For externalizing disorders (e.g., ASPD, CD and SUD), it is assumed that there might exist common underlying etiological mechanisms. This idea is supported by behavioral genetics studies showing significant heritability of a general externalizing factor (Beauchaine, Zisner, & Sauder, 2017; Blonigen, Hicks, Krueger, Patrick, & Iacono, 2005; Hicks et al., 2007). In addition, in large samples derived from healthy as well as clinical populations, it could be shown that externalization is a dimensionally distributed characteristic that can be found also in the non-pathological range of variation (Krueger et al., 2007; Markon & Krueger, 2005). Such individuals show disinhibitory traits like increased aggression as well as impulsivity or elevated scores in disagreeableness and unconscientiousness within the five-factor model of personality (DeYoung et al., 2008; John et al., 1994). Furthermore, negative emotionality, low fearfulness and low effortful control has been considered as most relevant pathways from individual temperament to externalizing psychopathology (Krieger & Stringaris, 2016).

From a developmental perspective, externalizing problems are related to deficits in emotion regulation as well as frequent stress exposure over the lifespan (Herts, McLaughlin, & Hatzenbuehler, 2012). The ability to react adequately to and cope with everyday emotion and stressful events has been considered one of the most relevant factors promoting mental health (Gross & Munoz, 1995). Thus, it can be hypothesized that externalization is linked to alterations in the psychobiological stress response not only in externalizing disorders but also within its non-clinical variation.

Regarding externalization-specific HPA axis stress responses, the majority of studies so far has shown an inverse relationship between externalization and acute cortisol stress responses. These

studies are based on clinical samples, e.g., children and adolescents with conduct disorder and disruptive behavior disorder (Fairchild et al., 2008; Van Goozen et al., 2000), violent adult offenders (Virkkunen, 1985), and substance misusers (Couture et al., 2008). Interestingly, reduced HPA axis responses to acute stress in patients with high externalization seem to coincide with higher emotional reactivity (McLaughlin, Hatzenbuehler, Mennin, & Nolen-Hoeksema, 2011). However, other previous studies could not confirm HPA axis hypo-responsivity in externalization (Alink et al., 2008). It can be speculated that, at least in part, such results might be explained by the fact that some stress induction paradigms failed to reliably induce robust cortisol responses. When it comes to externalization in a non-clinical range, we are aware of only one study in healthy college students, which reports a negative association between cortisol responses to the TSST and psychopathy scores (O'Leary et al., 2007). Taken together, there is still a lack of data on psychobiological stress responses in healthy adults exhibiting externalizing behaviors within a subclinical range.

Besides the cortisol stress response to stress, other endocrine parameters appear to be linked to the externalizing spectrum as well. In particular, testosterone levels were shown to be higher in clinical samples with conduct disorder (Pajer et al., 2006) and antisocial behavior (Yildirim & Derksen, 2012). Thus, it could be hypothesized that testosterone concentrations are also elevated in non-clinical participants showing higher externalization in the non-clinical range.

Based on lesion and other animal studies, it is assumed that HPA axis (dys)regulation in humans is mediated by an influence of higher order brain and limbic areas (e.g., medial prefrontal cortices, amygdala, hippocampus). While activity of the hippocampus and ACC seem to reduce glucocorticoid reactivity, amygdala activity potentiates cortisol stress responses (for review see Herman et al., 2005). The fact that these regions also play a key role in psychosocial stress regulation in humans could be confirmed during the last years with brain imaging studies (Akdeniz et al., 2014; Dahm et al., 2017; Henze et al., 2020; Lederbogen et al., 2011). However, both the specific pattern of activation and deactivation and the direction of cortisol-related associations appear to depend strongly on the used paradigm (for review see Noack et al., 2019). In sum, to the best of our

knowledge, so far no study focused on neural stress regulation in relation to externalization in the non-clinical range.

Previous studies already demonstrated that youth with conduct disorder (for review see Fairchild et al., 2019), patients with substance (mis)use (Koob & Volkow, 2010) as well as with antisocial behavioral tendencies (Oberlin et al., 2012) showed reduced activity in response to rewarding stimuli and acute threat in orbitofrontal cortex (OFC), ACC, striatum, and amygdala. Therefore, the externalizing spectrum seems to be linked to dysfunctional inhibitory control of limbic and mesolimbic regions by the orbitofrontal and prefrontal cortex as well as anterior cingulate cortex. Taken together, there is a notable overlap between neural networks that are altered in externalization and areas that are involved in HPA axis regulation.

The objective of the present study was fourfold. First, we aimed at examining psychobiological stress responses in healthy men and women with high versus low externalizing behavior in the non-clinical range. We expected a lower cortisol response as well as higher emotional reactivity to acute stress exposure in participants with high externalization. Second, testosterone levels were hypothesized to be generally higher in participants with high externalization. Third, we assumed that participants exhibiting high compared to low externalizing behavior show reduced neural responses to Scan*STRESS*, especially in the amygdala, striatum (nucleus accumbens, nucleus caudatus, and putamen) as well as prefrontal cortex (OFC, ACC). This expectation is based on the above-referenced blunted neural responses in these regions in externalizing disorders as well as their partly ascertained involvement in HPA axis regulation. Fourth, we aimed to identify brain regions associated with the assumed reduced cortisol stress response in participants with high compared to low externalization.

3.2 Material and methods

3.2.1 Participants

Sixty-five participants (32 men, 33 women) were initially recruited in this quasi-experimental fMRI study. Prior to experimental sessions, eligibility of volunteers was ascertained with an online assessment, including questionnaires on demographic variables, health status, MRI scanner contraindications, the TriPM and screening questions for the diagnostic modules of the Structured Clinical Interview for DSM-IV (SCID-II, Wittchen, Zaudig, & Fydrich, 1997). Participants were selected from a pool of volunteers (N = 784) based on their scores in the two subscales disinhibition and meanness of the TriPM and they were assigned to a group with either high (n = 33) or low (n = 32) externalization. Participant selection was based on data gained in a validation study for the TriPM in the general population (N = 1,476) (Eisenbarth, Castellino, Alpers, Kirsch, & Flor, 2012). For the present study, selected participants scored within the scopes of the highest (Q_{75}) or lowest quartile (Q_{25}) of the subscales disinhibition (Q_{75} = 36, Q_{25} = 29) and meanness (Q_{75} = 33, Q_{25} = 26) as derived from the aforementioned report. To avoid inclusion of individuals scoring high on psychopathy, volunteers scoring within the upper quartile of the TriPM subscale boldness (Q_{75} = 55) were not eligible for the present study. Further exclusion criteria were acute and chronic illness, mental and psychiatric disorders as assessed with the German version of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and Axis II Personality Disorders (SCID-II; Wittchen et al., 1997), criminal history, current use of drugs and medication containing glucocorticoids as well as MRI scanner incompatibility. Two participants dropped out during the screening due to exclusion criteria. Women not taking contraceptives (n = 15) were scheduled for the MRI sessions during the calculated luteal phase of the menstrual cycle (Wolfram, Bellingrath, & Kudielka, 2011) based on a chromatographic urinary ovulation test kit (gabmed GmbH, Köln, Germany).

Two further participants had to be excluded from the analysis after participation due to technical problems during the test session (n = 1) or poor image acquisition (n = 1). Thus, the final sample consisted of 61 participants (mean age 23.62 years, SD = 3.81, range: 18 – 34) comprising 31 men

(16 with high externalization) and 30 women (15 with high externalization). Heart rate (HR) data were only available for $n = 54$ ($n = 12$ men, $n = 13$ with low externalization) because of insufficient data acquisition. One participant missed to fill in the PANAS questionnaire at the last time point (see section 3.2.3).

Prior to participation, all participants gave written informed consent. Afterwards, they received a monetary compensation of 100 € or course credits. This experiment was approved by the ethics committee of the University of Regensburg.

3.2.2 Procedure

For the scanner session, participants arrived 90 min prior to the test session inside the MRI scanner. During the initial 45 min relaxation phase, participants were watching a neutral movie. Fourty-five min prior to the start of the Scan*STRESS* paradigm (not before 1:00 pm), participants were administered 75 g customary glucose mixed with 200 ml of camomile tea to facilitate cortisol reactivity (Henze et al., 2020; Zänkert, Kudielka, & Wüst, 2020). The following one-hour MRI session consisted of the Scan*STRESS* paradigm (see section 3.2.3), an 18 min resting state, and a 12 min anatomy sequence. Subsequently, participants filled out a questionnaire package including the German version of the Barratt Impulsiveness Scale – Short Version (BIS-15; Spinella, 2007) and the "Wortschatztest" (WST; German for vocabulary test; Schmidt & Metzler, 1994) (see section 3.2.5).

During the experimental session, we collected saliva samples by Cortisol Salivettes (Sarstedt, Nümbrecht, Germany) at ten time points: -75 min, -15 min, -1 min before the start of Scan*STRESS* as well as +15 min, +30 min, +50 min, +65 min, +80 min, +95 min and +110 min (C1 – C10) thereafter (see Figure 1). At each time point, participants completed the state version of the PANAS. For later testosterone analysis, native saliva was collected by passive drool using polypropylene tubes at three time points: -15 min before Scan*STRESS* onset and +65 min as well as +95 min thereafter (pooled for analysis). During Scan*STRESS,* we assessed HR in beats per minute (bpm) (see Figure 7).

On a second test day, participants performed a monetary mTAP, a second resting state and diffusion tensor imaging (DTI) sequence and completed another questionnaire package (results to be reported elsewhere).

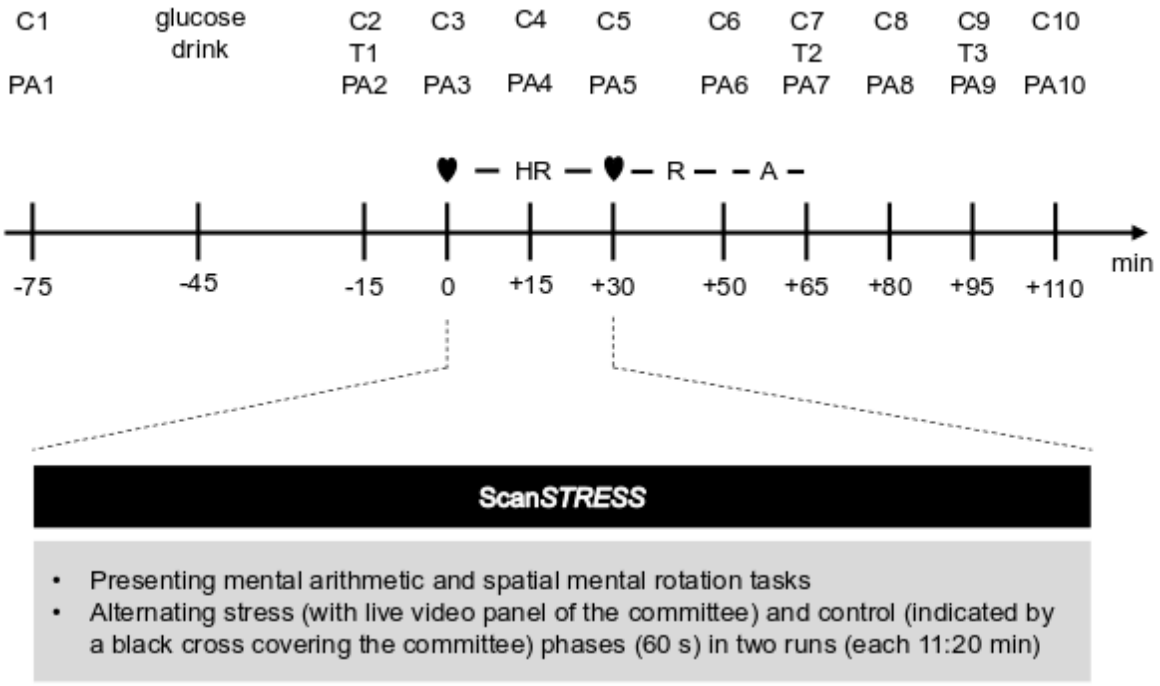

Figure 7. Flow chart of the experimental procedure during the MRI session including the ScanSTRESS exposure. *C1 – C10*: cortisol samples, T1 - T3: testosterone samples, PA: Positive and Negative Affect Schedule, HR = heart rate, R = resting state, A = anatomy.

3.2.3 Scan*STRESS*

Scan*STRESS* is an adaption of the TSST for scanner environments consisting of two runs. According to Streit et al. (2014), participants were instructed to respond to a mental arithmetic and spatial mental rotation task (see section 1.3.4) under a stress condition (performance trials, confrontation with an observation panel giving feedback) and a control condition (no observation and feedback by panel). We modified the initial protocol by prolonging the relaxing phase, administering the glucose drink, and shortening the transition time into the scanner (see Henze et al., 2020). For detailed information, we refer to section 1.3.4.

3.2.4 Materials, biochemical analysis, and data acquisition

After test sessions, saliva samples were stored at -20 °C. Analyses were performed by the biochemical laboratory at the University of Trier, Germany. Cortisol was assayed in duplicate using a time-resolved immunoassay with fluorometric detection (DELFIA); testosterone was assessed by a commercially available assay kit (Demeditec Diagnostics GmbH, Kiel, Germany). Inter- and intra-assay coefficients of variation were below 10%, respectively. Visual inspection of individual cortisol trajectories and testosterone concentrations did not reveal any unphysiological levels.

HR was extracted for four second intervals with the MR compatible pulse oximeter Nonin Model 7500FO (Nonin Medical B.V., Minnesota, United States).

Participants were scanned in a MAGNETOM 3T Prisma scanner (Siemens AG; Erlangen, Germany) equipped with a 64-channel head coil. A T2*-weighted echo-planar imaging (EPI) sequence (TR = 2000 ms; TE = 30 ms, flip angle = 90°, FOV = 192 x 192 mm^2, matrix size = 64 x 64 mm^2, 37 slices, slice thickness = 3.0 mm, 1.0 mm gap, voxel size = 3 x 3 x 3 mm^3, interleaved) was used resulting in 340 functional scans per run (in total 680) per participant and a T1-weighted magnetization-prepared rapid gradient-echo (MP-RAGE) sequence (TR = 2400 ms, TE = 2.18 ms, flip angle = 9°, voxel size = 0.8 x 0.8 x 0.8 mm^3, distance factor: 50%).

3.2.5 Psychometric measures

We used the TriPM with 58 items comprising the three subscales disinhibition, meanness, and boldness for the classification of our quasi-experimental groups (see above). Items of the TriPM are mainly adopted from the ESI (Krueger et al., 2007). Items are answered on a 4-point scale ranging from 1 = *not true at all* to 4 = *completely true*. In a community sample a Cronbach's α of .87 for the total score proved as excellent (van Dongen, Drislane, Nijman, Soe-Agnie, & van Marle, 2017). For the assessment of psychological responses to acute stress exposition, we applied the PANAS. The PANAS is a widely used 20 item-questionnaire assessing both positive and negative affect using a 5-point scale ranging from 1 (*not at all*) to 5 (*very much*). Finally, the questionnaire package included the German version of the BIS-15 consisting of three subscales (nonplanning, motor

impulsivity, attention impulsivity) and a total score as well as the WST. The WST is a vocabulary test which was formerly introduced as a proxy for verbal intelligence.

3.2.6 Data analysis

To assess potential differences between the two quasi-experimental groups in demographic variables, questionnaire scores, and testosterone concentrations, we conducted Bonferroni-corrected independent Welch-test comparisons using R (version 3.5.1; R Core Team, 2018) with the packages afex (Singmann, Bolker, Westfall, Aust, & Ben-Shachar, 2020), car (Fox & Weisberg, 2019), haven (Wickham & Miller, 2019), psych (Revelle, 2019), and sjstats (Lüdecke, 2020). Cortisol increase was defined by the difference between the individual cortisol peak (based on cortisol samples C5, C6, C7) and the pre-stress cortisol level (sample C3). Additionally, participants were grouped into responders versus non-responders according to the 1.5-nmol/l-criterion (Miller, Plessow, Kirschbaum, & Stalder, 2013).

Using R, repeated measures analyses of variance (ANOVAs, Greenhouse-Geisser corrected) were performed for salivary cortisol ('time' [10 cortisol samples] x gender [woman, man] x 'externalization' [high, low]), HR ('condition' [stress, control] x 'externalization' [high, low]), and affective stress response ('time' [10 assessments] x 'externalization' [high, low]).

fMRI data analysis. Imaging data were analyzed with FSL 6.0 (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl, Oxford, UK) using FEAT (FMRI Expert Analysis Tool). The following processing steps were conducted: motion correction by means of MCFLIRT, slice timing correction, non-brain removal using BET, intrasubject coregistration and registration to standard space defined by the Montreal Neurological Institute (MNI) using FLIRT and FNIRT, grand-mean intensity normalization, spatial smoothing with a Gaussian kernel of 8 mm full-width-at-half-maximum, high-pass filter correction of 120 s, time-series statistical analysis using FILM and region of interest (ROI) analysis applying fslmaths and featquery. The *z* (Gaussianized *t/F*) statistic images were thresholded nonparametrically using clusters determined by either $z > 3.1$ or $z > 2.3$.

For each run (first level analysis), a general linear model (GLM) was fitted with the two regressors of interest, namely stress and control, as well as eight regressors of no interest, namely the two announcement phases and the six realignment parameters. In a next step, data from each participant and run were entered into between-session analysis (second level) estimating mean responses for each participant. Subsequently, subject's mean responses were analyzed with a between-subject group model (third level) resulting in group mean responses.

For condition- and group-specific effects on whole brain level, the statistical images were thresholded (two-tailed combined test) with family-wise error rate (FWE) $p < .025$ (two-tailed combined test, FWE < .05). The SPM Anatomy toolbox (Eickhoff et al., 2005) was appropriated for anatomy labeling. For condition-specific effects we performed one-sample paired *t*-tests and for externalization effects two-sample *t*-tests. *Post-hoc*, Bonferroni-corrected ROI analyses (repeated measures ANCOVAs, between-subject factor 'externalization' [high, low]) were conducted using six predefined masks in amygdala (520 voxels), anterior cingulate cortex (1531 voxels), nucleus accumbens (126 voxels), putamen (1517 voxels), nucleus caudatus (901 voxels) and OFC (568 voxels) (all bilaterally) guided by the empirically-derived hypothesis mentioned above (see introductory section). These binarized masks were created with the Harvard-Oxford cortical atlas (http://www.cma.mgh. harvard.edu), using fslmaths. During this binarization process, all voxels get the value one which have at least the probability of 50 % being part of the specific ROI; the other voxels contain the value 0. Subsequently, we extracted mean beta-values for each ROI.

Finally, a two-group model with continuous covariate using the grand mean centered cortisol increase was conducted to test whether the relationship between cortisol increase and the neural response differs between the high versus low externalization groups (high > low and low > high, thresholded at FWE < .05). For *post-hoc* ROI analyses (Bonferroni-corrected), the parameter estimates in significant brain locations (two masks) were entered into a multiple regression analysis with the categorical variable externalization and cortisol increase as continuous predictor. To

capture potential gender effects whole brain analyses were conducted separately for men and women using one mask from the previous contrast (whole sample, FWE corrected at .05).

3.3 Results

3.3.1 Descriptives

The comparison between the two quasi-experimental groups (high vs. low externalization) in demographic (age), behavioral (BIS-15, WST), and psychometric variables showed that participants solely differed significantly regarding the two expected dimensions of the TriPM, disinhibition (Cronbach's α: .88) and meanness (Cronbach's α: .91) as well as the four subscales of the BIS-15 (see Table 4).

3.3.2 Endocrine, physiological, and psychological stress responses

For salivary cortisol responses, we found a significant main effect of 'time' (F(3.11,178.54) = 9.06, p < .001, η_p^2 = .13) as well as a significant interaction 'time by externalization' (F(3.11,178.54) = 4.21, p = .006, η_p^2 = .07) with the high externalization group showing lower cortisol responses to stress. The main effect of 'externalization' did not reach significance (F(1,59) = 2.74, p = .100, η_p^2 = .04). Due to enhanced SEM in cortisol trajectories (see Figure 8), we performed cook distance analyses for cortisol increase and externalization in order to identify potential effects of variational cases. Analyses revealed no value over the cut-off of 1 while the more conservative 4/N-criterion was exceeded by three participants (3 men with low externalization) pointing to a possible moderate predominance of these cases. However, when excluding the three participants from cortisol analyses, the interaction 'time by externalization' remained significant (F(2.98,166.64) = 2.97, p = .034, η_p^2 = .05).

Table 4

Mean ± SD of demographic, psychometric, behavioral, and hormonal data and results of welch tests comparing subjects with high versus low externalization groups (study 2)

	High externalization		Low externalization					
	Man (*n* = 16)	Woman (*n* = 15)	Man (*n* = 15)	Woman (*n* = 15)	*t*	*df*	*p*-value	*d*
Age (yrs.)	24.31 (± 3.38)	21.87 (± 4.31)	25.67 (± 4.34)	22.53 (± 2.20)	0.98	58.94	.330	0.25
TriPM								
Disinhibition	40.25 (± 4.57)	43.07 (± 6.03)	27.07 (± 1.62)	27.20 (± 1.82)	-14.16	35.99	.001*	-3.60
Meanness	40.31 (± 5.45)	39.27 (± 3.85)	23.20 (± 1.86)	22.80 (± 1.86)	-18.52	39.25	.001*	-4.69
Boldness	47.75 (± 5.52)	47.60 (± 4.55)	51.53 (± 2.97)	48.67 (± 3.75)	2.17	54.85	.034	0.56
BIS-15								
Non-planning impulsivity	13.31 (± 3.11)	12.93 (± 3.37)	10.53 (± 2.67)	9.80 (± 2.40)	-4.03	56.75	.001*	-1.03
Motor impulsivity	11.56 (± 2.16)	11.87 (± 3.04)	10.13 (± 2.03)	9.73 (± 1.87)	-3.05	55.46	.001*	-0.78
Attentional impulsivity	11.38 (± 2.06)	11.33 (± 1.40)	9.8 (± 1.97)	9.40 (± 1.45)	-3.97	58.98	.004*	-1.02
Sum	36.25 (± 5.36)	36.13 (± 5.71)	30.47 (± 4.81)	28.93 (± 3.56)	-5.22	56.42	.001*	-1.33
WST raw	33.00 (± 2.10)	32.40 (± 1.64)	33.33 (± 4.24)	32.53 (± 2.00)	0.33	45.94	.746	0.08
Cortisol increase (nmol/l)	2.36 (± 2.09)	1.43 (± 2.85)	6.28 (± 6.46)	1.80 (± 1.65)	2.04	41.46	.047	0.53
Responder/non-responder	10/6	6/9	11/4	4/11				

Notes: yrs. = years, nmol/l = nanomol per liter, TriPM = Triarchic-Psychopathy-Measure, BIS-15 = Barratt Impulsiveness Scale – Short Version, WST = "Wortschatztest". * comparison survived Bonferroni correction at $p < .05$.

Further covariance analyses revealed a significant effect of 'gender' ($F(1,57) = 6.95$, $p = .011$, $\eta_p^2 = .11$) and a 'time by gender interaction' ($F(3.22,177.39) = 3.32$, $p = .020$, $\eta_p^2 = .06$), indicating higher cortisol to stress responses in men (see Figure 8). The three-way interaction 'time by externalization by gender' missed the level of significance ($F(3.22,177.39) = 2.34$, $p = .072$, $\eta_p^2 = .04$).

Post-hoc-ANOVAs conducted separately for men and women yielded a 'time' effect in both men ($F(2.88,83.46) = 7.37$, $p < .001$, $\eta_p^2 = .20$) and women ($F(3.14,87.84) = 3.56$, $p = .020$, $\eta_p^2 = .11$). The interaction 'time by externalization' reached significance for men ($F(2.88,83.46) = 4.20$, $p < .001$, $\eta_p^2 = .13$; $F(2.29,59.49) = 2.43$, $p = .089$ if the three variational cases are excluded) but not for women ($F(3.14,87.84) = 0.76$, $p = .522$, $\eta_p^2 = .03$) (see Figure 8A and 8B). Table 4 displays the number of responders and non-responders in the experimental groups. A chi-square test of independence with the variables group and (non)responder (see Table 4), testing whether the 'externalization' effect could be causally associated with an unequal number of cortisol responders ($n = 31$) and non-responders ($n = 30$) in the two externalization groups, rendered non-significant ($\chi^2(df = 1) = 0$, $p = 1$). Welch-test comparisons regarding testosterone concentrations did not show any significant differences between externalization groups ($t(58.79) = -0.55$, $p = .584$, $d = - 0.14$) (see Figure 8D).

Results from HR analysis yielded a significant main effect of 'condition' ($F(1,52) = 86.75$, $p < .001$, $\eta_p^2 = .63$), but no significant difference between the two externalization groups ($F(1,52) = 0.08$, $p = .779$, $\eta_p^2 = .00$) (see Figure 8C) and no interaction.

Analysis of affective stress responses showed a 'time' effect for positive ($F(4.91,284.79) = 30.17$, $p < .001$, $\eta_p^2 = .34$) as well as negative affect ($F(3.64,214.69) = 29.89$, $p < .001$, $\eta_p^2 = .34$). Moreover, a main effect of 'externalization' could be found for positive affect ($F(1,58) = 5.52$, $p = .020$, $\eta_p^2 = .09$) showing a significantly more positive affective state in the low externalization group. No other effects were significant.

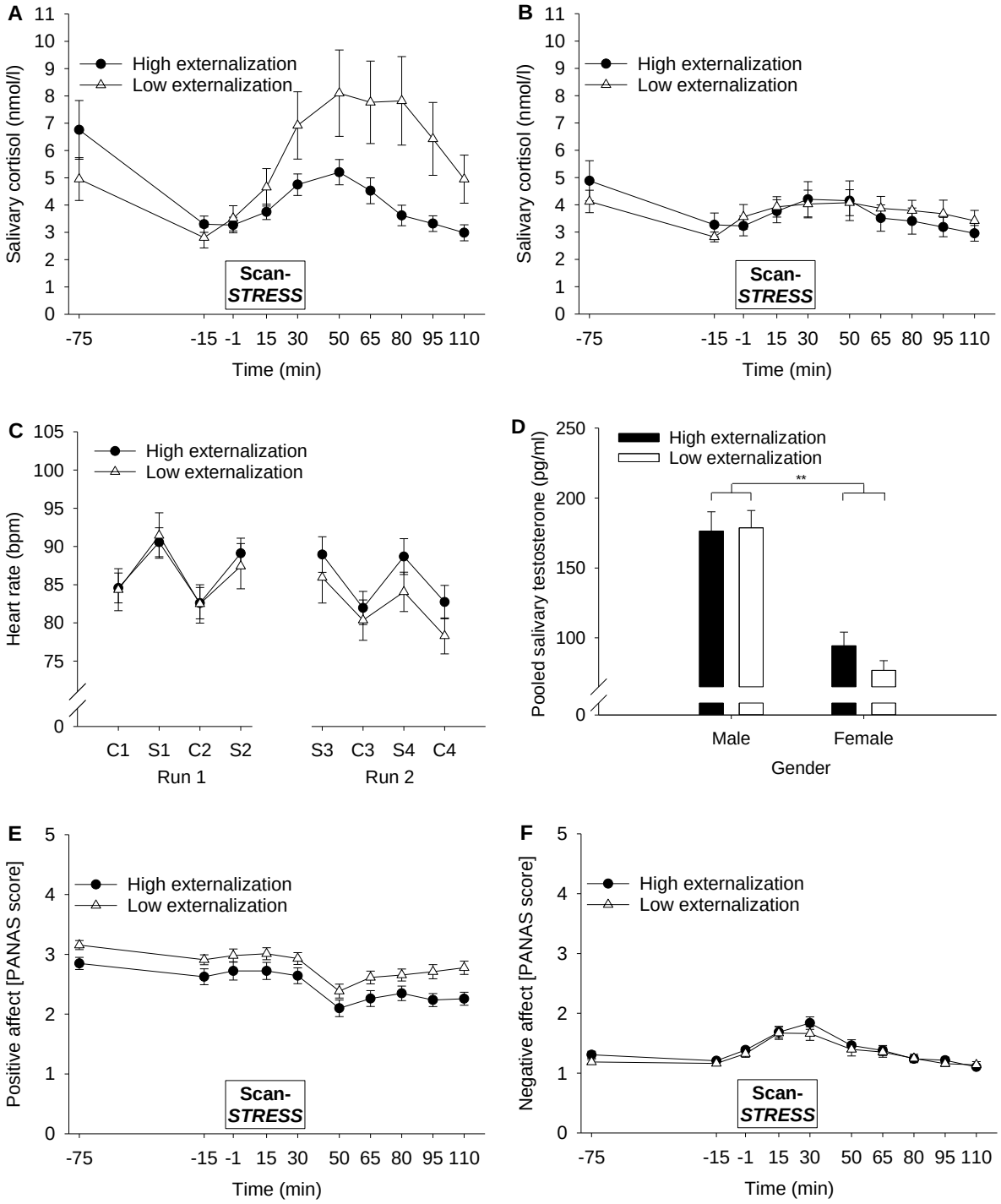

Figure 8. Hormonal, physiological and affective data comparing the high versus low externalization group. Salivary cortisol responses in (A) males and (B) females. Mean HR responses across conditions (C1 – C4: control phases, S1 – S4: stress phases) in run 1 and 2 (C). Pooled salivary testosterone concentrations in males and females (D). PANAS scores for positive (E) and negative (F) affect. Error bars represent standard error of the mean. ** $p < .01$, * $p < .05$.

3.3.3 Manipulation check: neural stress response

Contrasting stress versus control blocks (stress > control) on whole brain level, we observed a differential activation network including the left and right insula, the triangularis part of the left inferior frontal gyrus as well as left and right thalamus (see Appendix C). The opposite contrast (control > stress) was related to four clusters including the left and right prefrontal cortex (superior frontal gyrus, orbital gyrus), left posterior cingulate cortex, and left and right insula (see Appendix D). Affective stress responses are illustrated in Figures 8E and 8F.

3.3.4 Neural correlates of externalization

On whole brain level, no suprathreshold cluster could be observed by contrasting the high versus low externalization groups (high > low, low > high). None of the *post-hoc* ROI analyses revealed a significant externalization effect in the expected regions amygdala, anterior cingulate cortex, nucleus accumbens, putamen, nucleus caudatus, and OFC bilaterally (*F*s < 1.06, *p*s > .308). Further, analyses showed no significant externalization by gender interactions (*F*s < 1.27, *p*s > .265).

To account for the association between central nervous system and HPA axis responsivity to acute psychosocial stress, the individual cortisol increase (grand mean centered) was used as a covariate (FWE-corrected p < .05). The linear relationship between the cortisol increase (see Table 4) and three cluster including the left putamen and left nucleus caudatus (dorsal striatum) was found to be different between the two groups showing more activation in the low externalization group (low > high, see Figure 9A and Appendix E). For the opposite contrast (high > low) no suprathreshold cluster survived. *Post-hoc* ROI analysis revealed a significant interaction effect 'externalization by cortisol increase' in the bilateral putamen ($F(1,57) = 6.83$, $p = .011$, $\eta_p^2 = .11$) and nucleus caudatus ($F(1,57) = 5.62$, $p = .021$, $\eta_p^2 = .09$), surviving correction for the two significance tests. When excluding the above identified variational cases (see section 3.3.2), this effect remains significant for the putamen ($F(1,54) = 4.18$, $p = .046$, $\eta_p^2 = .07$) and nucleus caudatus ($F(1,54) = 5.70$, $p = .021$, $\eta_p^2 = .10$), although the p-value for the putamen surpassed the Bonferroni corrected significance level. In order to test whether these effects might be mainly driven by high externalization, correlational analysis were performed. Indeed, a

significant negative correlation between beta estimates and cortisol increase in the high externalization group (n = 31) could be observed in the putamen (r = -.41, p = .016, one-tailed) and nucleus caudatus (r = -.39, p = .012, one-tailed) surviving Bonferroni-correction. In contrast, in the low externalization group (n = 30), this correlation was non-significant (putamen: r = .20, p = .140, one-tailed; nucleus caudatus: r = .23, p = .113, one-tailed). Further, for both the putamen (z = 2.66, p = .008) and nucleus caudatus (z = 2.37, p = .018) correlation coefficients of the two externalization groups differed significantly from each other. Thus, in participants with high externalization, higher beta estimates during stress (vs. control) in the nucleus caudatus and putamen were associated with lower cortisol increases. In contrast, in participants with low externalization, beta estimates in these regions were not linked to cortisol increases. Figure 9 depicts scatterplots of mean beta estimates and cortisol increases differentiated by externalization group in the bilateral putamen (Figure 9B) and nucleus caudatus (Figure 9C).

To explore gender differences in the covariation of externalization-specific neural stress responses and cortisol increases, the same procedure was applied for men and women separately with a mask including the bilateral nucleus caudatus and putamen. Only for women, results revealed two supra-threshold clusters (FWE corrected at .05) including the right putamen (Z = 3.88, x = 20, y = 16, z = -2) and right nucleus caudatus (Z = 3.35, x = 12, y = 10, z = 10). For men, no supra-threshold cluster survived corrections.

3.4 Discussion

Over the last decades, evidence has accumulated that externalization in the pathological range is associated with altered psychobiological stress regulation. Based on these findings, the aim of the present study was to investigate cortisol and neural stress responses in relation to externalization in the non-clinical range.

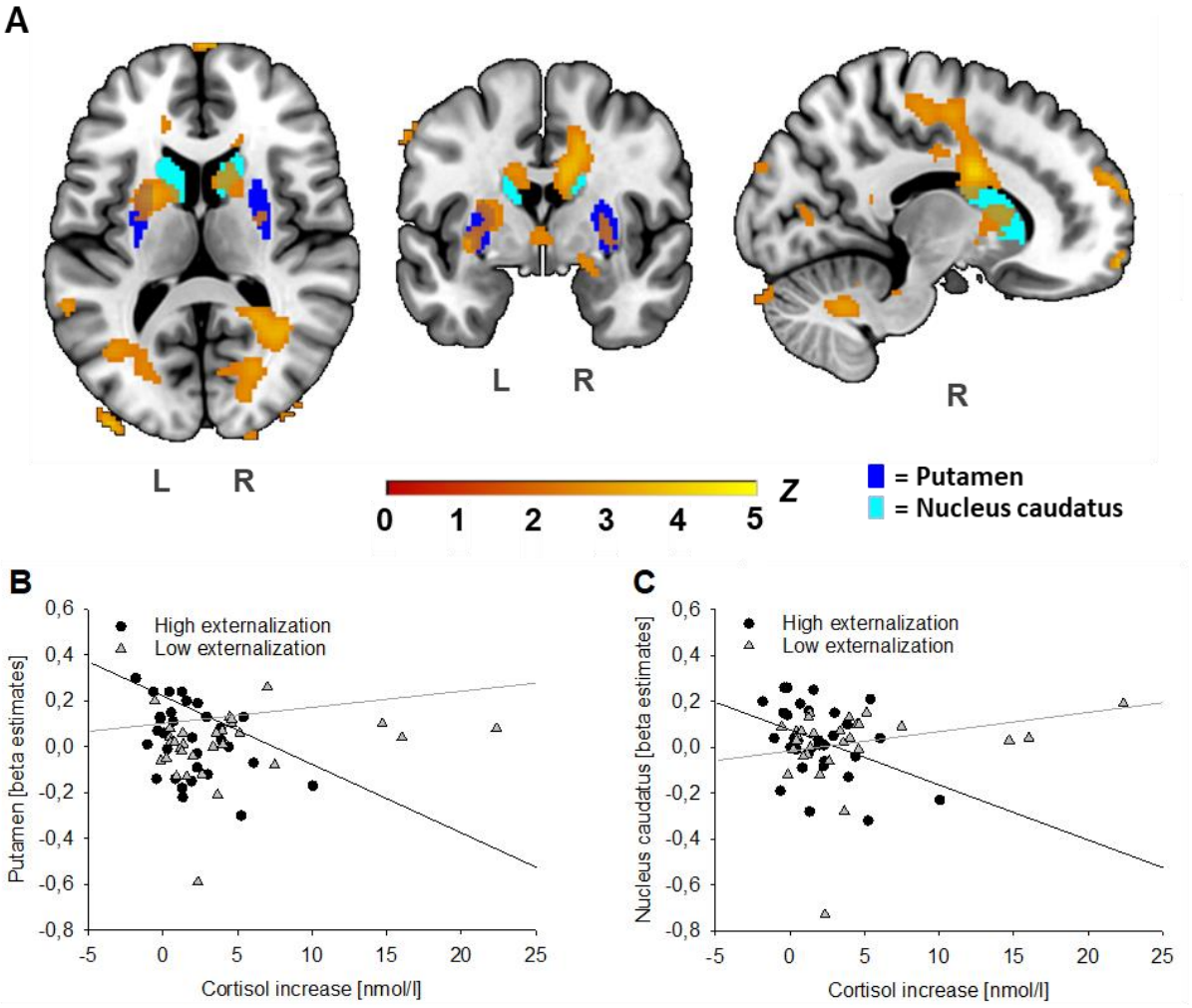

Figure 9. Externalization-group difference analysis (low vs. high externalization) with cortisol increase (continuous covariate) interaction (grand mean centered, two-tailed combined FWE-corrected $p < .05$) (A). Results of ROI analyses for 'cortisol increase by externalization' interaction in the bilateral putamen (B) and nucleus caudatus (C). L = left; R = right.

We observed reduced cortisol responses to stress in healthy participants with high compared to low externalizing behavior reflected by the interaction 'time by externalization'. This finding is consistent with previous clinical findings (Fairchild et al., 2008; Van Goozen et al., 2000). Based on these studies, it was hypothesized that altered HPA axis responsivity in participants with high externalization may be associated with the experience of chronic stress. According to the concept of allostasis and allostatic load (McEwen, 1998a) (see also section 1.3.3), it was suggested that significant stress exposure during childhood and adolescence triggers recurring cortisol surges to acute stress and, if chronic, this might lead to a downregulation of HPA axis functioning in the long run (Alink et al., 2008). Other proposed

concepts explaining the link between externalizing behavior and hyporeactivity are sensation-seeking (Robinson & Berridge, 1993) and low-fear (Raine, 2013) theories. They suggest that individuals with a physiological underarousal might engage in sensation-seeking behaviors in order to stimulate themselves by, e.g., antisocial behavior. Therefore, it is assumed that blunted cortisol responses are associated with reduced experiences of fear, which would be required for learning from negative consequences of one's actions (e.g., by punishment). Thus, such individuals should engage more frequently in externalizing behavior. Here, we showed for the first time that externalization is associated with altered HPA axis responsivity not only in the pathological but also in the sub-clinical range. Although highly speculative, HPA axis dysregulation could, at least, serve as a potential sensitive marker or even operate as risk factor for the development of externalizing problems on a clinically relevant level.

In the present study, men and women differed significantly in their salivary cortisol response to acute stress with men showing larger cortisol responses, a generally well-documented effect (Zänkert et al., 2019). It is assumed that gender differences are, for example, attributable to sexual dimorphisms in brain functioning (e.g., corticolimbic system), circulating gonadal steroids (e.g., estradiol, testosterone) and/or gender-specific interpretations of stressors (Kudielka & Kirschbaum, 2005). Based on this, it is not surprising that we also found a moderating influence of gender on externalization-specific HPA axis responses. A comparable results pattern was found in the study by O'Leary et al. (2007) examining gender differences in the association between stress-related cortisol responses and psychopathic personality traits in the general population. While women showed non-significant cortisol increases regardless of psychopathy, men with low psychopathic traits responded significantly higher to the TSST compared to men with high psychopathy. Potential candidates for explaining these gender differences in externalization-specific HPA axis reactivity might be differences in brain functioning or the influence of androgens, especially testosterone, at different levels of the HPA axis (O'Leary et al., 2007). Beyond such biological reasoning, it might also be possible that the much lower cortisol responses to stress in our woman subsample made it much more unlikely to detect potential group

differences in cortisol responses related to externalization in women than men. Thus, considering the observed gender effects in the present study and given the paucity of data on women with externalizing behavior, we recommend to focus more on woman samples in future investigations.

It was speculated earlier that testosterone acts as a potential biomarker regarding the externalizing spectrum (Pajer et al., 2006; Yildirim & Derksen, 2012). However, against our hypothesis, our current data revealed no difference in salivary testosterone concentrations between high and low externalizing participants, beyond the expected gender difference in testosterone levels. This is in line with a recent meta-analysis that reported only limited evidence for testosterone effects (Dekkers et al., 2019). The absence of such an effect could be due to the fact that we studied healthy participants showing externalizing behaviors within the non-clinical range. Thus, it might be speculated that differences in testosterone levels manifest only with larger deviations or only in the pathological range.

In respect to autonomic functioning, we did not observe any differences in stress-induced HR responses related to externalization. This appears to be in contrast to earlier studies that reported lower cardiovascular stress responses in clinical samples within the externalizing spectrum (McBurnett et al., 2005; Snoek et al., 2004). However, empirical evidence so far is rather inconsistent (Van Goozen et al., 1998).

At the affective experience level, exposure to Scan*STRESS* increased negative affect and declined positive affect. The current data shows that healthy participants with high externalizing behaviors showed a constantly lower positive affective state while the negative affect was comparable to participants with low externalization. This could be alluding to the hypothesis that participants with externalizing behavior in the subclinical range might be in particular characterized by a generally lowered positive affect while overreaching affective responses to acute stress might be (only) symptomatic for pathological externalizing behavior (Bohnert, Crnic, & Lim, 2003; McLaughlin et al., 2011).

One explanation for the absence of an effect of externalization in whole brain level analysis contrasting stress and control might be our non-clinical sample. This assumption is supported by studies

investigating group differences in healthy nonclinical participants scoring high versus low on externalization (Buckholtz et al., 2010; Gordon, Baird, & End, 2004). The majority of these studies also reported no group differences on whole brain level, whereas ROI analyses revealed at least strong associations between psychopathy scores and activity in amygdala, prefrontal cortex, and functional connectivity in the striatum. In the current experiment, ROI analyses in the expected regions of interest, however, likewise disclosed no significant difference between externalizing groups in neural activity.

To identify brain regions associated with the observed reduced HPA axis responses to acute psychosocial stress in high externalizing participants, we additionally entered the variable cortisol increase to the statistical models. Results indicate that the relationship between cortisol increase and neural activity including the putamen and nucleus caudatus (dorsal striatum) differed between externalization groups. *Post-hoc* ROI analyses revealed that this relationship was negative in participants with high externalization, as verified by negative correlations between beta estimates and cortisol increases. In contrast, respective correlations remained nonsignificant in participants with low externalization. This points to the idea that the effects are mainly driven by high externalization. It appears relevant that externalization (in particular trait impulsivity), also within the non-clinical range, was previously found to be associated with a hypo-responsive mesolimbic dopamine system (e.g., nucleus accumbens, ventral regions of the nucleus caudatus, putamen) as well as an inappropriate modulation of mesolimbic pathways by prefrontal areas (Beauchaine et al., 2017). Indeed, earlier studies provided some evidence that the mesolimbic dopamine system is sensitive to environmental adversity during development (Gatzke-Kopp, 2011), which might explain alterations in its functioning related to high externalization. Moreover, a positive correlation between cortisol stress responses and dopamine release in the mesolimbic pathway, especially the nucleus accumbens, could be observed (Oswald et al., 2005; Pruessner, Champagne, Meaney, & Dagher, 2004). Following from here, it could be assumed that the observed lower HPA axis response in the high externalization group is associated with a cortisol-related altered dorsal striatum activation as part of the reactive mesolimbic system. *Post-hoc* analyses revealed that this 'externalization by cortisol increase' effect in the dorsal striatum could

mainly be found in women. This is supported by studies suggesting a gender-specific neural activation network underlying the central stress response, especially regarding the involvement of the limbic system and prefrontal cortex (Dedovic, Duchesne, Andrews, Engert, & Pruessner, 2009; Wang et al., 2007). As shown recently, the stress response in the dorsal striatum seems to be greater in women than men (Goldfarb, Seo, & Sinha, 2019) while evidence regarding the direction is rather inconsistent (Kogler, Gur, & Derntl, 2015). Taken together, the man-specific HPA axis hyporeactivity in the high externalization group and the altered cortisol-related striatum responsivity in high externalizing women appear to be a complex pattern, in the first place. Potentially, due to the generally lower cortisol response in women, we could not detect a significant externalization effect. However, at least a descriptive difference between externalization groups in women could be observed (see Table 4), predicting cortisol-related activation in dorsal striatum known to be involved in woman stress regulation.

Based on these different lines of research, one might raise the question whether a phenotypic characteristic on the behavioral level, namely high versus low externalization, could be linked to the observed differences in HPA axis and striatal functioning. In this context, studies should be acknowledged that suggest a reallocation of neural network activity towards the salience network in response to acute stress, resulting in enhanced automatic and habitual responses mediated by the dorsal striatum (Vogel et al., 2015). It is assumed that this shift is coordinated, among others, by cortisol (Schwabe, Tegenthoff, Höffken, & Wolf, 2013). The difference in the neural network shift between high and low externalizing participants consisting of an altered cortisol-related neural activity in the dorsal striatum could result in a maladaptive recruitment of resources in high externalizing participants. Such resources might be essential for an appropriate response to acute stress and (un)availability could reflect individual deviations on the behavioral and emotional level (e.g., impulsivity, emotional reactivity). Taken together, this might indicate an insufficient neural network shift to salience reflecting a state of underarousal as proposed by sensation-seeking and low-fear theories.

Several limitations should be taken in account for the current study. First, we only included university students resulting in a limited generalizability of our results. Second, we can also not preclude a potential impact of any physical or psychological inconvenience experienced by participants inside the scanner on our data. However, we attempted to catch such possible unsystematic effects by the within block design of the Scan*STRESS* paradigm. Third, due to our strict inclusion criteria (e.g., high vs. low externalization, MR-compatibility, health status) other potentially confounding variables like use of contraceptives (treated as covariates) could not be kept constant leading to a partly heterogeneous sample.

3.4.1 Conclusions

In sum, the applied Scan*STRESS* paradigm reliably induced acute stress responses in terms of cortisol, HR and affective responses. As a main finding, we observed that high externalization, though still in the non-clinical range, came along with reduced acute cortisol stress responses as well as a generally lower positive affective state. This finding might provide further evidence for the assumption that externalization is a dimensionally distributed characteristic ranging from variations within the non-clinical range to more extreme extents of externalization (e.g., antisocial personality disorder, conduct disorder, and substance (mis)use). During psychosocial stress exposure, no differences in neural activity between participants with high versus low externalization could be observed. However, in the high externalization group, cortisol increases correlated negatively with dorsal striatum activity. This observation raises the idea that differences in the neural network shift mediated by the dorsal striatum might be associated with externalizing behavior via stress-related HPA axis regulation. Additionally, a modulating influence of gender was disclosed in HPA axis regulation showing greater externalization effects in men. In contrast, on the neural level results indicated altered cortisol-related dorsal striatum activity notable in high externalizing women, which did not lead to a significant 'externalization' effect in cortisol responses to stress.

Chapter IV:

Externalization and reactive aggression

4.1 Introduction

The externalizing spectrum model was originally developed for explaining the coincidence of a set of heterogeneous personality traits and behavioral patterns encompassing antisocial behavior, disinhibition, and substance (mis)use (Krueger et al., 2002; Krueger et al., 2007; Patrick et al., 2013). However, empirical work has shown that externalization is a dimensional characteristic, distributed across the general population (Krueger et al., 2007; Markon & Krueger, 2005). For example, the externalizing spectrum comprises variation in trait impulsivity as a personality characteristic as well as more extreme, and clinically-relevant, behavioral expressions which are indicative for, e.g., ADHD, CD, ASPD, and SUD (Zisner & Beauchaine, 2016). In particular, high trait impulsivity represents a risk factor for the development of externalizing disorders (Beauchaine et al., 2017). Furthermore, individuals exhibiting high externalization often show emotional hyperreactivity as well as aggressive behavior (Bohnert, Crnic, & Lim, 2003; McLaughlin et al., 2011). However, most studies reporting increased aggressive behavior in the context of externalization were derived from clinical populations while evidence from the general population is still scarce (Brislin et al., 2019; White et al., 2013).

A well-established laboratory measure for behavioral aggression, in particular reactive aggression, is the TAP. In the original version by Taylor (1967), participants played a fictitious reaction time task against a mock opponent with default win and lose trials. In win trials, the participant was instructed to administer an electric shock to the mock opponent. Thereby, the selected intensities served as indicator for reactive aggression as provoked by respective punishment selections of the mock opponent in previous lose trials. Over the last decades, this paradigm was modified quantitatively (e.g., in terms of provocation, duration, number of trials) and qualitatively (e.g., stimuli; Elson et al., 2014). Meanwhile, a mTAP was established using monetary stimuli (subtraction of money from a fictitious account) instead of electric shocks, noise or heat stimuli (Kogan-Goloborodko et al., 2016; Konzok et al., 2020; Wagels et al., 2018; Weidler et al., 2019). Recent research supports the validity of the monetarily modified TAP by showing a dose-response effect of preceding provocation on the behavioral responses by the participant (Repple et al., 2017; Schneider et al., 2015) as well as by

showing correlations with other measures of reactive aggression (e.g., self-reported aggression by questionnaire) (Konzok et al., 2020; Weidler et al., 2018). Furthermore, moderating effects of other factors known to influence reactive aggression (e.g., gender) was shown empirically (Konzok et al., 2020; Weidler et al., 2019). Finally, recent studies applied for the first time trial-by-trial analysis in order to account for individual aggression trajectories using multilevel approaches (Chester, 2019; Konzok et al., 2020; Wagels et al., 2018).

Investigating the neurobiology of retaliation, results from neuroimaging studies using mTAPs with noise, thermal or pneumatic stimuli indicated that reactive aggression is associated with altered activity in medial prefrontal cortex (mPFC), the OFC, and superior temporal gyrus reflecting cognitive control processes (for review see Fanning et al., 2017). Increased provocation by the mock opponent induced greater activation in the amygdala, insula, ACC, thalamus, and OFC (Buades-Rotger et al., 2016; Krämer et al., 2007; Lotze et al., 2007). Particularly interesting are recent findings with the monetary mTAP showing that the mPFC, posterior parts of the superior and middle frontal gyrus as well as cingulate cortex (ACC, middle cingulate cortex), and insula are activated during the active selection of a punishment (reactive aggression). In contrast to the neural activity during reactive aggressive behavior, observing the subtraction of money from the own account by the mock opponent (provocation) was related to activity in the ACC, thalamus, nucleus caudatus, mPFC, and insula (Repple et al., 2017; Wagels et al., 2019; Weidler et al., 2018). In sum, findings from mTAP-studies with different types of stimuli suggest that the ACC, mPFC, and OFC play a key role in both, provocation processing and active aggression, while insula and amygdala activity rather reflects the experience of negative emotions during provocation.

Examining neural circuits mediating abnormal reactive aggression, the majority of studies focused on psychiatric disorders for which aggression is a main diagnostic criterion including externalizing disorders like ASPD and CD. These studies suggest that reactive aggressive behavior is associated with abnormalities in three neural systems implicated in the experience of aggression, decision making, and regulation of emotions (for review see Coccaro et al., 2011). Especially in externalizing disorders,

neuroimaging studies revealed reduced neural activity during reward as well as emotion processing (including acute threat) and (poor) decision making in the OFC, ACC, striatum, amygdala, and insula as shown in patients with CD (for review see Fairchild et al., 2019), ADHD (Plichta et al., 2009), SUD (Koob & Volkow, 2010) and antisocial behavioral tendencies (Oberlin et al., 2012). These dysfunctions are hypothesized to enhance the probability for impulsive behavior and reactive aggression (Blair, Veroude, & Buitelaar, 2018). Furthermore, in a social provocation paradigm, individuals exhibiting disruptive behavior problems and low callous-unemotional traits are associated with reduced OFC activity during retaliation and blunted amygdala-OFC-connectivity during high provocation compared to healthy controls (White et al., 2016). With this, the question arises whether externalization in the non-clinical normal range might also be related to altered neural processing of provocation and reactive aggression. To this point, only one study revealed aberrant activity in threat and reward systems (e.g., amygdala, nucleus accumbens) in response to (un)pleasant pictures associated with trait disinhibition within a healthy sample. However, the authors did not explicitly control for psychiatric disorders (Foell et al., 2016).

Another prime candidate for explaining potential variance in reactive aggression in respect to externalization is testosterone. From a theoretical point of view, situations challenging social status evoke testosterone secretion which, in turn, leads to behavior obtaining interpersonal standing, including aggression (Dekkers et al., 2019). To date, extensive research has been conducted to deliver empirical evidence. Though, meta-analyses demonstrated only a modest correlation between testosterone concentrations and aggressive behavior and rather highlighted the moderating role of other factors like gender (Archer, 2006). However, using laboratory paradigms, several studies revealed a positive relationship between testosterone levels and aggressive behavior in men with the TAP (Berman, Gladue, & Taylor, 1993; Denson, Mehta, & Tan, 2013), the Point Subtraction Aggression Paradigm (Carré & McCormick, 2008; Geniole, Carré, & McCormick, 2011) and the Ultimatum Game (Mehta & Beer, 2010). In women, findings are much more inconsistent showing positive (Archer, 2006; Mehta & Beer, 2010), negative (Buades-Rotger et al., 2016) or no associations (Carré & McCormick,

2008). According to the dual-hormone hypothesis, the relationship between testosterone and aggression is thought to be potentially modulated by cortisol. However, a recent meta-analysis provided only marginal support for this interaction (Dekkers et al., 2019). In respect to testosterone levels in externalization, we suspect that it might be of relevance if subjects are healthy or belong to a psychopathological, delinquent population (Rowe, Maughan, Worthman, Costello, & Angold, 2004) or exhibit personality traits like psychopathy (Welker, Lozoya, Campbell, Neumann, & Carré, 2014). For instance, in our own data within this research project (see below), we revealed no group differences in salivary testosterone concentrations between healthy participants exhibiting high versus low externalizing behavior during the Scan*STRESS* paradigm (study 2) while earlier studies raised the idea of higher testosterone concentrations in clinically-relevant externalization (Dmitrieva, Oades, Hauffa, & Eggers, 2001). On the neural level, testosterone seems to increase activity in threat and reward systems including the amygdala (Hermans, Ramsey, & van Honk, 2008) and striatum (for a review see Welker, Gruber, & Mehta, 2015) and suppresses, among others, OFC activation (Mehta & Beer, 2010) which in sum facilitates aggressive behavior (Coccaro et al., 2011). To the best of our knowledge, there is only one study in healthy participants investigating gender-specific relationships between aggression and endogenous testosterone as well as neural correlates using a laboratory aggression paradigm (Buades-Rotger et al., 2016). Thus, the role of testosterone in high versus low externalization in a healthy population has not been examined, especially applying a multilevel approach accounting for individual trajectories on the behavioral as well as neural level.

4.1.1 Research questions

Our main research aim was to investigate the relationship between externalization and reactive aggressive behavior and its neural correlates in a non-clinical, healthy sample.

First, we expected a dose-response effect of provocation on reactive aggression in the monetary mTAP accompanied by activations in the neural inhibition network including the ACC as well as OFC. Limbic and mesolimbic systems including amygdala, insula, and ACC should be involved during higher levels of provocation.

In participants scoring high on externalization, we assumed higher aggression levels, especially after higher levels of provocation. Furthermore, we presumed enhanced negative affect as well as reduced positive affect in high compared to low externalizing participants in response to the monetary mTAP.

According to findings from externalizing pathologies, healthy participants with high compared to low externalization were expected to manifest reduced activation in inhibitory control areas (ACC, OFC) during active aggressive behavior and blunted (meso)limbic activity while being provoked by the opponent as well as in response to the outcome (win vs. lose).

Owing to well-known gender differences in behavioral aggression as well as testosterone concentrations, we finally assumed a gender-specific relationship between testosterone levels and behavioral reactive aggression in the monetary mTAP, potentially connected to cortisol secretion. Additionally, neural correlates will be explored in the relationship between testosterone levels and aggression-related activation in men and women with high versus low externalization.

4.2 Methods

4.2.1 Participants

Sixty-three preselected subjects (age: M = 23.62, SD = 3.81, range 18 – 34; 31 men, 32 women) from the higher (n = 32) versus lower (n = 31) range of the normal variation in externalization were tested twice in the MRI scanner. The assignment to one of the two externalization groups was based on the scores in the TriPM (Patrick, 2010; Patrick, Fowles, & Krueger, 2009) within the scopes of the highest (Q_{75}) or lowest quartile (Q_{25}) of the subscales disinhibition (Q_{75} = 36, Q_{25} = 29) and meanness (Q_{75} = 33, Q_{25} = 26) as derived from a large online assessment in the general population (Eisenbarth, Castellino, Alpers, Kirsch, & Flor, 2012). Volunteers scoring high on the personality trait psychopathy were not eligible, i.e., volunteers scoring within the upper quartile of the TriPM subscale boldness (Q_{75} = 55) were not selected.

All participants reported to be free of acute and chronic illness, mental, and psychiatric disorders as assessed with the German version of the SCID-I and SCID-II (Wittchen et al., 1997), criminal history,

current use of drugs and medication with glucocorticoids as well as MRI scanner contraindications. Owing to poor image acquisition and missing saliva samples, two subjects were excluded from fMRI analyses rendering a final sample of N = 61. All subjects gave written informed consent and were compensated with 100 € or course credits. This experiment was approved by the ethics committee of the University of Regensburg.

4.2.2 Procedure

The current research project consisted of two fMRI sessions on two different days. At the first scanning session, neural stress responses were investigated applying the Scan*STRESS* paradigm (reported in Konzok et al., under review; study 2). The second scanning session (reported here) comprised the monetary mTAP. Participants arrived at least 30 min before the beginning of the scanning session (starting after 12 p.m.). In order to increase the credibility of the cover story, subjects were introduced to a confederate of the same sex and informed that they will act as opponents in a competitive reaction time task while being in adjacent scanner rooms. After a monetary mTAP training session consisting of five trials, participants were transferred into the scanner room and passed a resting state (18 min) and DTI (22 min) sequence (results not presented in the present manuscript). Subsequently, we conducted the monetary mTAP (see below). After the scanning session, participants filled in an in-house deception check questionnaire to identify participants who expressed suspicion regarding the cover story (see Konzok et al., 2020; study 1).

During the experimental session, saliva was collected by Cortisol Salivettes (Sarstedt, Nümbrecht, Germany) at eight time points: -75 min, -55 min, -1 min before the start of the monetary mTAP as well as +1 min, +10 min, +20 min, +30 min, and +45 min (C1 – C8) thereafter. At each saliva sampling point, subjects completed the state version of the Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988) or answered it per hand signal while lying in the scanner. For testosterone analysis, we collected native saliva by passive drool using polypropylene tubes at three time points: -55 min before the start of the monetary mTAP and +1 min as well as +30 min thereafter (pooled for analysis).

One week after the scanning session, participants received a link to an online assessment including a trait aggression questionnaire (K-FAF) (Heubrock & Petermann, 2008), the Buss Perry Aggression Questionnaire (Buss & Perry, 1992), the Brown-Goodwin Lifetime History of Aggression Scale (BGHA; Brown, Goodwin, Ballenger, Goyer, & Major, 1979), and the Reactive Proactive Questionnaire (RPQ; Raine et al., 2006).

4.2.3 Monetary modified Taylor Aggression Paradigm

The monetary mTAP is a mock competitive reaction time task with a fictional opponent. Each trial consists of a decision phase, the reaction time task itself and a feedback phase. At the beginning of each trial (decision phase), the participant has to set a stake between 0 and 90 euro cents on the computer with a given setting start point of 45 cents for each trial. The setting start point cannot be chosen as final stake. Following the recommendations by Tedeschi and Felson (1994), the amount of 0 cents is included as nonaggressive option (see also Elson et al., 2014). For the reaction time task, participants are instructed to press a button as quickly as possible as soon as a green circle appears on the computer screen. After reaction times slower than 500 ms, the feedback: "You haven't pressed a button. Please react as quickly as possible" is presented. In case the participant presses the button before the green circle appears the trial is repeated. It is emphasized that the selected amount will be subtracted from the opponent in case the opponent loses the reaction time task. If the participant loses the trial, the punishment level selected by the fictional opponent will be delivered to the participant (feedback phase). In lose trials, the level of provocation is defined by the amount of subtracted money by the opponent. In case of winning, the participant always receives 50 cents (see Figure 4). While performing the task, participants do not receive any feedback of their current account balance (see also Konzok et al., 2020; study 1).

Participants performed 100 randomly presented trials, with preprogrammed 40 win and 60 lose trials. In accordance with earlier work (Konzok et al., 2020), we categorized the level of provocation (amount of money subtracted by the mock opponent in the preceding lose trial) into low (0 - 20 cents),

medium (30 - 60 cents) and high (70 - 90 cents). The unprovoked first trial was omitted from the analyses.

4.2.4 Materials, biochemical analysis, and data acquisition

The monetary mTAP was presented on a monitor via a stimulus computer using the software Presentation (Version 19.0; Neurobehavioral Systems, San Francisco, CA, USA). After test sessions, saliva samples were stored at -20 °C. Analyses were performed by the biochemical laboratory at the University of Trier, Germany. Cortisol was assayed in duplicate using a DELFIA; testosterone was assessed by a commercially available assay kit (Demeditec Diagnostics GmbH, Kiel, Germany). Inter- and intra-assay coefficients of variation were below 10%, respectively.

To collect fMRI data, we used a MAGNETOM Siemens 3T Prisma scanner (Siemens AG; Erlangen, Germany) and a 64-channel head coil. Between scalp and head coil foam paddings were positioned to prevent extensive movement. A T2*-weighted EPI sequence (TR = 2000 ms; TE = 30 ms, flip angle = 90°, FOV = 192 x 192 mm^2, matrix size = 64 x 64 mm^2, 37 slices, slice thickness = 3.0 mm, 1.0 mm gap, voxel size = 3 x 3 x 3 mm^3, interleaved) was used to create functional scans and a T1-weighted MP-RAGE sequence (TR = 2400 ms, TE = 2.18 ms, flip angle = 9°, voxel size = 0.8 x 0.8 x 0.8 mm^3, distance factor: 50%) for structural scans.

4.2.5 Psychometric measures

To identify the two quasi-experimental groups, we used the three subscales boldness, meanness, and disinhibition of the TriPM (Patrick, 2010; Patrick et al., 2009) (see above). The four-point answering format ranges from 1 = *not true at all* to 4 = *completely true*. The PANAS (Watson et al., 1988) is composed of 20 items and measures positive and negative affect on a 5-point answering scale. Trait aggression was evaluated by the German K-FAF (Heubrock & Petermann, 2008) comprising 49 items pooled to five subscales (spontaneous aggression, reactive aggression, excitability, aggression inhibition and self-aggression) and the 28-item BPAQ (Buss & Perry, 1992), which includes four subscales (physical aggression, verbal aggression, anger, and hostility). The BGHA (Brown, Goodwin,

Ballenger, Goyer, & Major, 1979) measures aggressive behavior by eleven questions to be answered on a 4-point rating scale across three developmental stages of life (childhood, adolescence, and adulthood). The 23-item RPQ (Raine et al., 2006) consists of two subscales, namely proactive and reactive aggression.

4.2.6 Statistical analysis

To assess potential differences between the two externalization groups, independent Welch-test comparisons regarding demographic variables and aggression questionnaires were performed using R (version 3.5.1; R Core Team, 2018) with the packages afex (Singmann et al., 2020), lme4 (Bates et al., 2014), lmerTest (Kuznetsova et al., 2017), psych (Revelle, 2019) and sjstats (Lüdecke, 2020). For cortisol analysis, we computed the area under the curve with respect to the ground (AUCg) following the recommendations of Pruessner, Kirschbaum, Meinlschmid, and Hellhammer (2003).

According to an already validated evaluation strategy of the monetary mTAP with LMMs (Konzok et al., 2020; study 1), we analyzed participants' aggression in response to the amount of money subtracted by the mock opponent in the previous trial (provocation, low vs. medium vs. high) and added a random intercept by subject accounting for interindividual variability as well as a random slope for provocation by subject. The dummy coded variable provocation consisted of two estimates: Low provocation served as the reference standard to which the effects of medium and high provocation were related to. In a next step, the predictors of interest, externalization (high vs. low), *z*-transformed pooled testosterone levels, *z*-transformed cortisol levels (AUCg) as well as the interaction term *z*-transformed pooled testosterone by gender (women vs. men) and *z*-transformed pooled testosterone by *z*-transformed cortisol (AUCg) were included. Additionally, the model was enhanced by the factors gender as sole factor, trait reactive aggression assessed by the K-FAF and deception check ensuring correct variance allocations.

To evaluate the goodness of fit regarding the random and fixed effect structure, we built four different models. Model 1 only consisted of a random intercept for participant (null-model). Model 2 included a fixed effect for provocation. In model 3 and 4 a random slope for provocation by participant was added. In

addition, model 4 (final full model) contained also fixed effects for the predictors of interest and additional factors. All fixed effects were tested with an *F*-Tests type III using a Satterthwaite approximation of the degrees of freedom (Kenward & Roger, 1997).

A repeated measures analysis of variance (ANOVA) was performed for positive and negative affect (time [8 assessments] x externalization [low, high]). For this analysis, eight participants had to be removed because of missing time points.

4.2.7 Functional MRI data analysis.

Imaging analysis was performed with SPM12 (Wellcome Department of Imaging Neuroscience, University College London, London, UK). Data preprocessing started with the realignment by registering the scans to the mean images. Afterwards, slice timing was conducted with the first slice as reference. Then, functional mean images and the T1-weighted scan were coregistered. In the context of segmentation and normalization, functional and structural images were transformed into the standard space defined by the MNI. At the end of the preprocessing, we smoothed the functional images with an isotropic Gaussian kernel of 8 mm full-width-at-half-maximum. Additionally, a high-pass filter correction of 128 s was arranged removing low-frequency drifts.

Whole brain analysis. For each single subject (first level analysis), a GLM was fitted with the four regressors of interest (decision phase after high and low provocation, feedback phase with high and low provocation) as well as eleven regressors of no interest (decision phase after winning trials as well as winning feedback phase, the reaction time task, the six realignment parameters). To enhance statistical power, decision phase after medium provocation and feedback phase with medium provocation were also excluded from the analyses as regressors of no interest.

On the second level, we performed four factor-analysis of variance (full factorial design) for decision and feedback phase, respectively one model in response to provocation (provoked trials only: preceding trial was lost) and one model as function of outcome (preceding winning and losing trial). The four models included the factors previous provocation (low vs. high), respectively previous outcome (won vs. lost) and externalization (low vs. high). Thresholded family-wise error was corrected at $p = .05$, *F*-

contrasts for main effect of provocation respectively outcome and externalization as well as the interaction effects were conducted for decision and feedback phase.

ROI analysis. For *post-hoc* ROI analyses, masks in the ACC, OFC, amygdala, insula, nucleus accumbens, nucleus caudatus, and putamen were created using the SPM Anatomy toolbox (Eickhoff et al., 2005). Average beta estimates were extracted employing the MarsBaR toolbox for SPM (http://marsbar.sourceforge.net/). The parameter estimates entered repeated ANOVAs, Bonferroni corrected at $p = .05$.

Parametric modulation. Exploring activations in brain regions covarying with the amount of money selected by the participants during decision phases, aggression-response-related BOLD analysis was conducted by building a second first level model for each participant. Further, differing decision phases in terms of the outcome of the previous trial (win > lost), this model included a parametric modulator for the amount of money subtracted by the participants in each trial. On the second level, we performed a two-sample *t*-test regarding the parametric modulator after losing trials (reactive aggression modulated neural response) with gender as group variable (women > men, men > women) and *z*-transformed testosterone concentrations as continuous covariate. Additionally, high and low externalizing participants were contrasted regarding the aggression-response-related activation pattern (high > low, low > high). Due to missing variance in aggressive response selection ($SD = 0$), five participants had to be removed from these analyses ($N = 56$). For this explorative analysis, an original threshold was set at $p = .001$ and family-wise error was corrected on cluster level at $p = .05$. For anatomy localization, the SPM Anatomy toolbox (Eickhoff et al., 2005) was used. *Post-hoc* ROI analyses were conducted within the significant cluster from the previous contrast.

4.3 Results

4.3.1 Descriptive results

Consistent with phenotypic characteristic of externalization, the high externalization group showed significant greater values in the disinhibition and meanness scales of the TriPM compared to the low

externalization group. Significant group differences also emerged in four out of five scales of the K-FAF (except aggression inhibition), the five scales of the BPAQ, the two scales of the RPQ and the adolescence scale of the BGHA (see Table 5). The aggression inhibition subscale of the K-FAF showed that low externalizing participants scored significantly higher. The deception check questionnaire revealed that 21 participants (33%) reported at least some suspicion regarding the cover story.

4.3.2 Behavioral and affective responses

Within the context of model building for behavioral aggression levels, inclusion of provocation as fixed effect (model 1 vs. model 2; $\Delta\chi^2(df = 2) = 135.80$, $p < .001$, $\Delta BIC = 119.41$) and as random effect (model 2 vs. model 3; $\Delta\chi^2(df = 5) = 181.18$, $p < .001$, $\Delta BIC = 140.13$) resulted in an improved model fit using likelihood ratio tests, indicating a significant impact of provocation on behavioral responses allover (fixed effect) as well as on individual level (random effect). Comparing model 3 and model 4 including all predictors as fixed effects, we also observed an improvement of the goodness fit (model 3 vs. model 4; $\Delta\chi^2(df = 12) = 75.14$, $p < .001$, $\Delta BIC = 51.14$). For detailed information regarding the different models, we refer to Table 6.

Evaluating fixed effects of the final model 4, Table 7 depicts *t*-tests for individual fixed-effect estimates. *F*-tests (Satterthwaite approximation of the degrees of freedom) for overall effects of fixed factors confirmed a main effect of provocation ($F(2, 108.75) = 12.34$, $p < .001$, $\eta_p^2 = .63$), though no main effect externalization ($F(1, 51.56) = 0.12$, $p = .733$, $\eta_p^2 < .01$) nor interaction of provocation and externalization ($F(2, 108.75) = 0.51$, $p = .602$, $\eta_p^2 = .03$) (see Figure 10A). Beyond this, the interaction gender by testosterone ($F(1, 51.04) = 5.48$, $p = .023$, $\eta_p^2 = .14$) reached significance, however, this does not apply to the main effect testosterone ($F(1, 55.03) = 0.06$, $p = .801$, $\eta_p^2 = .01$) nor cortisol ($F(1, 51.03) = 0.20$, $p = .731$, $\eta_p^2 = .01$). Owing to variational cases in testosterone concentrations, one female participant was identified with a value above the twofold *SD*. Nevertheless, when excluding this participant from analyses, the interaction gender by testosterone remained significant ($F(1, 50.05) = 4.99$, $p = .030$, $\eta_p^2 = .14$). None of the other interaction effects reached significance (Fs < 1.77, ps > .190, $\eta_p^2 < .05$). Thus, the significant

relationship between testosterone and behavioral aggression levels differed between men and women. In men, reactive aggression correlated negatively with salivary testosterone concentrations while this correlation was positive in women (see Figure 10B).

Table 5

Mean ± SD of demographic, behavioral, and psychometric variables and results of welch tests and Cohen's d contrasting the two externalization groups (study 3)

		High externalization		Low externalization					
		Men (n = 16)	Women (n = 16)	Men (n = 15)	Women (n = 16)	t	df	p-value	d
Age in yrs.		24.31 (± 3.38)	21.93 (± 4.17)	25.67 (± 4.34)	22.68 (± 2.21)	1.05	60.93	.298	-0.27
Pooled testosterone (pg/ml)		179.56 (± 57.43)	76.67 (± 46.87)	164.08 (± 61.15)	74.46 (± 24.17)	-0.68	58.91	.498	0.18
TriPM									
	Disinhibition	40.25 (± 4.57)	42.87 (± 5.88)	27.07 (± 1.62)	26.94 (± 2.05)	-14.56	38.34	**.001****	3.68
	Meanness	40.31 (± 5.45)	38.88 (± 4.03)	23.20 (± 1.86)	22.87 (± 1.82)	-18.31	40.04	**.001****	4.63
	Boldness	47.75 (± 5.52)	47.75 (± 5.51)	51.53 (± 2.97)	49.06 (± 3.96)	2.41	57.37	**.019***	-0.61
K-FAF									
	Spontaneous aggression	17.13 (± 8.89)	20.38 (± 7.63)	7.20 (± 6.21)	3.88 (± 2.55)	-7.74	50.57	**.001****	1.97
	Reactive aggression	27.44 (± 8.97)	28.38 (± 8.39)	15.80 (± 7.15)	11.69 (± 4.69)	-7.55	56.81	**.001****	1.92
	Excitability	20.19 (± 10.20)	24.38 (± 8.19)	11.33 (± 6.99)	11.88 (± 5.73)	-5.34	54.35	**.001****	1.36
	Self-aggression	18.06 (± 10.12)	23.88 (± 9.45)	8.87 (± 5.37)	8.25 (± 5.27)	-6.17	46.95	**.001****	1.56
	Inhibition	17.94 (± 5.99)	16.44 (± 5.19)	20.66 (± 4.89)	21.81 (± 4.10)	3.21	58.97	**.002****	-0.82
BPAQ									
	Physical aggression	20.50 (± 6.65)	21.00 (± 8.18)	14.87 (± 2.50)	11.50 (± 2.58)	-5.42	41.54	**.001****	1.37
	Verbal aggression	13.50 (± 3.16)	14.25 (± 4.40)	11.07 (± 2.46)	10.75 (± 1.61)	-3.89	47.87	**.001****	0.99
	Anger	18.31 (± 4.76)	20.38 (± 5.98)	13.80 (± 4.06)	12.88 (± 2.96)	-5.25	53.31	**.001****	1.34
	Hostility	19.81 (± 6.63)	22.13 (± 6.13)	14.20 (± 5.16)	13.13 (± 4.51)	-5.16	57.41	**.001****	1.32
	Sum	72.13 (± 14.34)	77.75 (± 21.28)	53.93 (± 9.03)	48.25 (± 8.11)	-6.70	45.52	**.001****	1.70
BGHA									
	Childhood	17.31 (± 5.10)	14.63 (± 2.83)	14.60 (± 5.11)	12.13 (± 1.86)	-2.56	60.84	**.013***	0.65
	Adolescence	20.19 (± 4.02)	18.38 (± 5.06)	15.20 (± 4.23)	13.94 (± 1.81)	-4.75	55.70	**.001****	1.21
	Adulthood	15.38 (± 3.56)	13.63 (± 2.22)	13.07 (± 2.22)	12.38 (± 2.19)	-2.68	56.37	**.010***	0.68
RPQ									
	Proactive aggression	13.13 (± 1.45)	14.00 (± 2.48)	12.27 (± 0.46)	12.19 (± 0.40)	-3.61	33.75	**.001****	0.91
	Reactive aggression	16.94 (± 4.12)	18.75 (± 3.70)	14.47 (± 3.00)	14.56 (± 1.50)	-4.09	50.17	**.001****	1.04

Notes: yrs. = years, pg/ml = picogram per milliliter. ** $p < .01$, * $p < .05$.

Table 6

Four LMMs with reactive aggression as dependent variable in the monetary mTAP (study 3)

	Model 1	Model 2	Model 3	Model 4
Intercept	**51.35 *****	**46.97 *****	**47.00 *****	**56.92*****
	(2.44)	(2.48)	(2.58)	(5.48)
Provocation (medium)		**4.18 *****	**4.15 *****	3.04
		(0.77)	(1.10)	(1.56)
Provocation (high)		**9.02 *****	**8.99 *****	**7.44*****
		(0.77)	(1.81)	(2.57)
Externalization (0=low, 1=high)				-4.12
				(7.02)
Provocation (medium) x externalization				2.22
				(2.20)
Provocation (high) x externalization				3.09
				(3.63)
Cortisol (normalized)				-6.60
				(5.17)
Testosterone (normalized)				**-12.00****
				(4.51)
Testosterone x gender				**35.44***
				(15.14)
Cortisol x testosterone				6.82
				(4.01)
Cortisol x gender (0=man, 1=woman)				20.33
				(16.09)
Cortisol x testosterone x gender				10.40
				(17.93)
Gender				12.27
				(13.13)
Deception check (0=no suspicion, 1=suspicion)				-4.28
				(5.16)
K-FAF trait reactive aggression (normalized)				4.89
				(3.45)
ICC: participant	.49			
AIC	32 402.65	32 270.83	32 099.66	32 048.52
BIC	32 421.27	32 301.86	32 161.73	32 185.06
Log Likelihood	-16 198.32	-16 130.42	-16 039.83	-16 002.26
Num. obs.	3664	3664	3664	3664
Num. groups: participant	62	62	62	62
Var: participant (Intercept)	363.55	363.71	396.95	357.96
Var: residual	378.66	365.09	336.87	336.93
Var: provocation (medium)			41.22	40.91
Var: provocation (high)			169.12	169.95
Cov: participant (Intercept) provocation (medium)			-44.47	-43.56
Cov: participant (Intercept) provocation (high)			-67.10	-65.89
Cov: participant (Intercept) provocation (medium and high)			83.13	82.99
Marginal R^2		.02	.02	.13
Conditional R^2	.49	.50	.58	.58

Notes. Model 1 only consists of a random intercept for participant (null-model); model 2 includes a fixed effect for provocation; in model 3 and 4 a random slope for provocation by participant is added; in addition, model 4 (final full model) contains also fixed effects for externalization, gender, pooled testosterone, deception check and the trait reactive aggression scale of the K-FAF; *** $p < .001$, ** $p<.01$, * $p < .05$.

Table 7

Parameter estimates for the final model (model 4) with reactive aggression as dependent variable in the monetary mTAP (study 3)

	beta	SEM	*t*	*p*-value
Intercept	56.92	5.48	10.39	**.001**
Provocation (medium)	3.04	1.56	1.95	.055
Provocation (high)	7.45	2.57	2.90	**.005**
Externalization (0=low, 1=high)	-4.12	7.02	-0.59	.559
Provocation (medium) x externalization	2.22	2.20	1.01	.317
Provocation (high) x externalization	3.09	3.63	0.85	.398
Cortisol AUCg (normalized)	-6.60	5.18	-1.28	.208
Testosterone (normalized)	-12.00	4.51	-2.66	**.010**
Testosterone x gender	35.44	15.14	2.34	**.023**
Cortisol x testosterone	6.82	4.01	1.70	.095
Cortisol x gender (0=man, 1=woman)	20.34	16.09	1.26	.212
Cortisol x testosterone x gender	10.40	17.93	0.58	.564
Gender	12.28	13.13	0.94	.354
Deception check (0=no suspicion, 1=suspicion)	4.89	3.45	1.42	.162
K-FAF trait reactive aggression (normalized)	-4.28	5.16	-0.83	.411

The additional factors gender ($F(1, 51.04) = 0.87$, $p = .354$, $\eta_p^2 = .02$), deception check ($F(1, 51.04) = 0.69$, $p = .411$, $\eta_p^2 = .02$) and *z*-normalized trait reactive aggression assessed by the K-FAF ($F(1, 51.04) = 2.02$, $p = .161$, $\eta_p^2 = .05$) manifested no predictive value for the aggression levels in the laboratory paradigm. Since externalization and trait reactive aggression appeared to be no independent factors (see Table 5), we subsequently excluded externalization from the analysis revealing a marginally significant main effect of trait reactive aggression ($F(1, 52.01) = 2.97$, $p = .091$, $\eta_p^2 = .08$). Conversely, externalization surpassed clearly the significance level ($F(1, 52.03) = 0.97$, $p = .330$, $\eta_p^2 = .02$) when removing trait reactive aggression.

PANAS scores for positive affect showed a significant decline over time ($F(3.74, 198.37) = 8.78$, $p < .001$, $\eta_p^2 = .14$) and a significant externalization effect ($F(1, 53) = 5.70$, $p = .021$, $\eta_p^2 = .10$) exhibiting a higher positive affective state in the low externalization group but no interaction effect ($F(3.74, 198.37) = 1.39$, $p = .240$, $\eta_p^2 = .03$). For negative affect, we found a significant time effect ($F(4.26, 225.55) = 10.88$, $p < .001$, $\eta_p^2 = .17$) indicating a decrease over the experimental session, however, no

group differences ($F(1, 53) = 0.11$, $p = .744$, $\eta_p^2 = .00$) nor interaction effect ($F(4.26, 225.55) = 1.18$, $p = .319$, $\eta_p^2 = .02$) emerged.

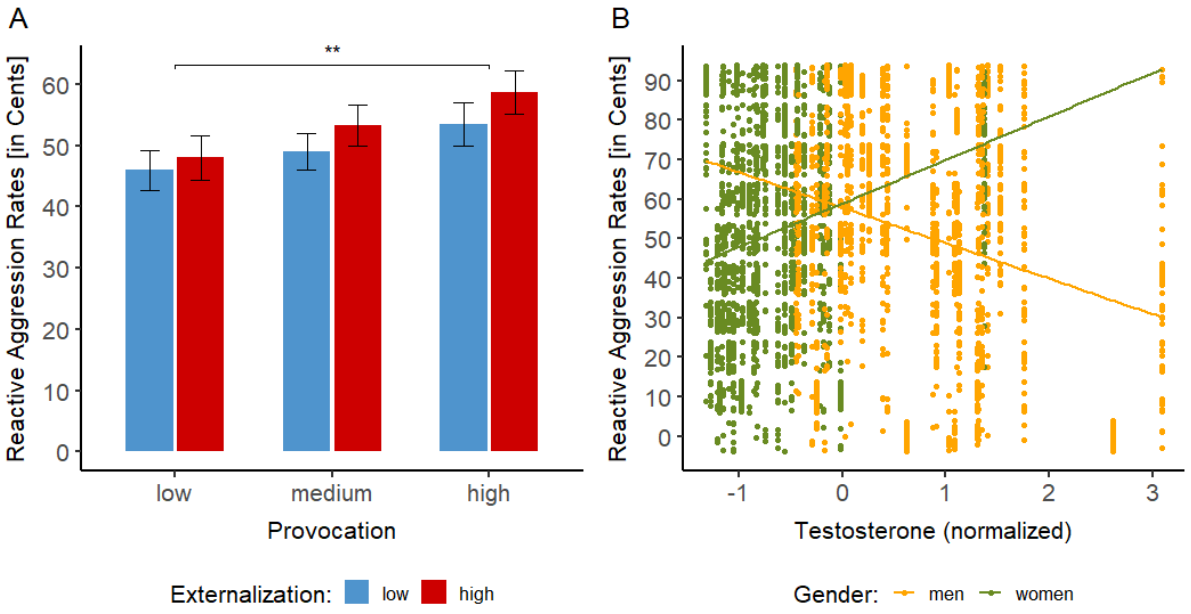

Figure 10. Mean (± *SEM*) aggression levels in response to the provocation of the fictional opponent, separated for the low and high externalization group (A). Linear relationship between aggression and *z*-normalized testosterone levels differing between men and women (interaction effect gender by testosterone; 60 data points per participant) (B). ** $p < .001$.

4.3.3 Neural results

Decision phase. Based on provoked trials (i.e., the preceding trial was lost), whole brain results for the decision phase (reflecting behavioral reactive aggression) revealed no suprathreshold cluster ($k > 20$) for the main effects of provocation and externalization nor the interaction provocation by externalization.

Post-hoc ROI analyses with repeated measure ANOVAs revealed a significant effect of the previous provocation (low vs. high, $F(1, 60) = 6.46$, $p = .014$, $\eta_p^2 = .10$) on beta estimates and the interaction provocation by externalization ($F(1, 60) = 7.65$, $p = .008$, $\eta_p^2 = .11$) in the rostral part of the ACC during the decision phase, surviving Bonferroni correction for multiple comparisons ($n = 2$). Activation peaks for this provocation effect are located in the pregenual region and for the interaction

effect in the subgenual region of the rostral ACC (rACC; see Figure 11A and B). The paired comparison of bilateral ACC beta estimates after high and low provocation showed a significant result in the low externalization group ($t(29) = -2.98$, $p = .006$), but not in the high externalization group ($t(32) = 0.23$, $p = .820$). Moreover, differences in bilateral ACC beta estimates after the high compared to the low provocation condition ($M_{high\ provocation} - M_{low\ provocation}$) differed significantly between externalization groups, indicating higher differences in the low externalization group ($M_{high\ ex} = -0.01$, $M_{low\ ex} = 0.19$, $t(45.34) = 2.72$, $p = .009$). In the OFC, there was no difference between low and high previous provocation ($F(1, 60) = 2.79$, $p = .100$, $\eta_p^2 = .04$), but a significant interaction effect provocation by externalization emerged which, however, did not survive Bonferroni correction ($F(1, 60) = 4.60$, $p = .036$, $\eta_p^2 = .07$). Therefore, increased activation in the ACC after high compared to low provocation during the decision phase was greater in the low than high externalization group (see Figure 11B and 11C).

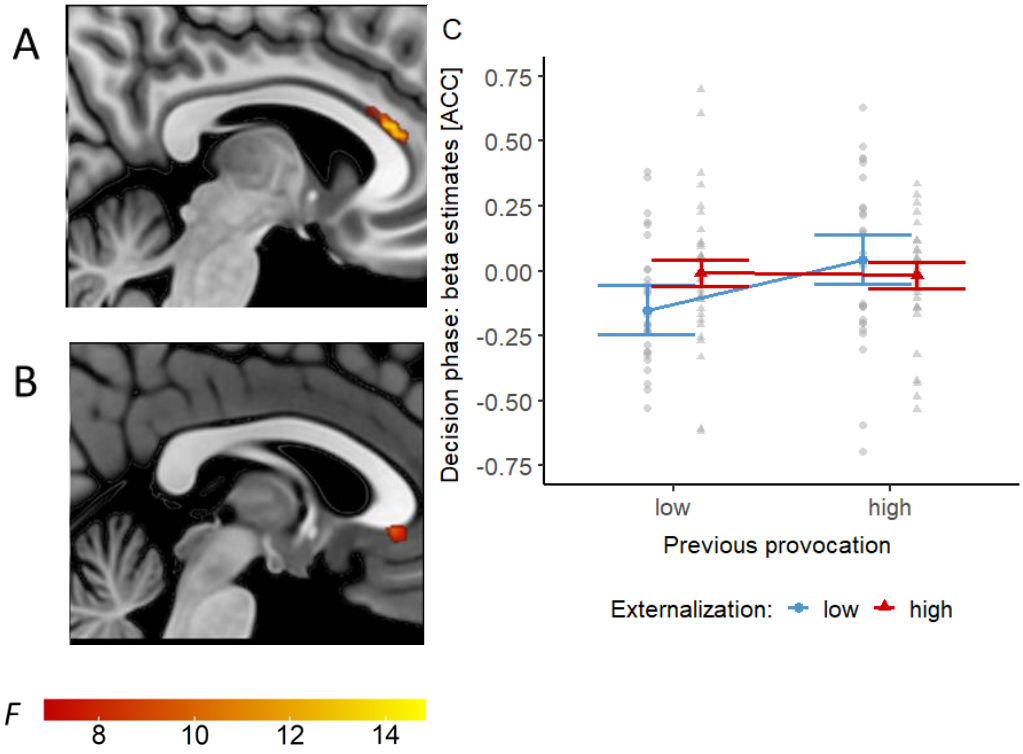

Figure 11. Provocation (high vs. low) x externalization (high vs. low) ROI analysis during decision phase. Location of the activation peaks for the provocation effect in the pregenual region of the rACC (A). Activation peaks for the provocation by externalization interaction in the subgenual region

of the rACC (B). Results of the ROI analyses for the provocation by externalization interaction in the bilateral ACC (C).

Thus, the high externalization group showed a reduced responsivity of the ACC to provocation compared to the low externalization group. This did not apply to the OFC. Figure 13 (Etkin, Egner, & Kalisch, 2011, p. 86) illustrates parcellation of the ACC.

Analyzing the decision phase depending on the pervious outcome (won vs. lost) of the previous trial (*F*-contrast), we found a significant main effect of previous outcome in one suprathreshold cluster including the right posterior cingulate cortex ($F = 38.05$; x = 9, y = -37, z = 11). This indicates more activation after win compared to lose trials as shown by a *post-hoc* *t*-contrast (previous outcome: won > lost). No externalization effect nor interaction of externalization and outcome could be found in whole brain analyses.

Parametric modulation of aggression-response-related activation. The BOLD response within the left precentral gyrus, left superior temporal gyrus, and the right nucleus caudatus covaried positively with the amount of money selected by the participants during decisions that followed lose trials (reactive aggression) (see Appendix F). There was no difference in aggression-response-related activation between the low and high externalization group nor between men and women. Beyond this, the relationship between aggression-response-related activation and *z*-transformed testosterone concentrations differed between women and men in one suprathreshold cluster ($k > 88$), indicating greater activation in the left superior frontal/precentral gyrus in women ($T = 5.06$, x = -21, y = -16, z = 71). *Post-hoc* ROI analyses showed a significant gender by testosterone interaction in the left precentral gyrus ($F(1, 53) = 9.40$, $p = .003$, $\eta_p^2 = .15$; see Figure 12). If excluding the above identified variational case (see section 4.3.2), this effect remained significant ($F(1, 52) = 5.27$, $p = .026$, $\eta_p^2 = .091$). Thus, aggression-response-related activation in the left precentral gyrus/supplementary motor area correlated negatively with testosterone levels in men and positively in women.

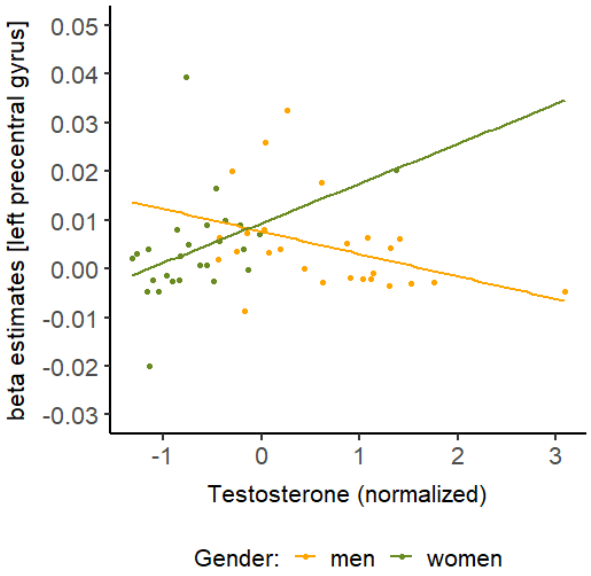

Figure 12. Aggression-response-related activation in the left precentral gyrus for the gender by testosterone interaction (study 3).

Feedback phase. Based on provocation trials (i.e., the preceding trial was lost), there was no externalization effect nor interaction of externalization and provocation for the feedback phase on whole brain level. Within ROI analyses, no differences between low and high provocation could be found in the amygdala ($F(1, 60) = 1.45$, $p = .233$, $\eta_p^2 = .02$) and ACC ($F(1, 60) = 0.03$, $p = .853$, $\eta_p^2 = .00$). The provocation effect in the insula reached significance ($F(1, 60) = 4.73$, $p = .034$, $\eta_p^2 = .07$), but did not survive Bonferroni correction ($n = 3$). Also, neural activation was neither affected by externalization nor the provocation by externalization interaction in the amygdala ($Fs < 1.45$, $ps > .234$), insula ($Fs < 2.11$, $ps > .151$), and ACC ($Fs < 0.97$, $ps > .328$).

Analyzing the feedback phase depending on the outcome of the previous trial (won vs. lost), three suprathreshold clusters with a peak maximum in the right ($F = 83.04$; x = 12, y = -82, z = 29) and left ($F = 71.06$; x = 0, y = -82, z = 26) cuneus, right nucleus caudatus ($F = 143.18$; x = 15, y = 8, z = -10), right ($F = 41.84$; x = 33, y = -13, z = -1) and left ($F = 120.19$; x = -12, y = 8, z = -10) putamen, and left ACC ($F = 50.10$; x = 0, y = 35, z = 11) were rendered by factorial analysis for the main effect of outcome. *Post-hoc* *t*-contrast (won > lost) yielded an activation pattern in these areas.

Externalization solely and the interaction outcome by externalization did not reveal suprathreshold clusters. *Post-hoc* ROI analyses did not show significant externalization or interaction effects in the nucleus accumbens (Fs < 0.29, ps > .592), nucleus caudatus (Fs < 1.13, ps > .292), and putamen (Fs < 1.92, ps > .167) during feedback.

4.4 Discussion

The current study investigated behavioral, affective, and neural processes mediating reactive aggression in externalization in the non-clinical range.

In line with previous results, the participants of our study responded on average more aggressively after high compared to low provocation in the monetary mTAP. On a neural level, this effect was associated with the ACC (activation peak in the pregenual part of the rACC) showing more activation during aggressive responses after high compared to low provocation in ROI analyses. This is in line with most of previous studies using the mTAP with either noise (Beyer, Münte, Erdmann, et al., 2014; Krämer et al., 2007; Krämer et al., 2011) or monetary stimuli (Repple et al., 2017).

While the caudal or dorsal ACC is linked to cognitive processes, the rACC is involved in emotional processing (for review see Etkin, Egner, & Kalisch, 2011; Stevens, Hurley, & Taber, 2011). In particular, the pregenual division of the rACC plays a key role in top-down control to regulate an over-activated emotional response. Thus, higher provocation in the current study led to more activation in the pregenual rACC during aggression, although speculative, indicating the demand to regulate an emotional response.

Additionally, in the current study, the BOLD response in the nucleus caudatus covaried significantly with the amount of money selected by the participants after lose trials (reactive aggression). This is in accordance with the majority of studies applying different versions of the mTAP (Beyer, Münte, Erdmann, et al., 2014; Gan, Sterzer, Marxen, Zimmermann, & Smolka, 2015; Krämer et al., 2007; Lotze et al., 2007) showing an association between activity of the nucleus

caudatus and reactive aggression levels. This link between reactive aggression and reward-related areas prompted the idea that aggressive behavior might be reinforcing (Fanning et al., 2017).

Unexpectedly, we found neither an externalization effect nor a significant impact of the interaction provocation by externalization for aggression levels in this monetary mTAP. This was astonishing, given that our healthy participants with higher externalization scores showed significantly higher self-reported trait aggression (K-FAF, BPAQ, RPQ), with comparable high effect sizes for proactive and reactive aggression scales (see Table 5). Results from previous studies showed a positive correlation between the scores of the TriPM subscales and self-report aggression questionnaires in a community sample, with proactive aggression subscales correlating more strongly with the TriPM than reactive aggression subscales (van Dongen et al., 2017). Thus, the current data support the view that participants with high externalizing within a non-clinical range indeed describe themselves as more aggressive in a self-report questionnaire. However, within the scope of a laboratory behavioral aggression task presented as competitive reaction time task, their inhibitory control appears to still be strong enough to control reactive aggression to a similar extent as in low externalizing participants.

Contrary to our predictions, we found a moderating effect of externalization only on positive affective states, but not on negative affect. However, results from a recent affective dynamics study in women pointed to a similar pattern showing association between externalizing disorders and less persistent positive affect as well as more variable positive emotionality (Scott et al., 2020). Thus, our results are consistent with Scott et al.'s conclusion that positive emotions are more transient in externalizing disorders, potentially also in non-clinical externalization. Further, this finding might also reflect hypersensitivity to emotionally rewarding and punishing cues as described in clinically-relevant externalization (Scott et al., 2020).

In the current study, high externalizing participants showed reduced responsivity in the ACC (activation peak in the subgenual region of the rACC) to higher provocation compared to the low externalization group. Former studies revealed that reduced activation in the subgenual portion of

the rACC is associated with deficits in emotion regulation (Etkin & Wager, 2007). Moreover, in affective decision making, blunted ACC activation was associated with reduced emotional control in CD patients (Cappadocia, Desrocher, Pepler, & Schroeder, 2009; Stadler et al., 2007). Consequently, within the non-clinical range the reduced rACC activation in high externalizing participants may reflect the problems in regulating emotional responses, respectively emotional hyperreactivity.

In response to outcome (win vs. lose), there were no neural differences between externalization groups. Also, on the behavioral level (reactive aggression rates), we did not observe any difference between the quasi-experimental groups, pointing to sustained emotional control. As suggested by the dual-system model (Beauchaine et al., 2017), two neurobiological systems are involved in promoting externalizing problems across the lifespan: the emotional circuitry localized in mesolimbic and limbic areas as well as the control circuitry (e.g., PFC/ ACC). It is hypothesized that the normal developmental delay of the control compared to the emotional system is responsible for a period of increased vulnerability to externalizing problems and can lead to poor affective decision making as well as unfavorable reward and sensation seeking. However, in our study, externalization was not associated with aberrant neural processes in the emotional circuitry including reward-related areas during provocation or reactive aggression. Thus, it can be speculated that dysfunctions in one of these circuitries might lead to manifestations of externalizing problems only within the non-clinical range while deficits in both circuitries enhance the likelihood for developing externalization within the psychopathological range. In line with this reasoning, Gordon et al. (2004) also yielded no behavioral group differences during an affective recognition task in non-clinical participants scoring high on psychopathy, but observed some susceptibility in the emotion circuitry. Further, fMRI studies demonstrated an increase in neural activity in the mesolimbic system and an improvement of clinical symptoms after the administration of the dopamine agonist methylphenidate in patients with ADHD (Vles et al., 2003). Additionally, reduced frontal brain activation in an inhibitory task predicted later alcohol problems in pre-symptomatic adolescents (Norman et al., 2011). Likewise, mesolimbic

dopamine dysfunction has been identified as neural correlate of trait impulsivity conferring vulnerability to externalizing pathologies (see Gatzke-Kopp, 2011). In order to clarify these associations, future work should focus on externalization in clinical and non-clinical samples applying experimental paradigms that are able to activate these neurobiological systems differentially.

In the current study, testosterone levels predicted gender-specific differences in reactive aggression rates, indicating a positive relationship in women and a negative relationship in men. Although previous studies showed rather positive correlations between testosterone and behavioral aggression in men, a meta-analysis from Archer et al. (2005) demonstrated a variety of moderating factors (e.g., age, source of testosterone, population) resulting in a significant attenuation of this relationship. Their findings underpin the view that respective relationships are relatively complex and strongly depend on contextual and personal factors. Thus, keeping in mind that results regarding the relationship between testosterone and aggression in women are rather inconsistent (for review, see Dekkers et al., 2019) it is possible, then, that testosterone facilitates aggression in women in the context of a mock reaction time task like the monetary mTAP task; in particular, since the mTAP presents a socially more agreeable form of taking revenge compared to e.g. physical aggression. For instance, it has been shown that men are prone to more overt physical forms of aggression while women rather tend to exert aggressive behavior in nonphysical ways (Björkqvist, Österman, & Lagerspetz, 1994; Taylor & Epstein, 1967). Consistently, earlier studies using the mTAP provided inconclusive results regarding gender differences in reactive aggression (Dambacher et al., 2015; Lawrence & Hutchinson, 2014; Repple et al., 2018) suggesting that gender-specific performance depends on moderating factors (e.g., contextual, personal and biological influences). In the present study, the precentral gyrus/SMA proved to be a part of the neural correlates of this gender-specific relationship between testosterone and aggression. It is known that the SMA plays an important role in motor planning and updating of motor plans (Mostofsky & Simmonds, 2008). Furthermore, SMA and pre-SMA are implicated in response inhibition, the ability to suppress

undesirable responses. Impairments in response inhibition are related to impulsivity and aggression (Aron, Behrens, Smith, Frank, & Poldrack, 2007). In addition, testosterone reduces the connectivity of the inferior frontal gyrus with the ACC and SMA, implicated in action preparation during social behavior (Bos et al., 2016). In sum, testosterone seems to facilitate aggression by influencing functional connectivity of the response inhibition circuit. Thus, a challenge for future studies will be to investigate this gender-specific link between aggression and testosterone levels with respect to functional connectivity.

In line with previous investigations (Geniole, Busseri, & McCormick, 2013), the relationship between reactive aggression and testosterone was not moderated by cortisol secretion in the current study. Additionally, a recent meta-analysis provided only marginal support for the dual-hormone hypothesis of testosterone and cortisol (Dekkers et al., 2019).

At this point, some limitations of our study have to be mentioned. First, prior to the monetary mTAP, participants went through a resting state and DTI sequence (to be reported elsewhere). Although these sequences demanded no specific task performance and did not evoke specific affective states conferring to monetary mTAP performance, a nonsystematic effect of fatigue cannot be excluded. Second, although our sample size is suitable for a fMRI study, larger sample sizes enhancing statistical power might be needed to detect group differences in the monetary mTAP. Third, we mainly tested University students potentially limiting the generalizability to other populations within the non-clinical range.

4.4.1 Conclusions

We successfully induced provocation dependent aggressive behavior linked with increased activation in the ACC. Although high externalizing participants did not behave more aggressively compared to the low externalization group, aberrant activation in the cognitive control network including the ACC could be observed even within this non-clinical range. Based on these findings, it might be speculated that additional dysfunctional regulation in other circuits that mediate pathological externalizing symptoms (e.g., in the emotional circuit), are essential for developing

externalizing disorders like ADHD, CD, and SUD. Furthermore, the gender-specific relationship between testosterone and behavioral aggression, which was mediated by SMA activation, points to potential correlates of gender-specific aggression responses.

Chapter V:

General discussion

In Chapter I, an overview of previous research on stress and aggression in the externalizing spectrum was presented. The theoretical background, assessment approaches as well as empirical neurobiological results were reviewed. In summary, multiple lines of evidence suggest that (1) the externalizing spectrum is continuously distributed including disinhibitory personality traits (Krueger et al., 2002; Walton et al., 2011), that (2) externalization, at least in the pathological range, is linked to altered decision making and emotion processing including acute threat response and reward processing (e.g., Cappadocia et al., 2009; Fairchild et al., 2019; Lee et al., 2009), and that (3) these neurobiological alterations are associated with HPA axis (stress reactivity) dysfunctioning (e.g., McBurnett et al., 2005; Virkkunen, 1985) and might explain higher rates of reactive aggression.

Regarding the externalizing spectrum, there is a lack of research in biobehavioral substrates of externalization within the non-clinical range, especially in stress response and aggression processing.

The present work aimed at investigating whether interindividual differences in externalizing behavior within the non-clinical range are associated with different affective, psychoendocrine and neural responses to acute stress exposure and provocation in a monetary variant of the mTAP.

In Study 1, a version of the monetary mTAP was developed as well as validated. Additionally, chapter II provided recommendations regarding a standardized procedure. Subsequently, participants from the higher versus lower range of the non-clinical variation in externalization were investigated (behavioral and endocrine parameter as well as neural response) in scanner environments during Scan*STRESS* (study 2) and the monetary mTAP (study 3).

In the following general discussion, the author will discuss and integrate the results of the three studies and take a closer look at strengths, limitations as well as implications for future research.

5.1 Summary of the findings

The first study was designed to validate a new version of a monetary mTAP (see Chapter I). This validation study aimed at examining effects of provocation on reactive aggression, gender differences as well as the relationship between self-reported trait reactive aggression and behavioral reactive aggression in a monetary mTAP using LMMs. Furthermore, provocation stimuli were

presented either randomly or in a fixed sequence to assess the role of provocation sequence in an experimental between-subjects design. Results revealed that provocation influenced reactive aggression significantly. Moreover, gender differences could be observed indicating lower aggression levels in women than in men. Provocation did not modulate the relationship between gender and reactive aggression. In terms of convergent validity, self-report trait reactive aggression was associated with aggression levels in the paradigm. Unlike predicted, the fixed order of monetary stimuli generated mainly as a sequence of triplets of the same provocation category (see section 2.2.2 and Appendix B) has not conferred any advantage on aggression induction. This confirms the assumption that reactive aggression appeared to be mostly affected by the immediately preceding provocation trial. In short, the findings provided new evidence supporting the view that the monetary mTAP is a valid paradigm for the induction and measurement of behavioral reactive aggression in the laboratory.

In the second study (see chapter III), healthy participants from the higher versus lower range of the normal, non-clinical variation in externalization, as assessed by the subscales disinhibition and meanness of the TriPM, were recruited and exposed to the Scan*STRESS* paradigm following a quasi-experimental design. In both groups, Scan*STRESS* induced a significant rise in salivary cortisol levels. The high externalization group showed lower cortisol responses to acute stress compared to the low externalizing group. *Post-hoc* analysis demonstrated that this effect could mainly be observed in men. Regarding imaging data, the detected activation and deactivation patterns including prefrontal cortex and limbic areas during acute stress (compared to the control condition) are well known in playing a key role in HPA axis regulation. No differences in neural activity in response to stress could be found between externalization groups on whole brain level and within the predefined ROI masks (bilateral amygdala, ACC, nucleus accumbens, putamen, nucleus caudatus, and OFC). However, when individual cortisol increases were used as continuous covariate, the group-specific cluster including the nucleus caudatus and putamen (dorsal striatum) reached significance. In the high externalization group, the dorsal striatum activity correlated negatively with

the cortisol stress response while this correlation remained nonsignificant in participants with low externalization. *Post-hoc* analysis showed that this effect could mainly be observed in women.

In the third study (see Chapter IV), participants with high and low externalization levels performed the monetary mTAP, which was validated in study 1, inside the MRI scanner. Results show that participants behaved more aggressively after higher levels of provocation, replicating results from study 1 (validation study). Neurally, this provocation effect was mediated by increased activity in the rostral part of the ACC, known as affective division, with extensive connections to the amygdala (Etkin, Egner, & Kalisch, 2011; Stevens, Hurley, & Taber, 2011). During the feedback phase (watching outcome), there were no differences in the activation pattern between high and low provocation. Testing the moderating effect of externalization on aggression levels revealed no significant result. On neural level, compared to low externalizing participants, high externalizing participants showed a reduced increase in the activity of the ACC after high versus low provocation. This finding indicates alterations in integrating emotion and cognition (e.g., resolving of emotional conflicts) in the high externalization group also within the non-clinical range (Etkin & Wager, 2007). In former studies, usage of an anger induction paradigm showed that activation in the rACC was closely linked to the experience of anger (Dougherty et al., 2004).

Analyzing pooled salivary testosterone concentrations revealed a gender-specific relationship between aggression levels and testosterone concentrations. In men, aggression levels correlated negatively with salivary testosterone concentrations while this correlation was positive in women. In parametric modulation analyses, the supplementary motor area emerged as one of the neural correlates of this relationship. Unlike predictions of the dual-hormone-hypothesis, salivary cortisol did not moderate the correlation between aggression and testosterone. The additional factors gender and trait reactive aggression did not significantly predict aggression levels in the laboratory paradigm.

5.2 Discussion and integration of the findings

5.2.1 Methodological advances relative to the monetary mTAP

Recalling critics of the CRTT, main aspects deal with the questions of validity and standardization. Over the past decades, a methodological and quantitative flexibility in using the mTAP has been observed in research literature. That is the case particularly for the type of stimuli, number of trials or quantification strategies. Besides, influences of other factors such as presentation of the provocation stimuli (e.g., random or fixed order) or other motives on reactivity of the participants are unclear, which casts doubt on the comparability of studies using different analytic strategies of the TAP (Elson et al., 2014). Study 1 focused on validation of a variant of the monetary mTAP and considered the above-mentioned aspects of critics. In study 3, the validated variant of the monetary mTAP was applied in scanner environments. The following paragraphs provide the methodological advances regarding the monetary mTAP, which resulted from the present thesis.

First, in terms of validity, study 1 showed a significant influence of provocation on aggression levels in a dose-response manner, a significant association of self-reported trait reactive aggression and aggression levels (convergent validity) as well as gender differences (external validity). One prime candidate for undermining validity of the mTAP is a constrained credibility related to the cover story (Anderson et al., 2000; Anderson et al., 2004). Consequently, chapter II provides a standard procedure of the monetary mTAP including a deception check questionnaire. In both studies, applying the monetary mTAP (study 1 and 3), suspicion rates reached a stable level of 24 to 33%, which seems a high proportion, compared to previous studies (e.g., 5 to 11%) (see section 2.4). However, the low-threshold questionnaire was utilized to include the variable deception check as factor to the main model. Empirical evidence of study 1 and 2 revealed that suspicion had no significant impact on aggression levels.

Second, aggression researchers stated that outcome parameters in the TAP are confounded by other motives such as the tendency to act competitively, thereby limiting the findings regarding validity of the TAP (Hyatt, Zeichner, & Miller, 2019; Tedeschi & Felson, 1994). One of the first

publications focusing this topic was the work of Bernstein et al. (1987) showing no associations between reactive aggression levels in a mTAP with shock stimuli and measures of competitive behavior. In this thesis, study 1 provided recommendations for dealing with confounding motives. Competitiveness could be identified as a relevant factor in explaining portions of the individual subject variance (random intercept). Thus, the individual aggression level can partly be attributed to the tendency to act competitively. Therefore, we recommend to include and validate potentially influencing factors step by step. If, thereby, variance allocation is more clearly differentiated, causal relations (e.g., influence of substances, part of making clinical diagnosis or predictions) can be derived from the performance of the monetary mTAP. In short, these findings support the validity of the monetary mTAP und fill a gap in aggression research.

Third, results from study 1 suggest a predominant influence of the last provocation since no difference between stimulus sequences could be found. The findings did not point to a significant role of the second to last and third to last provocation stimulus in the participants' response selection. Thus, it can be concluded that the specific provocation sequence, which we applied, had no impact on the behavioral aggression outcome. Based on the presented data, we would recommend applying the random order version to avoid unfavorable side-effects (e.g., systematic inattention effects in specific paradigm phases).

Finally, study 1 and 3 could show that the random and fixed effects structures of the mTAP model are complex. For instance, results of study 1 indicate that the widely published provocation effect is supposed to be divided into a general provocation effect over all participants and an effect which is ascribed to the characteristics of the individual participant. Thus, without such a distinction, findings might be misleading and can result in an overestimation of the eliciting character of the aggression paradigm. Moreover, researchers would be well-advised to consider the structure of the random components. These portions of variances are classified as person-specific and cannot be explained by gender or trait aggression. Focusing on these components may also contribute to a better understanding of the mechanisms underlying TAP performance.

In the context of misleading conclusions, it is noteworthy to have a look at the effect sizes and statistical power: Study 1 elicited a clear gender effect in aggression levels within a sample of 209 participants. In contrast, this effect was absent in study 3 (sample size of 61 participants). Thus, it can be concluded that the effect sizes of aggression-related variables, which are typically reported in aggression paradigms, are limited, and enhanced statistical power is demanded to detect subtle effects (see also Hyatt et al., 2019). Both, appropriate variance allocation and the partly limited effect sizes should be taken in account when interpreting results of the mTAP.

In short, study 1 and 3 provide empirical evidence that the monetary mTAP is a valid paradigm to capture reactive aggression in the laboratory, comparable to other aggression induction paradigms. Study 1 also includes recommendations towards a standard procedure in conducting and analyzing the monetary mTAP.

5.2.2 Biobehavioral substrates of non-clinical externalization

As reported in section 1.3.5 and 1.4.3, externalizing disorders are associated with functional changes in brain areas involved in emotion, reward and threat processing as well as decision making. These changes are results of a multifactorial interaction of risk factors such as genetic vulnerabilities, temperament, environmental potentiators and in particular stress throughout the lifespan (Hinshaw & Beauchaine, 2016). Following the concept of allostatic load (McEwen, 1998b), chronic stress (e.g., early adversity) might, in the long run, lead to a dysfunctioning HPA axis and neural aberrations. In externalizing pathologies, most studies revealed a hypoactivity of the HPA axis during acute stress. The question whether these deviations can partly be observed in externalization within the non-clinical range has not been answered, yet.

Results from functional imaging studies focusing on externalization within the non-clinical range are rare. One study provided support for the hypothesis that aberrant reward and threat responses can also be found within a healthy population exhibiting externalizing behavior assessed with the ESI. Foell et al. (2016) revealed reduced activation of the nucleus accumbens (ventral striatum) during a preparation phase and enhanced amygdala reactivity during viewing of pleasant and

aversive pictures. As highlighted in section 1.3.5, the acute threat response is also linked to stress regulation and HPA axis functioning.

Similar to externalizing disorders (Couture et al., 2008; Fairchild et al., 2008), study 2 showed that hyporeactivity of HPA axis in response to acute stress can also be found in externalization within the non-clinical range. At the neural level, the cortisol-related dorsal striatum activation differed between the higher and lower externalization group in this thesis (study 2, see chapter III). Thus, altered activity in the dorsal striatum appears to be one of the key mechanisms underlying the hyporeactivity of the HPA axis in externalization within the non-clinical range. This is supported by earlier findings showing that the mesolimbic dopamine system (including the dorsal striatum) is sensitive to stress exposure over the life span. Moreover, the presence of glucocorticoid receptors in the striatum has also been proved in former studies. Additionally, in externalizing disorders, behavioral abnormalities (e.g., reward seeking) are also associated with mesolimbic dysfunctioning (for review see Gatzke-Kopp, 2011). Furthermore, recent findings showed that cortisol mediates stress-induced functional connectivity between the centromedial amygdala and the dorsal striatum (Vogel et al., 2015). The authors concluded that cortisol plays a crucial role in the stress-induced shift towards the salience network during vigilance processing, leading to a habitual response. Following this, maladaptive stress coping in externalization, also within the non-clinical range, could be a result of dysfunctional shifting of neural resources towards the salience network mediated by dorsal striatum and cortisol (see chapter IV). This is in line with observations that externalization is associated with a strengthened attention on threat-related cues and reduced cognitive flexibility (Gatzke-Kopp, 2011), which, thereby, promote maladaptive stress coping and impulsive behavior.

Post-hoc gender analyses showed the externalization effect in cortisol response was driven by the male sample, whereas the cortisol-related dorsal striatum activation differing between low and high externalization group could mainly observed in women. This is in line with previous studies revealing a gender-specific neural activation network underlying the central stress response. According to that, female stress response is related to striato-limbic activation and frontal activation

is enhanced in men (Goldfarb et al., 2019; Wang et al., 2007). Thus, it can be concluded that stress response in female externalization is mediated by altered activation in known stress-related brain areas. However, in men, more research is needed to clarify neural basis of the observed externalization effect in cortisol responses.

Beside the aberrant dorsal striatum activity in non-clinical externalization, reduced rACC (subgenual region) responsivity to provocation was found in the high externalization group during reactive aggression. The rACC is involved in emotional processing (for review see Etkin et al., 2011) which is known to be deficient in participants exhibiting externalizing problems. In particular, the pregenual division of the rACC plays a key role in top-down control to regulate an over-activated emotional response (Stevens et al., 2011). Interestingly, several lines of research showed that patients with ADHD and CD exhibit hypoactivation in the fronto-striatal network including the ACC and in the nucleus caudatus during goal-directed behavior (Bayard et al., 2018; Shannon, Sauder, Beauchaine, & Gatzke-Kopp, 2009). In one study, abstinent drug users showed risky decision making mediated by abnormal ACC functioning (Fishbein et al., 2005). Furthermore, apathy is associated with reduced ACC activation in participants with antisocial tendencies (Veit et al., 2002). It can be noted that deficient ACC activation related to several externalizing disorders can also be found in a non-clinical sample which, however, did not result in different aggression levels. The only study investigating externalization within the non-clinical range revealed a positive association between the degree of externalization and physical aggression during a mTAP with electrical shocks, however, only under high provocation conditions (Subramani et al., 2019). In short, provoking characteristics of the monetary mTAP could be sensitive enough to reveal neural aberration in high externalization within the non-clinical range, however, insufficient to induce behavioral differences.

Taken together, it should be noted that, to my knowledge shown for the first time, during acute stress response and reactive aggression, participants exhibiting externalizing behavior within the non-clinical range displayed a partly gender-specific fronto-striatal alterations (which is well-known

in ADHD and CD patients) as well as, abnormalities in HPA axis. These aberrations can be described as part of the biobehavioral substrates of the externalizing spectrum.

5.2.3 Testosterone and aggression

Several meta-analyses in the past decade showed that testosterone and aggression correlate only moderately (Archer et al., 2005; Dekkers et al., 2019). This modest correlation remains unchanged when considering the moderating effects of cortisol secretion on the relationship between testosterone and aggression (dual-hormone-hypothesis) (Dekkers et al., 2019). Moreover, revealing conflicting results, previous studies reported a positive (Denson et al., 2013; Welker et al., 2014), negative (Mehta & Josephs, 2010) or no (Platje et al., 2015) correlation between testosterone and aggression. One prime candidate for explaining these conflicting results is the hypothesis that the relationship between testosterone and aggression is in many cases influenced by moderating factors (e.g., contextual, personal, and biological influences). Support for this assumption is provided by studies from different fields showing a moderating effect of trait anxiety, gender, type of population, social inclusion or exclusion as well as perceived threat or reward (for review see Dekkers et al., 2019). This is in accordance with the fitness model of testosterone dynamics (FMTD; Geniole & Carré, 2018) proposing an influence of individual factors (e.g., sex, genotype, personality, other hormones) and situational factors (e.g., status, setting, social hierarchy) on the testosterone metabolism. Dekkers et al. (2019) also stated that some conflicting results are caused by differences in methodology. In the following, the respective empirical results of study 2 and 3 are interpreted and integrated in the light of these reports (Dekkers et al., 2019).

Study 3 revealed a positive relationship between salivary testosterone and aggression levels in the mTAP in women and a negative association in men. Earlier findings, related to gender differences in reactive laboratory aggression and testosterone, are mixed. Although a number of studies showed a negative relationship between testosterone and aggression in women (Buades-Rotger et al., 2016) or null-finding (Geniole et al., 2013) during a mTAP with noise blasts, one previous study, for instance, with healthy women, reported a positive predictive effect of

testosterone on aggression levels in a mTAP with noise stimuli (Denson et al., 2013). In men, to the best of my knowledge, only one study exists testing the testosterone - aggression relationship with a monetary mTAP revealing higher provocation-related aggression levels in an exogenous testosterone group (Wagels et al., 2019). Most studies using other stimuli in the mTAP also support the view of a facilitating effect of testosterone on aggression in men (Berman et al., 1993). Overall, it should be emphasized that the non-physical, more competitive character of the current variant of the monetary mTAP (e.g., non-physical type of aggression) as situational factor seem to elicit testosterone-related reactive aggression more in women than in men. According to the recommendations of Dekkers et al. (2019), a study with different laboratory paradigms in women and men is required to clarify influencing effects of the different situational factors of each aggression paradigm.

An extensive body of literature investigated the influential effect of testosterone, and partly cortisol, on aggression-related brain regions. Several fMRI studies indicate that testosterone facilitates activation of subcortical areas (e.g., amygdala) and suppresses prefrontal activity in response to threat or provocation, leading to reduced inhibitory control and reactive aggression (Hermans et al., 2008; Mehta & Beer, 2010). The current results identified the precentral gyrus/SMA as the neural correlate of this gender-specific relationship between testosterone and aggression. Figure 13 (Etkin, Egner, & Kalisch, 2011, p. 86) illustrates parcellation of the ACC and mPFC including SMA. Furthermore, SMA and pre-SMA as subdivision of the mPFC play a crucial role in response inhibition, i.e. the ability to suppress undesirable responses (Floden & Stuss, 2006; Nachev, Wydell, O'neill, Husain, & Kennard, 2007; Picton et al., 2007). It was shown that testosterone reduces the connectivity of the inferior frontal gyrus with the ACC and SMA during social behavior (Bos et al., 2016). In short, this thesis provides support for a gender-specific influencing effect of testosterone on parts of the response inhibition circuit resulting in aggressive behavior.

Summarizing the contribution of the present thesis to the area of testosterone and aggression, it should be emphasized that non-physical provocation during monetary mTAP elicited testosterone-

related aggression more in women than in men. Moreover, SMA seems to play a relevant role in the neural network mediating the testosterone - aggression relationship.

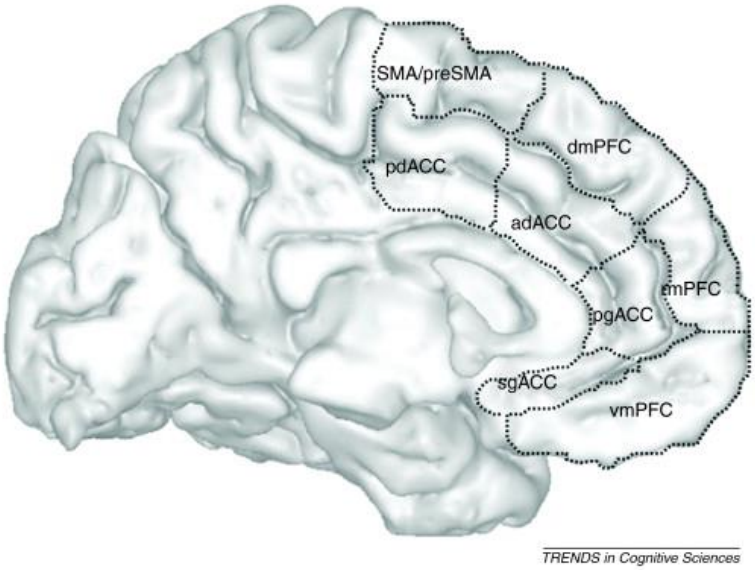

Figure 13. Subdivisions of ACC and mPFC. Notes: sg = subgenual, pg = pregenual, vm = ventromedial, rm = rostromedial, dm = dorsomedial, ad = anterior dorsal, pd = posterior dorsal. Reprinted from "Emotional processing in anterior cingulate and medial prefrontal cortex," by A. Etkin, T. Egner, & R. Kalisch, 2011, *Trends in Cognitive Sciences*, 15, p. 86.

5.3 Strengths and limitations of the present thesis

The following section will initially cover the strengths and subsequently the limitations of the present thesis. First, the present thesis is the first one investigating the hypothesis that interindividual differences in externalizing behavior are associated with alterations in affective, psychoendocrine and neural stress response and reactive aggression processing. Results support the idea of a biobehavioral and dimensional perspective on trait externalization.

Second, in all three studies men and women were systematically recruited in comparable sample sizes to investigate and control for moderating effects of gender. Especially, the endocrine parameters (cortisol, testosterone) are susceptible to gender differences. Additionally, reactive aggression is known to be different in men and women.

Third, in study 2 and 3, which assessed endocrine parameters, the current work took into account endocrine fluctuations within the female menstrual cycle by testing only women in the luteal phase of their cycle, in case they were not taking contraceptives.

Fourth, the developed variant of the monetary mTAP passed a validation process (study 1) before application in the MR scanner, which marks a proof of quality and supports the conclusions which were drawn in the studies. Thereby, a standardized and experimentally verified procedure and an analytic strategy were provided responding to the broad critic of the TAP (chapter II).

Fifth, while an extensive body of literature failed to evoke satisfactory rises in cortisol in response to acute stress exposure in studies on stress and externalizing behavior, study 2 revealed valid cortisol releases using the well-validated Scan*STRESS*-paradigm.

Finally, in study 2 and 3, fMRI was used to examine neural mechanisms underlying externalization within the non-clinical range. This approach allows to record neural processing with an appropriate spatial solution which is suitable for tapping activity in subcortical regions being dominant in stress response and aggression processing.

The limitations of study 1-3 are described in the respective sections of each discussion. The limitations related to all three studies are outlined in the following paragraphs.

In the quasi-experimental studies 2 and 3, an extreme groups approach (EGA) within the non-clinical range was used including only participants scoring high versus low in the respective subscales of the TriPM. This design was realized in order to reach sufficient statistical power. It should be noted that with the EGA, information about the rest of the non-clinical range could not be collected, which is all the more relevant, since the present thesis pursues a dimensional perspective on trait externalization. However given the early stage of research in this field, it appears appropriate to follow the EGA (Preacher, 2014; Preacher, Rucker, MacCallum, & Nicewander, 2005).

Second, the samples of the three studies comprised young healthy adults. Thus, to what extent the results could be generalized to other populations within the non-clinical range remains unclear.

Third, due to the strict inclusion criteria (e.g., high vs. low externalization, MR compatibility, health status) in study 2 and 3, other potentially confounding variables like handedness or use of oral contraceptives could not be kept constant, leading to a partly heterogeneous sample. For instance, it could not be controlled for adverse experiences in childhood and adolescents (e.g., early life stress) although studies show that patients with externalizing disorders are more frequently exposed to early life stress (Deater-Deckard et al., 1998), which can potentially influence neural activity (Cohen et al., 2006; Sapolsky, 1999).

Finally, in MRI research spatial and temporal solution is always debated to enhance imaging quality and capture smaller brain structures (e.g., brain stem). In study 2 and 3, echo-planar imaging sequences were used including parallel acquisition techniques to reduce acquisition time. With the relatively widespread multiband sequences, which collect multiple slices simultaneously, the quality of the scans (signal-to-noise ratio) could have been kept stable while improving the temporal solution. By contrast, multiband sequences can lead to a higher physiological strain for the participants (e.g., noise, temperature) (Moeller et al., 2010).

5.4 Future directions

For the first time, the present thesis investigates alterations in affective, psychoendocrine and neural stress response as well as reactive aggression processing in a non-clinical sample exhibiting externalizing problems. The following last section of the thesis presents a couple of research questions based on our findings which could be addressed by future studies.

First, as outlined in preceding sections, the externalizing spectrum is a continuously distributed trait. Thus, biobehavioral correlates during Scan*STRESS* and the monetary mTAP should be investigated in a forensic sample to evaluate, if the present findings in participants from the non-clinical range can be generalized to clinical samples. Prof. Dr. Boris Schiffer and his work group will address these research questions.

Second, future studies should not only focus on extreme groups but take advantage of the full range (Preacher, 2014; Preacher et al., 2005).

Third, results of study 2 and 3 showed fronto-striatal alterations in non-clinical externalization during stress response and reactive aggression. Previous studies also revealed a reduced effective connectivity between the ACC and striatum in externalizing disorders (Fishbein et al., 2005; Shannon et al., 2009). Future studies should assess the connectivity and the mutual influence between frontal and striatal areas in non-clinical externalization. Thereby, one prime approach is dynamic causal modeling (DCM, Friston, 2009) integrating the advantage of directionality.

Fourth, another line of research is concerned with the resting state functional connectivity , which enables the exploration of interconnected neural networks, for instance the default mode network (DMN) under baseline activity (Raichle et al., 2001). Raichle et al. (2001) identified the medial prefrontal cortex, the posterior cingulate cortex and the precuneus as well as the bilateral inferior parietal lobule as components of the DMN. In externalizing disorders, studies revealed aberrant functioning of the DMN (Sun et al., 2012). Additionally, first studies exploring externalization within the non-clinical range show altered connectivity, for instance, within the insula networks (Abram et al., 2015). More research is required to examine non-clinical externalization with this promising network approach.

Finally, the aspects of the last two paragraphs apply to the relationship between testosterone and aggression. Functional connectivity approaches could clarify connections between the supplementary motor area revealed in study 3 and areas known to be involved in testosterone processing (e.g., amygdala, frontal areas).

5.5 Conclusions

The present thesis, for the first time, investigates the hypothesis that interindividual differences in externalizing behavior are associated with alterations in affective, psychoendocrine, and neural stress response and reactive aggression processing. Guidance is provided by previous studies revealing abnormalities in externalizing disorders (Fairchild et al., 2008; Shannon et al., 2009). The results of the three studies showed that (1) the applied monetary mTAP is a valid measure for reactive aggression, that (2) externalization predicts blunted cortisol responses, which are associated

with gender-specific striatal abnormalities, that (3) high externalizing participants showed reduced responsivity to provocation in the inhibitory control system, which did not lead to higher rates of reactive aggression. (4) Gender-specific salivary testosterone was related to reactive aggression reflected by supplementary motor area activation at the neural level.

These findings support the biobehavioral, dimensional perspective on externalization (ranging from non-clinical externalizing behavior to pathological abnormalities), which connects the behavioral trait with biological vulnerabilities. This is confirmed by previous studies also showing blunted cortisol response (Susman, 2006; Van Goozen & Fairchild, 2006; Van Goozen et al., 2000) and fronto-striatal dysfunctioning in externalizing disorders (Cubillo et al., 2010; Hart, Radua, Nakao, Mataix-Cols, & Rubia, 2013; Shannon et al., 2009). Future studies will benefit from the current findings through a better understanding of the etiological pathways underlying heterotypic trajectories of psychopathologies. Additionally, the current work confirms the hypothesis that the association between aggression and testosterone depends strongly on moderating factors (e.g., contextual, personal and biological influences) proposed by FMTD (Geniole & Carré, 2018).

References

Abram, S. V., Wisner, K. M., Grazioplene, R. G., Krueger, R. F., MacDonald III, A. W., & DeYoung, C. G. (2015). Functional coherence of insula networks is associated with externalizing behavior. *Journal of Abnormal Psychology, 124*, 1079. https://doi.org/10.1093/cercor/bhs126

Adams, R. D., & Victor, M. (1989). *Multiple sclerosis and allied demyelinative diseases*. New York: McGill-Hill.

Akdeniz, C., Tost, H., Streit, F., Haddad, L., Wüst, S., Schäfer, A., Schneider, M., Rietschel, M., Kirsch, P., & Meyer-Lindenberg, A. (2014). Neuroimaging evidence for a role of neural social stress processing in ethnic minority-associated environmental risk. *JAMA Psychiatry, 71*, 672-680. https://doi.org/10.1001/jamapsychiatry.2014.35

Alink, L. R., Cicchetti, D., Kim, J., & Rogosch, F. A. (2012). Longitudinal associations among child maltreatment, social functioning, and cortisol regulation. *Developmental Psychology, 48*, 224-236. https://doi.org/10.1037/a0024892

Alink, L. R., van Ijzendoorn, M. H., Bakermans-Kranenburg, M. J., Mesman, J., Juffer, F., & Koot, H. M. (2008). Cortisol and externalizing behavior in children and adolescents: mixed meta-analytic evidence for the inverse relation of basal cortisol and cortisol reactivity with externalizing behavior. *Developmental Psychobiology, 50*, 427-450. https://doi.org/10.1002/dev.20300

Anderson, C. A. (1997). Effects of violent movies and trait hostility on hostile feelings and aggressive thoughts. *Aggressive Behavior, 23*, 161-178. https://doi.org/10.1002/(SICI)1098-2337(1997)23:3<161::AID-AB2>3.0.CO;2-P

Anderson, C. A., Anderson, K. B., Dorr, N., DeNeve, K. M., & Flanagan, M. (2000). Temperature and aggression. *Advances in Experimental Social Psychology, 32*, 63-133. https://doi.org/10.1016/S0065-2601(00)80004-0

Anderson, C. A., & Bushman, B. J. (2001). Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: a meta-analytic review of the scientific literature. *Psychological Science, 12*, 353-359. https://doi.org/10.1111/1467-9280.00366

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology, 53*, 27-51. https://doi.org/10.1146/annurev.psych.53.100901.135231

Anderson, C. A., & Carnagey, N. L. (2009). Causal effects of violent sports video games on aggression: is it competitiveness or violent content? *Journal of Experimental Social Psychology, 45*, 731-739. https://doi.org/10.1016/j.jesp.2009.04.019

Anderson, C. A., Carnagey, N. L., Flanagan, M., Benjamin, A. J., Eubanks, J., & Valentine, J. C. (2004). Violent video games: specific effects of violent content on aggressive thoughts and behavior. *Advances in Experimental Social Psychology, 36*, 200-251. http://dx.doi.org/10.1016/S0065-2601(04)36004-1

Anderson, C. A., & Dill, K. E. (2000). Video games and aggressive thoughts, feelings, and behavior in the laboratory and in life. *Journal of Personality and Social Psychology, 78*, 772-790. http://dx.doi.org/10.1037/0022-3514.78.4.772

Archer, J. (2004). Sex differences in aggression in real-world settings: a meta-analytic review. *Review of General Psychology, 8*, 291-322. http://doi.org/10.1037/1089-2680.8.4.291

Archer, J. (2006). Testosterone and human aggression: an evaluation of the challenge hypothesis. *Neuroscience & Biobehavioral Reviews, 30*, 319-345. https://doi.org/10.1016/j.neubiorev.2004.12.007

Archer, J., Graham-Kevan, N., & Davies, M. (2005). Testosterone and aggression: a reanalysis of Book, Starzyk, and Quinsey's (2001) study. *Aggression and Violent Behavior, 10*, 241-261. https://doi.org/10.1016/j.avb.2004.01.001

Arend, M. G., & Schäfer, T. (2019). Statistical power in two-level models: a tutorial based on Monte Carlo simulation. *Psychological Methods, 24*, 1-19. http://dx.doi.org/10.1037/met0000195

Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J., & Poldrack, R. A. (2007). Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. *Journal of Neuroscience, 27*, 3743-3752. https://doi.org/10.1523/JNEUROSCI.0519-07.2007

Bandura, A., & Walters, R. H. (1977). *Social learning theory* (Vol. 1). New York: Prentice Hall.

Baron, R., & Richardson, D. (1994). *Human aggression* (2nd ed.). New York: Plenum Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 48. http://dx.doi.org/10.18637/jss.v067.i01.

Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: the dark side of high self-esteem. *Psychological Review, 103*, 5-33. https://doi.org/10.1037/0033-295X.103.1.5

Bayard, F., Thunell, C. N., Abé, C., Almeida, R., Banaschewski, T., Barker, G., Bokde, A. L., Bromberg, U., Büchel, C., & Quinlan, E. B. (2018). Distinct brain structure and behavior

related to ADHD and conduct disorder traits. *Molecular Psychiatry*, 25, 1-14. https://doi.org/10.1038/s41380-018-0202-6

Beauchaine, T. P., Shader, T. M., & Hinshaw, S. P. (2016). An ontogenic process model of externalizing psychpathology. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *The Oxford handbook of externalizing spectrum disorders*. New York: Oxford University Press.

Beauchaine, T. P., Zisner, A. R., & Sauder, C. L. (2017). Trait impulsivity and the externalizing spectrum. *Annual Review of Clinical Psychology, 13*, 343-368. https://doi.org/10.1146/annurev-clinpsy-021815-093253

Berkowitz, L. (1993). Pain and aggression: some findings and implications. *Motivation and Emotion, 17*, 277-293. https://doi.org/10.1007/BF00992223

Berkowitz, L., & LePage, A. (1967). Weapons as aggression-eliciting stimuli. *Journal of Personality and Social Psychology, 7*, 202-207. https://doi.org/10.1037/h0025008

Berman, M., Gladue, B., & Taylor, S. (1993). The effects of hormones, type A behavior pattern, and provocation on aggression in men. *Motivation and Emotion, 17*, 125-138. https://doi.org/10.1007/BF00995189

Bernstein, S., Richardson, D., & Hammock, G. (1987). Convergent and discriminant validity of the Taylor and Buss measures of physical aggression. *Aggressive Behavior, 13*, 15-24. http://dx.doi.org/10.1002/1098-2337(1987)13:1<15::AID-AB2480130104>3.0.CO;2-K

Bertsch, K., Böhnke, R., Kruk, M. R., Richter, S., & Naumann, E. (2011). Exogenous cortisol facilitates responses to social threat under high provocation. *Hormones and Behavior, 59*, 428-434. https://doi.org/10.1016/j.yhbeh.2010.12.010

Bettencourt, B., & Miller, N. (1996). Gender differences in aggression as a function of provocation: a meta-analysis. *Psychological Bulletin, 119*, 422-447. http://dx.doi.org/10.1037/0033-2909.119.3.422

Beyer, F., Buades-Rotger, M., Claes, M., & Krämer, U. M. (2017). Hit or run: exploring aggressive and avoidant reactions to interpersonal provocation using a novel Fight-or-Escape

Paradigm (FOE). *Frontiers in Behavioral Neuroscience, 11*, 190. https://doi.org/10.3389/fnbeh.2017.00190

Beyer, F., Münte, T. F., Erdmann, C., & Krämer, U. M. (2014). Emotional reactivity to threat modulates activity in mentalizing network during aggression. *Social Cognitive and Affective Neuroscience, 9*, 1552-1560. https://doi.org/10.1093/scan/nst146

Beyer, F., Münte, T. F., Göttlich, M., & Krämer, U. M. (2015). Orbitofrontal cortex reactivity to angry facial expression in a social interaction correlates with aggressive behavior. *Cerebral Cortex, 25*, 3057-3063. https://doi.org/10.1093/cercor/bhu101

Beyer, F., Münte, T. F., & Krämer, U. M. (2014). Increased neural reactivity to socio-emotional stimuli links social exclusion and aggression. *Biological Psychology, 96*, 102-110. https://doi.org/10.1016/j.biopsycho.2013.12.008

Björkqvist, K., Österman, K., & Lagerspetz, K. M. (1994). Sex differences in covert aggression among adults. *Aggressive Behavior, 20*, 27-33. https://doi.org/10.1002/1098-2337(1994)20:1<27::AID-AB2480200105>3.0.CO;2-Q

Blair, R. J., Leibenluft, E., & Pine, D. S. (2014). Conduct disorder and callous–unemotional traits in youth. *New England Journal of Medicine, 371*, 2207-2216. https://doi.org/10.1056/NEJMra1315612

Blair, R. J., Veroude, K., & Buitelaar, J. (2018). Neuro-cognitive system dysfunction and symptom sets: a review of fMRI studies in youth with conduct problems. *Neuroscience & Biobehavioral Reviews, 91*, 69-90. https://doi.org/10.1016/j.neubiorev.2016.10.022

Blair, R. J. (2003). Neurobiological basis of psychopathy. *British Journal of Psychiatry, 182*, 5-7. https://doi.org/10.1192/bjp.182.1.5

Blair, R. J. (2016). The neurobiology of impulsive aggression. *Journal of Child and Adolescent psychopharmacology, 26*, 4-9. https://doi.org/10.1089/cap.2015.0088

Blonigen, D. M., Hicks, B. M., Krueger, R. F., Patrick, C. J., & Iacono, W. G. (2005). Psychopathic personality traits: heritability and genetic overlap with internalizing and externalizing

psychopathology. *Psychological Medicine, 35*, 637-648. https://doi.org/10.1017/s0033291704004180

Bohnert, A. M., Crnic, K. A., & Lim, K. G. (2003). Emotional competence and aggressive behavior in school-age children. *Journal of Abnormal Child Psychology, 31*, 79-91. https://doi.org/10.1023/a:1021725400321

Böhnke, R., Bertsch, K., Kruk, M. R., Richter, S., & Naumann, E. (2010). Exogenous cortisol enhances aggressive behavior in women, but not in men. *Psychoneuroendocrinology, 35*, 1034-1044. https://doi.org/10.1016/j.psyneuen.2010.01.004

Bond, A., & Lader, M. (1986a). A method to elicit aggressive feelings and behaviour via provocation. *Biological Psychology, 22*, 69-79. https://doi.org/10.1016/0301-0511(86)90021-9

Bond, A., & Lader, M. (1986b). The relationship between induced behavioural aggression and mood after the consumption of two doses of alcohol. *British Journal of Addiction, 81*, 65-75. https://doi.org/10.1111/j.1360-0443.1986.tb00296.x

Bos, P. A., Hofman, D., Hermans, E. J., Montoya, E. R., Baron-Cohen, S., & van Honk, J. (2016). Testosterone reduces functional connectivity during the 'Reading the Mind in the Eyes' Test. *Psychoneuroendocrinology, 68*, 194-201. https://doi.org/10.1016/j.psyneuen.2016.03.006

Brislin, S. J., Patrick, C. J., Flor, H., Nees, F., Heinrich, A., Drislane, L. E., Yancey, J. R., Banaschewski, T., Bokde, A. L., & Bromberg, U. (2019). Extending the construct network of trait disinhibition to the neuroimaging domain: validation of a bridging scale for use in the European IMAGEN project. *Assessment, 26*, 567-581. https://doi.org/10.1177/1073191118759748

Brosschot, J. F. (2010). Markers of chronic stress: prolonged physiological activation and (un) conscious perseverative cognition. *Neuroscience & Biobehavioral Reviews, 35*, 46-50. https://doi.org/10.1016/j.neubiorev.2010.01.004

Brown, G. L., Goodwin, F. K., Ballenger, J. C., Goyer, P. F., & Major, L. F. (1979). Aggression in humans correlates with cerebrospinal fluid amine metabolites. *Psychiatry Research, 1*, 131-139. https://doi.org/10.1016/0165-1781(79)90053-2

Buchanan, T. W., al'Absi, M., & Lovallo, W. R. (1999). Cortisol fluctuates with increases and decreases in negative affect. *Psychoneuroendocrinology, 24*, 227-241. https://doi.org/10.1016/S0306-4530(98)00078-X

Buades-Rotger, M., Engelke, C., Beyer, F., Keevil, B. G., Brabant, G., & Krämer, U. M. (2016). Endogenous testosterone is associated with lower amygdala reactivity to angry faces and reduced aggressive behavior in healthy young women. *Scientific Reports, 6*, 38538. https://doi.org/10.1038/srep38538

Buckholtz, J. W., Treadway, M. T., Cowan, R. L., Woodward, N. D., Benning, S. D., Li, R., Ansari, M. S., Baldwin, R. M., Schwartzman, A. N., Shelby, E. S., Smith, C. E., Cole, D., Kessler, R. M., & Zald, D. H. (2010). Mesolimbic dopamine reward system hypersensitivity in individuals with psychopathic traits. *Nature Neuroscience, 13*, 419-421. https://doi.org/10.1038/nn.2510

Bufkin, J. L., & Luttrell, V. R. (2005). Neuroimaging studies of aggressive and violent behavior: current findings and implications for criminology and criminal justice. *Trauma, Violence, & Abuse, 6*, 176-191. https://doi.org/10.1177/1524838005275089

Bunce, S. C., Larsen, R. J., & Cruz, M. (1993). Individual differences in the excitation transfer effect. *Personality and Individual Differences, 15*, 507-514. https://doi.org/10.1016/0191-8869(93)90333-X

Bushman, B. J. (1995). Moderating role of trait aggressiveness in the effects of violent media on aggression. *Journal of Personality and Social Psychology, 69*, 950-960. http://dx.doi.org/10.1037/0022-3514.69.5.950

Bushman, B. J. (2002). Does venting anger feed or extinguish the flame? catharsis, rumination, distraction, anger, and aggressive responding. *Personality and Social Psychology Bulletin, 28*, 724-731. https://doi.org/10.1177/0146167202289002

Bushman, B. J., & Anderson, C. A. (2002). Violent video games and hostile expectations: a test of the general aggression model. *Personality and Social Psychology Bulletin, 28*, 1679-1686. https://doi.org/10.1177/014616702237649

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology, 63*, 452-459. http://dx.doi.org/10.1037/0022-3514.63.3.452

Calkins, S. D., Gill, K. L., Johnson, M. C., & Smith, C. L. (1999). Emotional reactivity and emotional regulation strategies as predictors of social behavior with peers during toddlerhood. *Social Development, 8*, 310-334. https://doi.org/10.1111/1467-9507.00098

Cappadocia, M. C., Desrocher, M., Pepler, D., & Schroeder, J. H. (2009). Contextualizing the neurobiology of conduct disorder in an emotion dysregulation framework. *Clinical Psychology Review, 29*, 506-518. https://doi.org/10.1016/j.cpr.2009.06.001

Carré, J. M., & McCormick, C. M. (2008). Aggressive behavior and change in salivary testosterone concentrations predict willingness to engage in a competitive task. *Hormones and Behavior, 54*, 403-409. https://doi.org/10.1016/j.yhbeh.2008.04.008

Castellanos-Ryan, N., & Séguin, J. R. (2016). Prefrontal and anterior cingulate cortex mechanisms of impulsivity. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *The Oxford handbook of externalizing spectrum disorders* (pp. 201-219). Oxford: Oxford University Press

Cherek, D. R. (1981). Effects of smoking different doses of nicotine on human aggressive behavior. *Psychopharmacology, 75*, 339-345. https://doi.org/10.1007/BF00435849

Chester, D. S. (2019). Beyond the aggregate score: using multilevel modeling to examine trajectories of laboratory-measured aggression. *Aggressive Behavior*, 45, 498-506. https://doi.org/10.1002/ab.21837

Chester, D. S., DeWall, C. N., Derefinko, K. J., Estus, S., Peters, J. R., Lynam, D. R., & Jiang, Y. (2015). Monoamine oxidase A (MAOA) genotype predicts greater aggression through impulsive reactivity to negative affect. *Behavioural Brain Research, 283*, 97-101. https://doi.org/10.1016/j.bbr.2015.01.034

Chester, D. S., & Lasko, E. N. (2018). Validating a standardized approach to the Taylor Aggression Paradigm. *Social Psychological and Personality Science*, 10, 1-12. https://doi.org/10.1177/1948550618775408

Chrousos, G. P. (1998). Stressors, stress, and neuroendocrine integration of the adaptive response: The 1997 Hans Selye Memorial Lecture. *Annals of the New York Academy of Sciences, 851*, 311-335. https://doi.org/10.1111/j.1749-6632.1998.tb09006.x

Chrousos, G. P., & Gold, P. W. (1992). The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis. *Journal of the American Medical Association, 267*, 1244-1252. https://doi.org/10.1001/jama.1992.03480090092034

Cleare, A. J. (2003). The neuroendocrinology of chronic fatigue syndrome. *Endocrine Reviews, 24*, 236-252. https://doi.org/10.1210/er.2002-0014

Clow, A., Hucklebridge, F., & Thorn, L. (2010). The cortisol awakening response in context. *International Review of Neurobiology, 93*, 153-175. https://doi.org/10.1016/S0074-7742(10)93007-9

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. *Biological Psychiatry, 62*, 168-178. https://doi.org/10.1016/j.biopsych.2006.08.024

Coccaro, E. F., Sripada, C. S., Yanowitch, R. N., & Phan, K. L. (2011). Corticolimbic function in impulsive aggressive behavior. *Biological Psychiatry, 69*, 1153-1159. https://doi.org/10.1016/j.biopsych.2011.02.032

Cohen, R. A., Grieve, S., Hoth, K. F., Paul, R. H., Sweet, L., Tate, D., Gunstad, J., Stroud, L., McCaffery, J., & Hitsman, B. (2006). Early life stress and morphometry of the adult anterior cingulate cortex and caudate nuclei. *Biological Psychiatry, 59*, 975-982. https://doi.org/10.1016/j.biopsych.2005.12.016

Couture, S., Brown, T. G., Ouimet, M. C., Gianoulakis, C., Tremblay, J., & Carbonneau, R. (2008). Hypothalamic-pituitary-adrenal axis response to stress in man DUI recidivists. *Accident Analysis & Prevention, 40*, 246-253. https://doi.org/10.1016/j.aap.2007.06.003

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*, 74. https://doi.org/10.1037/0033-2909.115.1.74

Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms in reactive and proactive aggression. *Child Development, 67*, 993-1002. https://doi.org/10.1111/j.1467-8624.1996.tb01778.x

Cubillo, A., Halari, R., Ecker, C., Giampietro, V., Taylor, E., & Rubia, K. (2010). Reduced activation and inter-regional functional connectivity of fronto-striatal networks in adults with childhood Attention-Deficit Hyperactivity Disorder (ADHD) and persisting symptoms during tasks of motor inhibition and cognitive switching. *Journal of Psychiatric Research, 44*, 629-639. https://doi.org/10.1016/j.jpsychires.2009.11.016

Dahm, A.-S., Schmierer, P., Veer, I. M., Streit, F., Görgen, A., Kruschwitz, J., Wüst, S., Kirsch, P., Walter, H., & Erk, S. (2017). The burden of conscientiousness? examining brain activation and cortisol response during social evaluative stress. *Psychoneuroendocrinology, 78*, 48-56.

Dambacher, F., Schuhmann, T., Lobbestael, J., Arntz, A., Brugman, S., & Sack, A. T. (2015). No effects of bilateral tDCS over inferior frontal gyrus on response inhibition and aggression. *PLoS One, 10*, e0132170. https://doi.org/10.1371/journal.pone.0132170

Deater-Deckard, K., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1998). Multiple risk factors in the development of externalizing behavior problems: group and individual differences. *Development and Psychopathology, 10*, 469-493. https://doi.org/10.1017/S0954579498001709

De Kloet, E. R., Joëls, M., & Holsboer, F. (2005). Stress and the brain: from adaptation to disease. *Nature Reviews Neuroscience, 6*, 463-475. https://doi.org/10.1038/nrn1683

Dedovic, K., Duchesne, A., Andrews, J., Engert, V., & Pruessner, J. C. (2009). The brain and the stress axis: the neural correlates of cortisol regulation in response to stress. *NeuroImage, 47*, 864-871. https://doi.org/10.1016/j.neuroimage.2009.05.074

Dekkers, T. J., van Rentergem, J. A. A., Meijer, B., Popma, A., Wagemaker, E., & Huizenga, H. M. (2019). A meta-analytical evaluation of the dual-hormone hypothesis: does cortisol moderate the relationship between testosterone and status, dominance, risk taking, aggression, and psychopathy? *Neuroscience & Biobehavioral Reviews, 96*, 250-271. https://doi.org/10.1016/j.neubiorev.2018.12.004

Denson, T. F., Mehta, P. H., & Tan, D. H. (2013). Endogenous testosterone and cortisol jointly influence reactive aggression in women. *Psychoneuroendocrinology, 38*, 416-424. https://doi.org/10.1016/j.psyneuen.2012.07.003

Devika, D., Tuvblad, C., & Baker, L. A. (2016). Behavioral genetics of the externalizing spectrum. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *The Oxford handbook of externalizing spectrum disorders* (pp. 94). New York: Oxford University Press.

DeYoung, C. G., Peterson, J. B., Séguin, J. R., & Tremblay, R. E. (2008). Externalizing behavior and the higher order factors of the Big Five. *Journal of Abnormal Psychology, 117*, 947-953. https://doi.org/10.1037/a0013742

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin, 130*, 355-391. http://dx.doi.org/10.1037/0033-2909.130.3.355

Dmitrieva, T. N., Oades, R. D., Hauffa, B. P., & Eggers, C. (2001). Dehydroepiandrosterone sulphate and corticotropin levels are high in young man patients with conduct disorder: comparisons for growth factors, thyroid and gonadal hormones. *Neuropsychobiology, 43*, 134-140. https://doi.org/10.1159/000054881

Dollard, J., Miller, N. E., Doob, L. W., Mowrer, O. H., & Sears, R. R. (1939). *Frustration and aggression*. New Haven: Yale Univ. Press.

Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., Moffitt, T. E., & Caspi, A. (2005). Low self-esteem is related to aggression, antisocial behavior, and delinquency. *Psychological Science, 16*, 328-335. https://doi.org/10.1111/j.0956-7976.2005.01535.x

Dougherty, D. D., Rauch, S. L., Deckersbach, T., Marci, C., Loh, R., Shin, L. M., . . . Fava, M. (2004). Ventromedial prefrontal cortex and amygdala dysfunction during an angerinduction positron emission tomography study in patients with major depressivedisorder with anger attacks. *Archives of General Psychiatry, 61*, 795-804. https://doi.org/10.1001/archpsyc.61.8.795

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage, 25*, 1325-1335. https:/doi.org/10.1016/j.neuroimage.2004.12.034

Eisenbarth, H., Castellino, N., Alpers, G. W., Kirsch, P., & Flor, H. (2012). *The German version of Triarchic Psychopathy Measure. Unpublished report*. Regensburg.

Elson, M., Mohseni, M. R., Breuer, J., Scharkow, M., & Quandt, T. (2014). Press CRTT to measure aggressive behavior: the unstandardized use of the competitive reaction time task in aggression research. *Psychological Assessment, 26*, 419-432. http://dx.doi.org/10.1037/a0035569

Eron, L. D., Gentry, J. H., & Schlegel, P. (1994). *Reason to hope: a psychosocial perspective on violence and youth*. Washington, DC: American Psychological Association.

Ersche, K. D., Barnes, A., Jones, P. S., Morein-Zamir, S., Robbins, T. W., & Bullmore, E. T. (2011). Abnormal structure of frontostriatal brain systems is associated with aspects of impulsivity and compulsivity in cocaine dependence. *Brain, 134*, 2013-2024. https://doi.org/10.1093/brain/awr138

Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences, 15*, 85-93. https://doi.org/10.1016/j.tics.2010.11.004

Etkin, A., & Wager, T. D. (2007). Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *American Journal of Psychiatry, 164*, 1476-1488. https://doi.org/10.1176/appi.ajp.2007.07030504

Fairchild, G., Hawes, D. J., Frick, P. J., Copeland, W. E., Odgers, C. L., Franke, B., Freitag, C. M., & De Brito, S. A. (2019). Conduct disorder. *Nature Reviews Disease Primers, 5*, 43. https://doi.org/10.1038/s41572-019-0095-y

Fairchild, G., van Goozen, S. H., Stollery, S. J., Brown, J., Gardiner, J., Herbert, J., & Goodyer, I. M. (2008). Cortisol diurnal rhythm and stress reactivity in man adolescents with early-onset or adolescence-onset conduct disorder. *Biological Psychiatry, 64*, 599-606. http://doi.org/10.1016/j.biopsych.2008.05.022

Fanning, J. R., Keedy, S., Berman, M. E., Lee, R., & Coccaro, E. F. (2017). Neural correlates of aggressive behavior in real time: a review of fMRI studies of laboratory reactive aggression. *Current Behavioral Neuroscience Reports, 4*, 138-150. https://doi.org/10.1007/s40473-017-0115-8

Ferguson, C. J., Rueda, S. M., Cruz, A. M., Ferguson, D. E., Fritz, S., & Smith, S. M. (2008). Violent video games and aggression: causal relationship or byproduct of family violence and intrinsic violence motivation? *Criminal Justice and Behavior, 35*, 311-332. https://doi.org/10.1177/0093854807311719

Ferguson, C. J., Smith, S., Miller-Stratton, H., Fritz, S., & Heinrich, E. (2008). Aggression in the laboratory: Problems with the validity of the modified Taylor Competitive Reaction Time Test as a measure of aggression in media violence studies. *Journal of Aggression, Maltreatment & Trauma, 17*, 118-132. https://doi.org/10.1080/10926770802250678

Fink, G. (2007). *Encyclopedia of stress* (2nd revised ed.). San Diego: Academic Press.

Fishbein, D. H., Eldreth, D. L., Hyde, C., Matochik, J. A., London, E. D., Contoreggi, C., Kurian, V., Kimes, A. S., Breeden, A., & Grant, S. (2005). Risky decision making and the anterior cingulate cortex in abstinent drug abusers and nonusers. *Cognitive Brain Research, 23*, 119-136. https://doi.org/10.1016/j.cogbrainres.2004.12.010

Floden, D., & Stuss, D. T. (2006). Inhibitory control is slowed in patients with right superior medial frontal damage. *Journal of Cognitive Neuroscience, 18*, 1843-1849. https://doi.org/10.1162/jocn.2006.18.11.1843

Foell, J., Brislin, S. J., Strickland, C. M., Seo, D., Sabatinelli, D., & Patrick, C. J. (2016). Externalizing proneness and brain response during pre-cuing and viewing of emotional pictures. *Social Cognitive and Affective Neuroscience, 11*, 1102-1110. https://doi.org/10.1093/scan/nsv080

Folkman, S., & Lazarus, R. S. (1984). *Stress, appraisal, and coping*. New York: Springer.

Fox, J., & Weisberg, S. (2019). *An (R) companion to applied regression*. California: Sage.

Friston, K. (2009). Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biology, 7*, e1000033. https://doi.org/10.1371/journal.pbio.1000033

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. https://doi.org/10.1038/npp.2015.141

Gatzke-Kopp, L. M. (2011). The canary in the coalmine: the sensitivity of mesolimbic dopamine to environmental adversity during development. *Neuroscience & Biobehavioral Reviews, 35*, 794-803. https://doi.org/10.1016/j.neubiorev.2010.09.013

Geen, R. (2001). *Human aggression* (2nd ed.). Philadelphia: Open University Press.

Geniole, S. N., Busseri, M. A., & McCormick, C. M. (2013). Testosterone dynamics and psychopathic personality traits independently predict antagonistic behavior towards the perceived loser of a competitive interaction. *Hormones and Behavior, 64*, 790-798. https://doi.org/10.1016/j.yhbeh.2013.09.005

Geniole, S. N., & Carré, J. M. (2018). Human social neuroendocrinology: review of the rapid effects of testosterone. *Hormones and Behavior, 104*, 192-205. https://doi.org/10.1016/j.yhbeh.2018.06.001

Geniole, S. N., Carré, J. M., & McCormick, C. M. (2011). State, not trait, neuroendocrine function predicts costly reactive aggression in men after social exclusion and inclusion. *Biological Psychology, 87*, 137-145. https://doi.org/10.1016/j.biopsycho.2011.02.020

Giancola, P. R., Godlaski, A. J., & Roth, R. M. (2012). Identifying component-processes of executive functioning that serve as risk factors for the alcohol-aggression relation.

Psychology of Addictive Behaviors, 26, 201. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3288328/pdf/nihms323884.pdf

Giancola, P. R., & Parrott, D. J. (2008). Further evidence for the validity of the Taylor Aggression Paradigm. *Aggressive Behavior, 34*, 214-229. https://doi.org/10.1002/ab.20235

Giancola, P. R., & Zeichner, A. (1995). Construct validity of a competitive reaction-time aggression paradigm. *Aggressive Behavior, 21*, 199-204. https://doi.org/10.1002/1098-2337(1995)21:3<199::AID-AB2480210303>3.0.CO;2-Q

Goldfarb, E. V., Seo, D., & Sinha, R. (2019). Sex differences in neural stress responses and correlation with subjective stress and stress regulation. *Neurobiology of Stress, 11*, 100177. https://doi.org/10.1016/j.ynstr.2019.100177

Gordon, H. L., Baird, A. A., & End, A. (2004). Functional differences among those high and low on a trait measure of psychopathy. *Biological Psychiatry, 56*, 516-521. https://doi.org/10.1016/j.biopsych.2004.06.030

Green, P., & MacLeod, C. J. (2016). SIMR: a R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7*. https://doi.org/10.1111/2041-210X.12504

Gross, J. J., & Munoz, R. F. (1995). Emotion regulation and mental-health. *Clinical Psychology Science and Practice, 2*, 151-164. https://doi.org/10.1111/j.1468-2850.1995.tb00036.x

Gustafson, R. (1985). Alcohol and aggression: a validation study of the Taylor Aggression Paradigm. *Psychological Reports, 57*, 667-676. https://doi.org/10.2466/pr0.1985.57.2.667

Hart, H., Radua, J., Nakao, T., Mataix-Cols, D., & Rubia, K. (2013). Meta-analysis of functional magnetic resonance imaging studies of inhibition and attention in attention-deficit/hyperactivity disorder: exploring task-specific, stimulant medication, and age effects. *JAMA Psychiatry, 70*, 185-198. https://doi.org/10.1001/jamapsychiatry.2013.277

Henze, G.-I., Konzok, J., Kreuzpointner, L., Bärtl, C., Peter, H., Giglberger, M., Streit, F., Kudielka, B. M., Kirsch, P., & Wüst, S. (2020). Increasing deactivation of limbic structures over

psychosocial stress exposure time. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 5, 697-704. https://doi.org/10.1016/j.bpsc.2020.04.002

Herman, J. P., Figueiredo, H., Mueller, N. K., Ulrich-Lai, Y., Ostrander, M. M., Choi, D. C., & Cullinan, W. E. (2003). Central mechanisms of stress integration: hierarchical circuitry controlling hypothalamo–pituitary–adrenocortical responsiveness. *Frontiers in Neuroendocrinology, 24*, 151-180. https://doi.org/10.1016/j.yfrne.2003.07.001

Herman, J. P., Ostrander, M. M., Mueller, N. K., & Figueiredo, H. (2005). Limbic system mechanisms of stress regulation: hypothalamo-pituitary-adrenocortical axis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 29*, 1201-1213. https://doi.org/10.1016/j.pnpbp.2005.08.006

Hermans, E. J., Ramsey, N. F., & van Honk, J. (2008). Exogenous testosterone enhances responsiveness to social threat in the neural circuitry of social aggression in humans. *Biological Psychiatry, 63*, 263-270. https://doi.org/10.1016/j.biopsych.2007.05.013

Herrmann-Lingen, C., Buss, U., & Snaith, R. P. (2011). *Hospital anxiety and depression scale: - Deutsche Version: HADS-D* Bern: Huber.

Herts, K. L., McLaughlin, K. A., & Hatzenbuehler, M. L. (2012). Emotion dysregulation as a mechanism linking stress exposure to adolescent aggressive behavior. *Journal of Abnormal Child Psychology, 40*, 1111-1122. https://doi.org/10.1007/s10802-012-9629-4

Hesser, H. (2015). Modeling individual differences in randomized experiments using growth models: recommendations for design, statistical analysis and reporting of results of internet interventions. *Internet Interventions, 2*, 110-120. https://doi.org/10.1016/j.invent.2015.02.003

Heubrock, D., & Petermann, F. (2008). *Kurzfragebogen zur Erfassung von Aggressivitätsfaktoren: K-FAF*. Göttingen: Hogrefe.

Hicks, B. M., Blonigen, D. M., Kramer, M. D., Krueger, R. F., Patrick, C. J., Iacono, W. G., & McGue, M. (2007). Gender differences and developmental change in externalizing

disorders from late adolescence to early adulthood: a longitudinal twin study. *Journal of Abnormal Psychology, 116*, 433-447. https://doi.org/10.1037/0021-843X.116.3.433

Hinshaw, S. P., & Beauchaine, T. P. (2016). The developmental psychopahology perspective on externalizing behavior dimensions and externalizing disorders. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *The Oxford handbook of externalizing spectrum disorders* (pp. 94). New York: Oxford University Press.

Hoaken, P. N. S., & Pihl, R. O. (2000). The effects of alcohol on aggressive responses in men and women. *Alcohol and Alcoholism, 35*, 471-477. http://dx.doi.org/10.1093/alcalc/35.5.471

Huesmann, L. R. (1998). *The role of social information processing and cognitive schema in the acquisition and maintenance of habitual aggressive behavior.* In R. G. Geen & E. Donnerstein (Eds.), *Human aggression: Theories, research, and implications for social policy* (p. 73–109). Academic Press. https://doi.org/10.1016/B978-012278805-5/50005-5

Hyatt, C. S., Chester, D. S., Zeichner, A., & Miller, J. D. (2019). Analytic flexibility in laboratory aggression paradigms: relations with personality traits vary (slightly) by operationalization of aggression. *Aggressive Behavior*, 45, 377-388. https://doi.org/10.1002/ab.21830

Hyatt, C. S., Zeichner, A., & Miller, J. D. (2019). Laboratory aggression and personality traits: a meta-analytic review. *Psychology of violence, 9*, 675-689. https://doi.org/10.1037/vio0000236

Hyde, J. S. (1984). How large are gender differences in aggression? a developmental meta-analysis. *Developmental Psychology, 20*, 722-736. http://dx.doi.org/10.1037/0012-1649.20.4.722

Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., & van Anders, S. M. (2018). The future of sex and gender in psychology: five challenges to the gender binary. *American Psychologist, 74*, 171-193. http://dx.doi.org/10.1037/amp0000307

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry*, 7, 748-51.

John, O. P., Caspi, A., Robins, R. W., Moffitt, T. E., & Stouthamer-Loeber, M. (1994). The "little five": Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Development, 65*, 160-178. https://doi.org/10.1111/j.1467-8624.1994.tb00742.x

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics, 53*, 983-997. http://dx.doi.org/10.2307/2533558

Kirschbaum, C., Pirke, K.-M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28*, 76-81. https://doi.org/10.1159/000119004

Kobak, R., Zajac, K., & Levine, S. (2009). Cortisol and antisocial behavior in early adolescence: the role of gender in an economically disadvantaged sample. *Development and Psychopathology, 21*, 579-591. https://doi.org/10.1017/S0954579409000315

Kogan-Goloborodko, O., Brügmann, E., Repple, J., Habel, U., & Clemens, B. (2016). Experimentally assessed reactive aggression in borderline personality disorder. *PLoS One, 11*, e0166737. http://dx.doi.org/10.1371%2Fjournal.pone.0166737

Kogler, L., Gur, R. C., & Derntl, B. (2015). Sex differences in cognitive regulation of psychosocial achievement stress: brain and behavior. *Human Brain Mapping, 36*, 1028-1042.

Konzok, J., Henze, G.-I., Peter, H., Giglberger, M., Bärtl, C., Massau, C., Kärgel, C., Schiffer, B., Eisenbarth, H., Wüst, S., & Kudielka, B. M. (2020). Externalizing behavior in healthy young adults is associated with lower cortisol responses to acute stress and altered neural activation in the dorsal striatum. Manuscript under review in *Psychophysiology.*

Konzok, J., Kreuzpointner, L., Henze, G.-I., Wagels, L., Kärgel, C., Weidacker, K., Schiffer, B., Eisenbarth, H., Wüst, S., & Kudielka, B. M. (2020). Validation of a monetary Taylor Aggression Paradigm: associations with trait aggression and role of provocation sequence. *Journal of Experimental Social Psychology, 88*, 103960. https://doi.org/10.1016/j.jesp.2020.103960

Koob, G. F., & Volkow, N. D. (2010). Neurocircuitry of addiction. *Neuropsychopharmacology, 35*, 217-238. https://doi.org/10.1038/npp.2009.110

Körner, A., Geyer, M., Roth, M., Drapeau, M., Schmutzer, G., Albani, C., Schumann, S., & Brähler, E. (2008). Personality assessment with the NEO-five-factor inventory: the 30-item-short-version (NEO-FFI-30). *Psychotherapie, Psychosomatik, Medizinische Psychologie, 58*, 238-245. https://10.1055/s-2007-986199

Krämer, U. M., Büttner, S., Roth, G., & Münte, T. F. (2008). Trait aggressiveness modulates neurophysiological correlates of laboratory-induced reactive aggression in humans. *Journal of Cognitive Neuroscience, 20*, 1464-1477. http://dx.doi.org/10.1162/jocn.2008.20103

Krämer, U. M., Jansma, H., Tempelmann, C., & Münte, T. F. (2007). Tit-for-tat: the neural basis of reactive aggression. *NeuroImage, 38*, 203-211. https://doi.org/10.1016/j.neuroimage.2007.07.029

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*, e27668. https://doi.org/10.1371/journal.pone.0027668

Krieger, F. V., & Stringaris, A. (2016). Temperament and vulnerability to externalizing behavior. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *The Oxford handbook of externalizing spectrum disorders* (pp. 170-179). New York: Oxford University Press.

Krigolson, O. E., Hassall, C. D., & Handy, T. C. (2014). How we learn to make decisions: rapid propagation of reinforcement learning prediction errors in humans. *Journal of Cognitive Neuroscience, 26*, 635-644. http://doi.org/10.1162/jocn_a_00509

Krueger, R. F., Hicks, B. M., Patrick, C. J., Carlson, S. R., Iacono, W. G., & McGue, M. (2002). Etiologic connections among substance dependence, antisocial behavior, and personality: modeling the externalizing spectrum. *Journal of Abnormal Psychology, 111*, 411-424. http://dx.doi.org/10.1037/11855-003

Krueger, R. F., Markon, K. E., Patrick, C. J., Benning, S. D., & Kramer, M. D. (2007). Linking antisocial behavior, substance use, and personality: an integrative quantitative model of

the adult externalizing spectrum. *Journal of Abnormal Psychology, 116*, 645-666. https://doi.org/10.1037/0021-843X.116.4.645

Krueger, R. F., Markon, K. E., Patrick, C. J., & Iacono, W. G. (2005). Externalizing psychopathology in adulthood: a dimensional-spectrum conceptualization and its implications for DSM-V. *Journal of Abnormal Psychology, 114*, 537-550. http://dx.doi.org/10.1037/0021-843X.116.4.645

Kudielka, B. M., Hellhammer, D. H., & Kirschbaum, C. (2007). Ten Years of Research with the Trier Social Stress Test - Revisited. In E. Harmon-Jones & P. Winkielman (Eds.), *Social neuroscience: integrating biological and psychological explanations of social behavior*. New York: The Guilford Press.

Kudielka, B. M., Hellhammer, D. H., & Wüst, S. (2009). Why do we respond so differently? reviewing determinants of human salivary cortisol responses to challenge. *Psychoneuroendocrinology, 34*, 2-18. https://doi.org/10.1016/j.psyneuen.2008.10.004

Kudielka, B. M., & Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: a review. *Biological Psychology, 69*, 113-132. https://doi.org/10.1016/j.biopsycho.2004.11.009

Kudielka, B. M., & Kirschbaum, C. (2007). Biological bases of the stress response. In M. al'Absi (Ed.), *Stress and addiction: biological and psychological mechanisms* (pp. 3-19). Amsterdam: Elsevier.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82*, 26. https://doi.org/10.18637/jss.v082.i13

Lawrence, C., & Hutchinson, L. (2014). The impact of non-aggressive behaviour early in aggressive interactions: sex differences in direct and indirect aggression in response to provocation. *British Journal of Psychology, 105*, 127-144. https://doi.org/10.1111/bjop.12020

Lederbogen, F., Kirsch, P., Haddad, L., Streit, F., Tost, H., Schuch, P., Wüst, S., Pruessner, J. C., Rietschel, M., Deuschle, M., & Meyer-Lindenberg, A. (2011). City living and urban

upbringing affect neural social stress processing in humans. *Nature, 474*, 498-501. https://doi.org/10.1038/nature10190

Lee, R., Kavoussi, R. J., & Coccaro, E. F. (2008). Placebo-controlled, randomized trial of fluoxetine in the treatment of aggression in man intimate partner abusers. *International Clinical Psychopharmacology, 23*, 337-341. https://doi.org/10.1097/YIC.0b013e32830fbdd2

Levine, S., & Ursin, H. (1991). What is stress? In R. M. Brown, G. F. Koob, & C. Rivier (Eds.), *Stress, neurobiology and neuroendocrinology* (pp. 3-21). New York: Dekker.

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*, 470-478. https://doi.org/10.1016/j.neuroimage.2006.09.028

Lüdecke, D. (2020). Sjstats: statistical functions for regression models (version 0.17.9). https://CRAN.R-project.org/package=sjstats

Markon, K. E., & Krueger, R. F. (2005). Categorical and continuous models of liability to externalizing disorders: a direct comparison in NESARC. *Archives of General Psychiatry, 62*, 1352-1359. https://doi.org/10.1001/archpsyc.62.12.1352

Mason, J. W. (1968). A review of psychoendocrine research on the pituitary-adrenal cortical system. *Psychosomatic Medicine, 30*, 576-607.

McBurnett, K., Raine, A., Stouthamer-Loeber, M., Loeber, R., Kumar, A. M., Kumar, M., & Lahey, B. B. (2005). Mood and hormone responses to psychological challenge in adolescent men with conduct problems. *Biological Psychiatry, 57*, 1109-1116. http://doi.org/10.1016/j.biopsych.2005.01.041

McCloskey, M. S., & Coccaro, E. F. (2003). Laboratory measures: the Taylor Aggression Paradigm. In E. F. Coccaro (Ed.), *Aggression: Psychiatric Assessment and Treatment* (1st ed., pp. 195-214). New York: CRC Press Taylor & Francis Group.

McEwen, B. S. (1998a). Protective and damaging effects of stress mediators. *New England Journal of Medicine, 338*, 171-179. https://doi.org/10.1056/NEJM199801153380307

McEwen, B. S. (1998b). Stress, adaptation, and disease: allostasis and allostatic load. *Annals of the New York Academy of Sciences, 840*, 33-44. https://doi.org/10.1111/j.1749-6632.1998.tb09546.x

McEwen, B. S., & Stellar, E. (1993). Stress and the individual: mechanisms leading to disease. *Archives of Internal Medicine, 153*, 2093-2101. https://doi.org/10.1001/archinte.1993.00410180039004

McLaughlin, K. A., Hatzenbuehler, M. L., Mennin, D. S., & Nolen-Hoeksema, S. (2011). Emotion dysregulation and adolescent psychopathology: a prospective study. *Behaviour Research and Therapy, 49*, 544-554. https://doi.org/10.1016/j.brat.2011.06.003

Mehta, P. H., & Beer, J. (2010). Neural mechanisms of the testosterone–aggression relation: the role of orbitofrontal cortex. *Journal of Cognitive Neuroscience, 22*, 2357-2368. https://doi.org/10.1162/jocn.2009.21389

Mehta, P. H., & Josephs, R. A. (2010). Testosterone and cortisol jointly regulate dominance: evidence for a dual-hormone hypothesis. *Hormones and Behavior, 58*, 898-906. https://doi.org/10.1016/j.yhbeh.2010.08.020

Miller, R., Plessow, F., Kirschbaum, C., & Stalder, T. (2013). Classification criteria for distinguishing cortisol responders from nonresponders to psychosocial stress: evaluation of salivary cortisol pulse detection in panel designs. *Psychosomatic Medicine, 75*, 832-840. https://doi.org/10.1097/PSY.0000000000000002

Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., & Uğurbil, K. (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magnetic Resonance in Medicine, 63*, 1144-1153. https://doi.org/10.1002/mrm.22361

Mostofsky, S. H., & Simmonds, D. J. (2008). Response inhibition and response selection: two sides of the same coin. *Journal of Cognitive Neuroscience, 20*, 751-761. https://doi.org/10.1162/jocn.2008.20500

Nachev, P., Wydell, H., O'neill, K., Husain, M., & Kennard, C. (2007). The role of the pre-supplementary motor area in the control of action. *NeuroImage, 36*, 155-163. https://doi.org/10.1016/j.neuroimage.2007.03.034

Noack, H., Nolte, L., Nieratschker, V., Habel, U., & Derntl, B. (2019). Imaging stress: an overview of stress induction methods in the MR scanner. *Journal of Neural Transmission*, 126, 1187-1202. https://doi.org/10.1007/s00702-018-01965-y

Norman, A. L., Pulido, C., Squeglia, L. M., Spadoni, A. D., Paulus, M. P., & Tapert, S. F. (2011). Neural activation during inhibition predicts initiation of substance use in adolescence. *Drug and Alcohol Dependence, 119*, 216-223. https://doi.org/10.1016/j.drugalcdep.2011.06.019

O'Leary, M. M., Loney, B. R., & Eckel, L. A. (2007). Gender differences in the association between psychopathic personality traits and cortisol response to induced stress. *Psychoneuroendocrinology, 32*, 183-191. https://doi.org/10.1016/j.psyneuen.2006.12.004

Oberlin, B. G., Dzemidzic, M., Bragulat, V., Lehigh, C. A., Talavage, T., O'Connor, S. J., & Kareken, D. A. (2012). Limbic responses to reward cues correlate with antisocial trait density in heavy drinkers. *NeuroImage, 60*, 644-652. https://doi.org/10.1016/j.neuroimage.2011.12.043

Oswald, L. M., Wong, D. F., McCaul, M., Zhou, Y., Kuwabara, H., Choi, L., Brasic, J., & Wand, G. S. (2005). Relationships among ventral striatal dopamine release, cortisol secretion, and subjective responses to amphetamine. *Neuropsychopharmacology, 30*, 821-832. https://doi.org/10.1038/sj.npp.1300667

Pajer, K., Tabbah, R., Gardner, W., Rubin, R. T., Kenneth Czambel, R., & Wang, Y. (2006). Adrenal androgen and gonadal hormone levels in adolescent girls with conduct disorder. *Psychoneuroendocrinology, 31*, 1245-1256. http://www.sciencedirect.com/science/article/pii/S0306453006001491

Patrick, C. J. (2010). *Operationalizing the triarchic conceptualization of psychopathy: preliminary description of brief scales for assessment of boldness, meanness, and disinhibition*. Florida State University. Tallahassee.

Patrick, C. J., Fowles, D. C., & Krueger, R. F. (2009). Triarchic conceptualization of psychopathy: developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology, 21*, 913-938. https://doi.org/10.1017/S0954579409000492

Patrick, C. J., Venables, N. C., Yancey, J. R., Hicks, B. M., Nelson, L. D., & Kramer, M. D. (2013). A construct-network approach to bridging diagnostic and physiological domains: application to assessment of externalizing psychopathology. *Journal of Abnormal Psychology, 122*, 902-916. https://doi.org/10.1037/a0032807

Picton, T. W., Stuss, D. T., Alexander, M. P., Shallice, T., Binns, M. A., & Gillingham, S. (2007). Effects of focal frontal lesions on response inhibition. *Cerebral Cortex, 17*, 826-838. https://doi.org/10.1093/cercor/bhk031

Platje, E., Popma, A., Vermeiren, R. R., Doreleijers, T. A., Meeus, W. H., van Lier, P. A., Koot, H. M., Branje, S. J., & Jansen, L. M. (2015). Testosterone and cortisol in relation to aggression in a non-clinical sample of boys and girls. *Aggressive Behavior, 41*, 478-487. https://doi.org/10.1002/ab.2158

Plichta, M. M., Vasic, N., Wolf, R. C., Lesch, K.-P., Brummer, D., Jacob, C., Fallgatter, A. J., & Grön, G. (2009). Neural hyporesponsiveness and hyperresponsiveness during immediate and delayed reward processing in adult attention-deficit/hyperactivity disorder. *Biological Psychiatry, 65*, 7-14. https://doi.org/10.1016/j.biopsych.2008.07.008

Popma, A., Jansen, L. M. C., Vermeiren, R., Steiner, H., Raine, A., Van Goozen, S. H. M., Engeland, H. v., & Doreleijers, T. A. H. (2006). Hypothalamus pituitary adrenal axis and autonomic activity during stress in delinquent man adolescents and controls. *Psychoneuroendocrinology, 31*, 948-957. https://doi.org/10.1016/j.psyneuen.2006.05.005

Preacher, K. J. (2014). Extreme groups designs. *The encyclopedia of clinical psychology*1-4. https://doi.org/10.1002/9781118625392.wbecp190

Preacher, K. J., Rucker, D. D., MacCallum, R. C., & Nicewander, W. A. (2005). Use of the extreme groups approach: a critical reexamination and new recommendations. *Psychological Methods, 10*, 178-192. https://doi.org/10.1037/1082-989X.10.2.178

Pruessner, J. C., Champagne, F., Meaney, M. J., & Dagher, A. (2004). Dopamine release in response to a psychological stress in humans and its relationship to early life maternal care: a positron emission tomography study using [11C] raclopride. *Journal of Neuroscience, 24*, 2825-2831. https://doi.org/10.1523/JNEUROSCI.3422-03.2004

Pruessner, J. C., Dedovic, K., Khalili-Mahani, N., Engert, V., Pruessner, M., Buss, C., Renwick, R., Dagher, A., Meaney, M. J., & Lupien, S. (2008). Deactivation of the limbic system during acute psychosocial stress: evidence from positron emission tomography and functional magnetic resonance imaging studies. *Biological Psychiatry, 63*, 234-240. https://doi.org/10.1016/j.biopsych.2007.04.041

Pruessner, J. C., Kirschbaum, C., Meinlschmid, G., & Hellhammer, D. H. (2003). Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology, 28*, 916-931. https://doi.org/10.1016/S0306-4530(02)00108-7

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. *Proceedings of the National Academy of Sciences, 98*, 676-682. https://doi.org/10.1073/pnas.98.2.676

Raine, A. (2013). *The psychopathology of crime: criminal behavior as a clinical disorder*. Cambridge: Elsevier.

Raine, A., Dodge, K., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., Stouthamer-Loeber, M., & Liu, J. (2006). The reactive–proactive aggression questionnaire: differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior, 32*, 159-171. https://doi.org/10.1002/ab.20115

Repple, J., Habel, U., Wagels, L., Pawliczek, C. M., Schneider, F., & Kohn, N. (2018). Sex differences in the neural correlates of aggression. *Brain Structure and Function, 223*, 4115-4124. https://doi.org/10.1007/s00429-018-1739-5

Repple, J., Pawliczek, C. M., Voss, B., Siegel, S., Schneider, F., Kohn, N., & Habel, U. (2017). From provocation to aggression: the neural network. *BMC Neuroscience, 18*, 73. https://doi.org/10.1186/s12868-017-0390-z

Revelle, W. (2019). Psych: procedures for psychological, psychometric, and personality research. https://CRAN.R-project.org/package=psych

Richardson, D. S., & Hammock, G. S. (2007). Social context of human aggression: are we paying too much attention to gender? *Aggression and Violent Behavior, 12*, 417-426. https://doi.org/10.1016/j.avb.2006.11.001

Riva, P., Gabbiadini, A., Lauro, L. J. R., Andrighetto, L., Volpato, C., & Bushman, B. J. (2017). Neuromodulation can reduce aggressive behavior elicited by violent video games. *Cognitive, Affective, & Behavioral Neuroscience, 17*, 452-459. https://doi.org/10.3758/s13415-016-0490-8

Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Research Reviews, 18*, 247-291. https://doi.org/10.1016/0165-0173(93)90013-P

Rowe, R., Maughan, B., Worthman, C. M., Costello, E. J., & Angold, A. (2004). Testosterone, antisocial behavior, and social dominance in boys: pubertal development and biosocial interaction. *Biological Psychiatry, 55*, 546-552. https://doi.org/10.1016/j.biopsych.2003.10.010

Sapolsky, R. M. (1999). Glucocorticoids, stress, and their adverse neurological effects: relevance to aging. *Experimental Gerontology, 34*, 721-732. https://doi.org/10.1016/S0531-5565(99)00047-9

Schmidt, K.-H., & Metzler, P. (1994). *Wortschatztest - WST*. Göttingen: Beltz.

Schmidt, A. F., Zimmermann, P. S., Banse, R., & Imhoff, R. (2015). Ego depletion moderates the influence of automatic and controlled precursors of reactive aggression. *Social Psychology, 46*, 132-141. https://doi.org/10.1027/1864-9335/a000233

Schneider, I., Regenbogen, C., Kohn, N., Zepf, F. D., Bubenzer-Busch, S., Schneider, F., Gur, R. C., & Habel, U. (2015). Reduced responsiveness to social provocation in autism spectrum disorder. *Autism Research, 8*, 297-306. http://dx.doi.org/10.1002/aur.1446

Schommer, N. C., Hellhammer, D. H., & Kirschbaum, C. (2003). Dissociation between reactivity of the hypothalamus-pituitary-adrenal axis and the sympathetic-adrenal-medullary system to repeated psychosocial stress. *Psychosomatic Medicine, 65*, 450-460. https://doi.org/10.1097/01.PSY.0000035721.12441.17

Schwabe, L., Tegenthoff, M., Höffken, O., & Wolf, O. T. (2013). Mineralocorticoid receptor blockade prevents stress-induced modulation of multiple memory systems in the human brain. *Biological Psychiatry, 74*, 801-808. https://doi.org/10.1016/j.biopsych.2013.06.001

Scott, L. N., Victor, S. E., Kaufman, E. A., Beeney, J. E., Byrd, A. L., Vine, V., Pilkonis, P. A., & Stepp, S. D. (2020). Affective dynamics across internalizing and externalizing dimensions of psychopathology. *Clinical Psychological Science, 8*, 412-427. https://doi.org/10.1177/2167702619898802

Selye, H. (1936). A syndrome produced by diverse nocuous agents. *Nature, 138*, 32-32. https://doi.org/10.1038/138032a0

Selye, H. (1946). The general adaptation syndrome and the diseases of adaptation. *The Journal of Clinical Endocrinology, 6*, 117-230. https://doi.org/10.1210/jcem-6-2-117

Shannon, K. E., Sauder, C., Beauchaine, T. P., & Gatzke-Kopp, L. M. (2009). Disrupted effective connectivity between the medial frontal cortex and the caudate in adolescent boys with externalizing behavior disorders. *Criminal Justice and Behavior, 36*, 1141-1157. https://doi.org/10.1177/0093854809342856

Sherrill, A. M., Magliano, J. P., Rosenbaum, A., Bell, K. M., & Wallace, P. S. (2016). Trait aggressiveness and aggressive behavior in the context of provocation and inhibition.

Journal of Aggression, Maltreatment & Trauma, 25, 487-502. https://doi.org/10.1080/10926771.2015.1121192

Sherry, J. L. (2001). The effects of violent video games on aggression: a meta-analysis. *Human Communication Research, 27*, 409-431. https://doi.org/10.1111/j.1468-2958.2001.tb00787.x

Siever, L. J. (2008). Neurobiology of aggression and violence. *American Journal of Psychiatry, 165*, 429-442. https://doi.org/10.1176/appi.ajp.2008.07111774

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). Afex: analysis of factorial experiments. https://CRAN.R-project.org/package=afex

Singmann, H., & Kellen, D. (2017). An introduction to mixed models for experimental psychology. In D. H. Spieler & E. Schumacher (Eds.), *New Methods in Neuroscience and Cognitive Psychology*: Psychology Press.

Snoek, H., Van Goozen, S. H., Matthys, W., Buitelaar, J. K., & van Engeland, H. (2004). Stress responsivity in children with externalizing behavior disorders. *Development and Psychopathology, 16*, 389-406. https://doi.org/10.1017/S0954579404044578

Spinella, M. (2007). Normative data and a short form of the Barratt Impulsiveness Scale. *International Journal of Neuroscience, 117*, 359-368. https://doi.org/10.1080/00207450600588881

Stalder T., & Kirschbaum C. (2013). Cortisol. In Gellman M.D. & Turner J.R. (Eds.), *Encyclopedia of Behavioral Medicine*. New York: Springer.

Stalder, T., Kirschbaum, C., Kudielka, B. M., Adam, E. K., Pruessner, J. C., Wüst, S., . . . Hellhammer, D. H. (2016). Assessment of the cortisol awakening response: expert consensus guidelines. *Psychoneuroendocrinology, 63*, 414-432. https://doi.org/10.1016/j.psyneuen.2015.10.010

Stadler, C., Sterzer, P., Schmeck, K., Krebs, A., Kleinschmidt, A., & Poustka, F. (2007). Reduced anterior cingulate activation in aggressive children and adolescents during affective

stimulation: association with temperament traits. *Journal of Psychiatric Research, 41*, 410-417. https://doi.org/10.1016/j.jpsychires.2006.01.006

Sterling, P., & Eyer, J. (1988). Allostasis: a new paradigm to explain arousal pathology. In S. Fisher & J. Reason (Eds.), *Handbook of life stress, cognition and health*. New York: John Wiley & Sons.

Stevens, F. L., Hurley, R. A., & Taber, K. H. (2011). Anterior cingulate cortex: unique role in cognition and emotion. *The Journal of Neuropsychiatry and Clinical Neurosciences, 23*, 121-125. https://doi.org/10.1176/jnp.23.2.jnp121

Streit, F., Haddad, L., Paul, T., Frank, J., Schafer, A., Nikitopoulos, J., Akdeniz, C., Lederbogen, F., Treutlein, J., Witt, S., Meyer-Lindenberg, A., Rietschel, M., Kirsch, P., & Wüst, S. (2014). A functional variant in the neuropeptide S receptor 1 gene moderates the influence of urban upbringing on stress processing in the amygdala. *Stress, 17*, 352-361. https://doi.org/10.3109/10253890.2014.921903

Subramani, O. S., Parrott, D. J., Latzman, R. D., & Washburn, D. A. (2019). Breaking the link: Distraction from emotional cues reduces the association between trait disinhibition and reactive physical aggression. *Aggressive behavior, 45*, 151-160. https://doi.org/10.1002/ab.21804

Sun, L., Cao, Q., Long, X., Sui, M., Cao, X., Zhu, C., Zuo, X., An, L., Song, Y., & Zang, Y. (2012). Abnormal functional connectivity between the anterior cingulate and the default mode network in drug-naive boys with attention deficit hyperactivity disorder. *Psychiatry Research: Neuroimaging, 201*, 120-127. https://doi.org/10.1016/j.pscychresns.2011.07.001

Susman, E. J. (2006). Psychobiology of persistent antisocial behavior: stress, early vulnerabilities and the attenuation hypothesis. *Neuroscience & Biobehavioral Reviews, 30*, 376-389. https://doi.org/10.1016/j.neubiorev.2005.08.002

Swaab, D. F., Bao, A.-M., & Lucassen, P. J. (2005). The stress system in the human brain in depression and neurodegeneration. *Ageing Research Reviews, 4*, 141-194. https://doi.org/10.1016/j.arr.2005.03.003

Talmi, D., Slapkova, M., & Wieser, M. J. (2019). Testing the possibility of model-based Pavlovian control of attention to threat. *Journal of Cognitive Neuroscience, 31*, 36-48. http://doi.org/10.1162/jocn_a_01329

Taylor, S. P. (1967). Aggressive behavior and physiological arousal as a function of provocation and the tendency to inhibit aggression. *Journal of Personality, 35*, 297-310. https://doi.org/10.1111/j.1467-6494.1967.tb01430.x

Taylor, S. P., & Epstein, S. (1967). Aggression as a function of the interaction of the sex of the aggressor and the sex of the victim. *Journal of Personality, 35*, 474-486. https://doi.org/10.1111/j.1467-6494.1967.tb01441.x

Tedeschi, J. T., & Felson, R. B. (1994). *Violence, aggression, and coercive actions*. Washington, DC: American Psychological Association.

van Dongen, J. D., Drislane, L. E., Nijman, H., Soe-Agnie, S. E., & van Marle, H. J. (2017). Further evidence for reliability and validity of the triarchic psychopathy measure in a forensic sample and a community sample. *Journal of Psychopathology and Behavioral Assessment, 39*, 58-66. https://doi.org/10.1007/s10862-016-9567-5

Van Goozen, S., & Fairchild, G. (2006). Neuroendocrine and neurotransmitter correlates in children with antisocial behavior. *Hormones and Behavior, 50*, 647-654. https://doi.org/10.1016/j.yhbeh.2006.06.021

Van Goozen, S. H., Matthys, W., Cohen-Kettenis, P. T., Gispen-de Wied, C., Wiegant, V. M., & Van Engeland, H. (1998). Salivary cortisol and cardiovascular activity during stress in oppositional-defiant disorder boys and normal controls. *Biological Psychiatry, 43*, 531-539. https://doi.org/10.1016/S0006-3223(97)00253-9

Van Goozen, S. H., Van den Ban, E., Matthys, W., Cohen-Kettenis, P. T., Thijssen, J. H., & Van Engeland, H. (2000). Increased adrenal androgen functioning in children with oppositional

defiant disorder: A comparison with psychiatric and normal controls. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 1446-1451. https://doi.org/10.1097/00004583-200011000-00020

VanZomeren-Dohm, A., Xu, X., Thibodeau, E., & Cicchetti, D. (2016). Child maltreatment and vulnerability to externalizing spectru disorders. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *The Oxford handbook of externalizing spectrum disorders* (pp. 94). New York: Oxford University Press.

Veit, R., Flor, H., Erb, M., Hermann, C., Lotze, M., Grodd, W., & Birbaumer, N. (2002). Brain circuits involved in emotional learning in antisocial behavior and social phobia in humans. *Neuroscience Letters, 328*, 233-236. https://doi.org/10.1016/S0304-3940(02)00519-0

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*, 3365-3372. https://doi.org/10.1016/j.neuroimage.2009.11.040

Venables, N. C., & Patrick, C. J. (2012). Validity of the Externalizing Spectrum Inventory in a criminal offender sample: Relations with disinhibitory psychopathology, personality, and psychopathic features. *Psychological Assessment, 24*, 88-100. https://doi.org/10.1037/a0024703

Virkkunen, M. (1985). Urinary free cortisol secretion in habitually violent offenders. *Acta Psychiatrica Scandinavica, 72*, 40-44. https://doi.org/10.1111/j.1600-0447.1985.tb02568.x

Vles, J., Feron, F., Hendriksen, J., Jolles, J., van Kroonenburgh, M., & Weber, W. (2003). Methylphenidate down-regulates the dopamine receptor and transporter system in children with attention deficit hyperkinetic disorder (ADHD). *Neuropediatrics, 34*, 77-80. https://doi.org/10.1055/s-2003-39602

Vogel, S., Klumpers, F., Krugers, H. J., Fang, Z., Oplaat, K. T., Oitzl, M. S., Joëls, M., & Fernández, G. (2015). Blocking the mineralocorticoid receptor in humans prevents the stress-induced

enhancement of centromedial amygdala connectivity with the dorsal striatum. *Neuropsychopharmacology, 40*, 947-956. https://doi.org/10.1038/npp.2014.271

Wagels, L., Votinov, M., Kellermann, T., Eisert, A., Beyer, C., & Habel, U. (2018). Exogenous testosterone enhances the reactivity to social provocation in men. *Frontiers in Behavioral Neuroscience, 12*. https://doi.org/10.3389/fnbeh.2018.00037

Wagels, L., Votinov, M., Kellermann, T., Konzok, J., Jung, S., Montag, C., Schneider, F., Eisert, A., Beyer, C., & Habel, U. (2019). Exogenous testosterone and the monoamine-oxidase A polymorphism influence anger, aggression and neural responses to provocation in men. *Neuropharmacology, 156*, 107491. https://doi.org/10.1016/j.neuropharm.2019.01.006

Walton, K. E., Ormel, J., & Krueger, R. F. (2011). The dimensional nature of externalizing behaviors in adolescence: evidence from a direct comparison of categorical, dimensional, and hybrid models. *Journal of Abnormal Child Psychology, 39*, 553-561. https://doi.org/10.1007/s10802-010-9478-y

Wang, J., Korczykowski, M., Rao, H., Fan, Y., Pluta, J., Gur, R. C., McEwen, B. S., & Detre, J. A. (2007). Gender difference in neural response to psychological stress. *Social Cognitive and Affective Neuroscience, 2*, 227-239. https://doi.org/10.1093/scan/nsm018

Warburton, W. A., & Bushman, B. J. (2019). The competitive reaction time task: The development and scientific utility of a flexible laboratory aggression paradigm. *Aggressive Behavior*, 45, 389-396. https://doi.org/10.1002/ab.21829

Waschbusch, D. A., Pelham, W. E., Jennings, J. R., Greiner, A. R., Tarter, R. E., & Moss, H. B. (2002). Reactive aggression in boys with disruptive behavior disorders: Behavior, physiology, and affect. *Journal of Abnormal Child Psychology, 30*, 641-656. https://doi.org/10.1023/A:1020867831811

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063-1070. http://dx.doi.org/10.1037/0022-3514.54.6.1063

Webster, G. D., DeWall, C. N., Pond Jr, R. S., Deckman, T., Jonason, P. K., Le, B. M., Nichols, A. L., Schember, T. O., Crysel, L. C., & Crosier, B. S. (2014). The brief aggression questionnaire: psychometric and behavioral evidence for an efficient measure of trait aggression. *Aggressive Behavior, 40*, 120-139. https://doi.org/10.1002/ab.21507

Weidler, C., Habel, U., Hüpen, P., Akkoc, D., Schneider, F., Blendy, J. A., & Wagels, L. (2019). On the complexity of aggressive behavior: contextual and individual factors in the Taylor Aggression Paradigm. *Frontiers in Psychiatry, 10*. https://doi.org/10.3389/fpsyt.2019.00521

Weidler, C., Wagels, L., Regenbogen, C., Hofhansel, L., Blendy, J. A., Clemens, B., Montag, C., & Habel, U. (2018). The influence of the OPRM1 (A118G) polymorphism on behavioral and neural correlates of aggression in healthy men. *Neuropharmacology,* 156, 107467. https://doi.org/10.1016/j.neuropharm.2018.12.014

Welker, K. M., Gruber, J., & Mehta, P. H. (2015). A positive affective Neuroendocrinology (PANE) approach to Reward and Behavioral Dysregulation *Frontiers in Psychiatry, 18*, 1-13. https://doi.org/10.3389/fpsyt.2015.00093

Welker, K. M., Lozoya, E., Campbell, J. A., Neumann, C. S., & Carré, J. M. (2014). Testosterone, cortisol, and psychopathic traits in men and women. *Physiology & Behavior, 129*, 230-236. https://doi.org/10.1016/j.physbeh.2014.02.057

West, B. T., Welch, K. B., & Galecki, A. T. (2014). *Linear mixed models: a practical guide using statistical software*. Boca Raton: CRC Press Taylor & Francis Group.

White, B. A., Jarrett, M. A., & Ollendick, T. H. (2013). Self-regulation deficits explain the link between reactive aggression and internalizing and externalizing behavior problems in children. *Journal of Psychopathology and Behavioral Assessment, 35*, 1-9. https://doi.org/10.1007/s10862-012-9310-9

White, S. F., VanTieghem, M., Brislin, S. J., Sypher, I., Sinclair, S., Pine, D. S., Hwang, S., & Blair, R. J. R. (2016). Neural correlates of the propensity for retaliatory behavior in youths with

disruptive behavior disorders. *American Journal of Psychiatry, 173*, 282-290. https://doi.org/10.1176/appi.ajp.2015.15020250

Wickham, H., & Miller, E. (2019). Haven: import and export 'SPSS', 'Stata' and 'SAS' Files. https://CRAN.R-project.org/package=haven

Wilkowski, B. M., & Robinson, M. D. (2008). The cognitive basis of trait anger and reactive aggression: an integrative analysis. *Personality and Social Psychology Review, 12*, 3-21. https://journals.sagepub.com/doi/10.1177/1088868307309874

Wittchen, H.-U., Zaudig, M., & Fydrich, T. (1997). *Strukturiertes Klinisches Interview für DSM-IV Achse I und II (SKID)*. Göttingen: Hogrefe.

Wolf, O. T. (2009). Stress and memory in humans: twelve years of progress? *Brain Research, 1293*, 142-154. https://doi.org/10.1016/j.brainres.2009.04.013

Wolfram, M., Bellingrath, S., & Kudielka, B. M. (2011). The cortisol awakening response (CAR) across the woman menstrual cycle. *Psychoneuroendocrinology, 36*, 905-912. https://doi.org/10.1016/j.psyneuen.2010.12.006

Word Health Organization (2007). *Third milestones of a global campaign for violence prevention report, 2007: scaling up.* Geneva, Switzerland: World Health Organization.

Yehuda, R. (2002). Post-traumatic stress disorder. *New England Journal of Medicine, 346*, 108-114. https://doi.org/10.1056/NEJMra012941

Yildirim, B. O., & Derksen, J. J. (2012). A review on the relationship between testosterone and the interpersonal/affective facet of psychopathy. *Psychiatry Research, 197*, 181-198. https://doi.org/10.1016/j.psychres.2011.08.016

Young, S. E., Stallings, M. C., Corley, R. P., Krauter, K. S., & Hewitt, J. K. (2000). Genetic and environmental influences on behavioral disinhibition. *American Journal of Medical Genetics, 96*, 684-695. https://doi.org/10.1002/1096-8628(20001009)96:5<684::AID-AJMG16>3.0.CO;2-G

Zänkert, S., Bellingrath, S., Wüst, S., & Kudielka, B. M. (2019). HPA axis responses to psychological challenge linking stress and disease: what do we know on sources of intra-and

interindividual variability? *Psychoneuroendocrinology, 105*, 86-97. https://doi.org/10.1016/j.psyneuen.2018.10.027

Zänkert, S., Kudielka, B., & Wüst, S. (2020). Effect of sugar administration on cortisol responses to acute psychosocial stress. *Psychoneuroendocrinology*, 115, 104607. https://doi.org/10.1016/j.psyneuen.2020.104607

Zepf, F. D., Stadler, C., Demisch, L., Schmitt, M., Landgraf, M., & Poustka, F. (2008). Serotonergic functioning and trait-impulsivity in attention-deficit/hyperactivity-disordered boys (ADHD): influence of rapid tryptophan depletion. *Human Psychopharmacology: Clinical and Experimental, 23*, 43-51. https://doi.org/10.1002/hup.896

Zisner, A. R., & Beauchaine, T. P. (2016). Midbrain neural mechanisms of trait impulsivity. In T. P. Beauchaine & S. P. Hinshaw (Eds.), *The Oxford handbook of externalizing spectrum disorders* (Vol. Oxford University Press). New York.

Appendix

Appendix A

The four items of the deception check questionnaire (competitiveness scale) using 4-point scales anchored at 1 (*not true at all*) and 4 (*completely true*)

1. In general, it was important for me to provide a good performance.
2. It was important for me to achieve a high profit.
3. It was important for me to outpace my opponent.
4. It was important for me to achieve a higher profit than my opponent.

Appendix B

Provocation sequence in the fixed order condition

Trial	Low provocation	Medium provocation	High provocation	Win
1		- 50 cents		
2				+ 50 cents
3	- 10 cents			
4	0 cents			
5	- 20 cents			
6				+ 50 cents
7				+ 50 cents
8				+ 50 cents
9		- 60 cents		
10		- 60 cents		
11		- 40 cents		
12			- 80 cents	
13			- 70 cents	
14			- 90 cents	
15				+ 50 cents
16	- 10 cents			
17	0 cents			
18	- 20 cents			
19				+ 50 cents
20			- 90 cents	
21			- 70 cents	
22			- 80 cents	
23				+ 50 cents
24		- 50 cents		
25		- 30 cents		
26				+ 50 cents
27				+ 50 cents
56		- 40 cents		
29				+ 50 cents
30	- 20 cents			
31				+ 50 cents
32		- 60 cents		
33				+ 50 cents
34	0 cents			
35				+ 50 cents
36				+ 50 cents
37				+ 50 cents
38				+ 50 cents
39				+ 50 cents
40			- 90 cents	
41			- 80 cents	
42			- 90 cents	
43				+ 50 cents
44				+ 50 cents
45				+ 50 cents
46		- 30 cents		
47		- 30 cents		
48		- 50 cents		
49				+ 50 cents
50				+ 50 cents
51				+ 50 cents
52	- 10 cents			
53	0 cents			
54	- 10 cents			
55			- 70 cents	
56			- 70 cents	
57			- 90 cents	
58				+ 50 cents
59				+ 50 cents
60				+ 50 cents
61			- 90 cents	
62				+ 50 cents
63			- 80 cents	
64				+ 50 cents
65				+ 50 cents
66	0 cents			

67	- 20 cents			
68	- 20 cents			
69		- 30 cents		
70		- 40 cents		
71		- 40 cents		
72			- 70 cents	
73			- 70 cents	
74			- 70 cents	
75				+ 50 cents
76				+ 50 cents
77			- 80 cents	
78			- 80 cents	
79			- 90 cents	
80		- 60 cents		
81		- 40 cents		
82		- 30 cents		
83	0 cents			
84	- 20 cents			
85	- 20 cents			
86				+ 50 cents
87				+ 50 cents
88				+ 50 cents
89	0 cents			
90	- 10 cents			
91	- 10 cents			
92				+ 50 cents
93				+ 50 cents
94		- 50 cents		
95				+ 50 cents
96		- 60 cents		
97				+ 50 cents
98		- 50 cents		
99				+ 50 cents
100				+ 50 cents

Appendix C

Activation peaks within a significant cluster contrasting stress vs. control during Scan*STRESS* (two-tailed combined test, FWE corrected at .05)

Contrast	Cluster size (k)	Cytoarchitectonic location	*Z*	x	y	z
Stress > Control	30496	N/A	5.08	0	-28	-2
		L Thalamus	4.93	-8	-10	12
		L Middle occipital gyrus	4.91	-30	-70	30
		R Cerebellum (VI)	4.90	36	-36	-36
		L Thalamus	4.84	-4	-10	-2
		N/A	4.82	-26	-34	-38
		N/A	4.81	4	-40	-36
		Cerebellar vermis (9)	4.71	4	-52	-30
		N/A	4.64	-6	-18	-10
		R Thalamus	4.62	16	-12	6
		L Cerebellum (VI)	4.61	-32	-38	-36
	13537	L Posterior-medial frontal gyrus	5.02	-8	0	66
		R Precentral gyrus	4.87	40	-4	40
		L IFG (p. Triangularis)	4.76	-42	20	14
		L Insula lobe	4.70	-26	20	-6
		L Insula lobe	4.65	-30	18	-2
		L Superior frontal gyrus	4.60	-18	6	70
		L Precentral gyrus	4.58	-42	-6	36
		L Superior frontal gyrus	4.56	-18	10	68
		L Middle frontal gyrus	4.47	-38	50	2
		R Middle frontal gyrus	4.43	38	58	6
		R Insula lobe	4.39	32	22	-16

Notes: L = left, R = right. IFG = Inferior frontal gyrus, N/A = Not found in any probability map.

Appendix D

Activation peaks within a significant cluster contrasting control vs. stress (two-tailed combined test, FWE corrected at .05)

Contrast	Cluster size (k)	Cytoarchitectonic location	*Z*	x	y	z
Control > Stress	4105	N/A	4.47	-2	2	-12
		N/A	4.42	8	4	-12
		N/A	4.39	2	0	-20
		R Middle orbital gyrus	4.35	8	60	-12
		L Middle orbital gyrus	4.31	0	52	-8
		L Rectal gyrus	4.08	-4	36	-22
		L Superior orbital gyrus	3.99	-14	24	-22
		R Superior orbital gyrus	3.96	20	30	-26
		R Rectal gyrus	3.95	10	34	-26
		L Superior frontal gyrus	3.89	-12	46	34
		L Superior frontal gyrus	3.84	-12	46	40
	1753	L PCC	4.17	-10	-52	30
		L Precuneus	4.12	-12	-56	28
		L Cuneus	4.12	-8	-62	24
		R Precuneus	3.93	8	-58	20
	1483	R Insula lobe	4.62	40	-12	0
		R Rolandic operculum	4.60	42	-10	16
		R Heschls gyrus	4.33	50	-8	6
		R Superior temporal gyrus	3.99	60	-8	2
		R Heschls gyrus	3.37	38	-32	16
		R Rolandic operculum	3.29	62	-14	12
		R Superior temporal gyrus	3.27	66	-18	14
		R Superior temporal gyrus	3.12	48	-30	12
	1001	L Superior temporal gyrus	4.29	-58	-4	0
		L Rolandic operculum	4.17	-44	-16	16
		L Middle temporal gyrus	3.95	-60	-12	-18
		L Insula lobe	3.80	-40	-12	6
		L Superior temporal gyrus	3.69	-64	-24	6
		L Superior temporal gyrus	3.58	-56	-14	4
		L Superior temporal gyrus	3.30	-66	-12	2
		N/A	3.20	-68	-14	4

Notes: L = left, R = right. PCC = Posterior cingulate cortex, N/A = Not found in any probability map.

Appendix E

Activation peaks within a significant cluster (stress vs. control) covarying with cortisol increase contrasting subjects exhibiting low vs. high externalization (FWE $p < .05$)

Contrast	Cluster size (k)	Cytoarchitectonic location	*Z*	x	y	z
Low > High	2960	N/A	4.37	14	2	28
		L Nucleus caudatus	3.51	-12	8	18
		L Putamen	3.45	-22	4	8
		N/A	3.27	4	-12	22
		R MCC	3.26	8	-20	52
		L Olfactory cortex	3.22	0	4	-2
		N/A	3.16	-30	0	-4
		R Posterior medial Frontal	3.11	12	-8	52
		N/A	3.09	18	-10	40
		L Putamen	3.09	-26	-8	16
		N/A	3.02	22	-10	38
	2054	N/A	3.76	26	-48	4
		R Fusiform Gyrus	3.69	34	-58	0
		R Fusiform Gyrus	3.68	42	-34	-12
		N/A	3.59	36	-54	8
		R Calcarine Gyrus	3.54	22	-72	14
		N/A	3.53	34	-56	12
		R Lingual Gyrus	3.52	30	-52	0
		R Fusiform Gyrus	3.21	32	-46	-4
		N/A	3.03	22	-48	34
		R Fusiform Gyrus	2.74	26	-82	-2
		R Lingual Gyrus	2.69	24	-80	2
	1789	R Cerebellum (Crus 1)	3.15	26	-74	-32
		R Cerebellum (IX)	3.13	14	-56	-34
		L Cerebellum (III)	3.10	-12	-42	-18
		L Cerebellum (IX)	3.05	-8	-52	-34
		L Parahippocampal Gyrus	2.98	-18	-28	-10
		L Cerebellum (IV-V)	2.97	-10	-56	-12
		R Lingual Gyrus	2.9	20	-96	-8
		N/A	2.87	-8	-40	-26
		R Inferior Occipital Gyrus	2.78	24	-102	-6
		N/A	2.75	-10	-32	-24
		Cerebellar Vermis (4/5)	2.74	4	-56	-14

Notes: L = left, R = right, MCC = Middle cingulate cortex, N/A = Not found in any probability map.

Appendix F

Activation peaks within a significant cluster covarying positively with selected aggression levels (parametric modulation)

Cluster size (k)	Cytoarchitectonic location	*T*	x	y	z
217	N/A	4.43	-27	-22	53
	L Precentral Gyrus	4.35	-30	-28	62
	L Precentral Gyrus	4.21	-39	-25	65
	L Precentral Gyrus	4.20	-33	-19	68
201	L Superior Temporal Gyrus	4.34	-51	-28	17
	N/A	4.25	-33	-37	26
	N/A	4.22	-39	-40	23
	L Rolandic Operculum	3.97	-45	-28	23
	N/A	3.86	-36	-13	29
	N/A	3.73	-42	-10	29
110	R Nucleus Caudatus	5.26	18	17	17
	N/A	4.96	-3	-4	20
	N/A	4.61	-9	-10	29
	N/A	4.35	-12	8	26
	N/A	3.81	3	14	8
	N/A	3.77	6	20	11
	N/A	3.7	12	2	26
	N/A	3.63	6	2	20

Notes: L = left, R = right. N/A = Not found in any probability map.

Eidesstattliche Versicherung

Familienname: Konzok

Vorname: Julian

Geb. 07.07.1987

Ich erkläre hiermit an Eides Statt, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Textpassagen, Daten, Bilder oder Grafiken sind unter Angabe der Quelle gekennzeichnet.

Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise entgeltlich / unentgeltlich geholfen; dies ist auch in der Dissertation an den entsprechenden Stellen explizit ausgewiesen.

1. siehe Kapitel 1.5
2. siehe Kapitel 1.5
3. siehe Kapitel 1.5

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- beziehungsweise Beratungsdiensten (Promotionsberater oder anderer Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt. Ich versichere an Eides Statt, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Vor Aufnahme der obigen Versicherung an Eides Statt wurde ich über die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung belehrt.

_____ _____

(Ort, Datum) Julian Konzok