Understanding the Interplay between Empathy and Aggression across the Psychopathy Spectrum

Célia F. Camara

Doctor of Philosophy (PhD)

Department of Psychology

University of Essex

Publication date: August 2024

Review date: May 2025

Declaration Statement

I hereby declare that this thesis is my own original work. To the best of my knowledge, it does not contain any material previously published or written by another person, nor does it include substantial portions of material that have been accepted for the award of any other degree or diploma at the University of Essex or any other educational institution, except where due acknowledgment is made in the thesis. Contributions to the research made by others, with whom I have worked at the University of Essex or elsewhere, are explicitly acknowledged. Additionally, I declare that the intellectual content of this thesis is the result of my own efforts, except where assistance in project design, conception, style, presentation, and linguistic expression has been acknowledged.

Signed: ______________________

Acknowledgements

I am deeply grateful to everyone who has supported me throughout my PhD journey. First and foremost, I want to extend my heartfelt thanks to my supervisors, Dr. Alejandra Sel, Dr. Paul Hanel, and Dr. Carina de Clerk. Your ongoing support and flexibility have been crucial in navigating the various obstacles that arose during this project. Your expertise, patience, and encouragement have made an enormous difference, and I am thankful for your mentorship. I also want to express my sincere appreciation to my collaborators, Dr. Andrés Molero Chamizo, Dr. Carmen Sergiou, Prof. Michael Nitsche, Dr. Nathzidy Rivera Urbina, and Dr. Sebastian Halder. Thank you for willingly diving into my projects and bringing your invaluable insights. Your contributions have significantly enriched my research and pushed it forward in ways I could not have achieved alone.

I extend my sincere thanks to Arkadij Lobov and Hayley Warman, who have stood by me through the various challenges of my research endeavours. To the staff members in my department, especially Elena and Steve. Your help and dedication made the technical challenges that arose during data collection much more manageable, and I am truly grateful for your assistance. I am also grateful to the students I've had the privilege to supervise and work with. Your contributions in recruiting participants and collecting data have been invaluable to the success of my projects. A special thanks to all the participants who enthusiastically volunteered in my EEG experiment, offering their time and sacrificing their hairdo in exchange for engaging conversations and electrode gel-loaded scalp 'massages'.

Lastly, I want to express my heartfelt gratitude to my close friends and family members who have been my pillars throughout this journey. To my fellow PhD colleagues, Dr. Khetam Alfaraj, Dr. Olayinka Farris, Dr. Charli Sherman and soon-to-be Dr. Aisyah Shamshun Baharin, I am incredibly fortunate to have shared this journey with you. From day one, you have made this experience less solitary and far more enjoyable. I am proud to count you not only as fellow professionals but also as friends for life. Your camaraderie and support have meant the world to me, and I look forward to our future collaborations. To Sonia, Jose, Jema, Erick, Pau, Felix, Shehzadi, Sam, and Ravi, your patience in listening to my endless rants about the trials and triumphs of my PhD studies means more to me than words can convey. To my mom, sisters, *mémé*, and aunties. Thank you for your eagerness to hear about my work, even if you didn't understand a word. I am blessed to be surrounded by such loving and supportive women who consistently remind me how proud they are of me and make me feel unstoppable. Above all, I want to dedicate this work to my *abuelo*. Although he is no longer with us, his presence and guidance were my main source of inspiration to become a researcher. I will forever be thankful for his influence in my life.

To each and every one of you who has supported and encouraged me along this challenging but rewarding journey, I extend my deepest gratitude. Your belief in me and your unwavering support have made this thesis possible.

Content

List of abbreviations

ACC	Anterior Cingulate Cortex
ACME	Affective and Cognitive Measure of Empathy
ADD	Attention Deficit Disorder
ADHD	Attention Deficit and Hyperactivity Disorder
AIC	Akaike Information Criterion
AI	Anterior Insula
$BCCI_{95\%}$	95% Bootstrapped Confidence Intervals
BIC	Bayesian Information Criterion
BPAQ	Buss and Perry Aggression Questionnaire
Cal	Callousness
ccPAS	Cortico-cortical paired associative stimulation
$CI_{95\%}$	95% Confidence Interval
COG	Cognitive empathy subscale in ACME
CSP	Common Special Pattern
cTBS	Continuous Theta Burst Stimulation
Df	Degrees of Freedom
DIS	Affective dissonance
DLPFC	Dorsolateral Prefrontal Cortex
DSM–5	Diagnostic and Statistical Manual of Mental Disorders–5th edition
EC	Empathic Concern
EEG	Electroencephalography
ERDs	Event-related Dynamics
ERPs	Event-related Potentials

fMRI	Functional Magnetic Resonance Imaging
F1	Factor 1 in the SRP–SF, representing interpersonal callousness
F2	Factor 2 in the SRP–SF, representing social deviance
FFT	Fast Fourier Transform
HD-tACS	High Definition tACS
HD-tDCS	High Definition tDCS
HF-rTMS	High frequency rTMS
ICA	Independent Component Analysis
ICD–11	International Classification of Diseases–11th edition
ICU	Inventory of Callous-Unemotional Traits
IFG	Inferior Frontal Gyrus
IPL	Inferior Parietal Lobule
IRI	Interpersonal Reactivity Index
iTBS	Intermittent Theta Burst Stimulation
LDA	Linear Discriminant Analysis
LF-rTMS	Low frequency rTMS
LPP	Late Positive Potentials
LQ	Laterality Quotient
LTP	Long-term Potentiation
MDS	Moral Disengagement Scale
MFG	Middle Frontal Gyrus
NIBS	Non-invasive Brain Stimulation
OFC	Orbitofrontal Cortex
PAG	Periaqueductal Gray
PCA	Principal Component Analysis

PCL–R	Psychopathy Checklist–Revised
PEESE	Precision-Effect Estimate with Standard Error
PET	Precision-Effect Test
PICO	Population-Intervention-Comparison-Outcome framework
Phys	Physical Aggression
PMv	Ventral Premotor Cortex
PFC	Prefrontal Cortex
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
Proact	Proactive Aggression
PT	Perspective-taking
React	Reactive Aggression
rMT	Resting Motor Threshold
rTMS	Repetitive TMS
ROI	Regions of Interest
RPQ	Reactive and Proactive Aggression Questionnaire
SABIC	Sample-Size Adjusted BIC
SEM	Structural Equation Modelling
SMA	Supplementary Motor Area
SRP–4	Self-Report Psychopathy Scale–4th Edition
SRP–SF	Self-Report Psychopathy Scale–Short Form
tACS	Transcranial Alternating Current Stimulation
TBS	Theta Burst Stimulation
tDCS	Transcranial Direct Current Stimulation
TMS	Transcranial Magnetic Stimulation
TPJ	Temporo-parietal Junction

Unc	Uncaring behaviour
Une	Unemotional traits
VAS	Visual Analogue Scale
VIF	Variance Inflation Factor
VMPFC	Ventromedial Prefrontal Cortex
VLPFC	Ventrolateral Prefrontal Cortex

List of figures

List of tables

Preface

"Human nature is evil" – this provocative assertion by the Chinese philosopher Xunzi from the third century B.C. finds some support in contemporary research. Modern studies show that a person's actions and decisions can indeed be influenced by their biological predispositions. And while these insights do not provide the complete picture, the existence of ruthless and callous offenders such as those labelled *psychopaths* does suggest that some individuals may be inherently prone to cruelty[1]. My introduction to this cohort began with Ted Bundy – one of the most prolific serial killers in history. Born in November 24 (1946) in Burlington, Vermont, Theodore Robert Bundy grew up in a middle-class family and excelled academically. He was known for his charming demeanour, making him well-liked by those around him. However, beneath this façade laid a profound absence of empathy and remorse – traits that would later lead him to commit a series of brutal murders and sexual assaults to numerous young women and girls. His ability to deceive and manipulate – coupled with the brutality of his crimes – left many (including myself) questioning how such cruelty could exist in a seemingly ordinary person. This curiosity led to questions that became the foundation of my research endeavours: What can lead a person to deliberately harm others? And crucially, what can be done to prevent it?

While it is tempting to simply dismiss Bundy's actions as inexplicable anomalies, cruel forms of interpersonal harm are not merely the product of *evil* but are influenced by a myriad of interacting factors – as discussed by Simon Baron-Cohen in *The Science of Evil* (2011). Upbringing, societal influences, and psychological predispositions can converge in ways that push ordinary individuals towards actions they might otherwise find inconceivable. Interestingly, Bundy's story reveals a paradox in societal responses to extreme cruelty. While his actions were undeniably horrific, the public's reaction to his trial and execution mirrored a similar callousness. What I found particularly fascinating – and troubling – was the eagerness with which people anticipated his execution and the indifference some displayed towards his suffering. This paradox highlights the complexity of human aggression; it suggests that the potential for cruelty resides within the spectrum of human behaviour rather than being confined to a select few with clinical disorders like psychopathy. In fact, traits of psychopathy are not solely relegated to notorious figures like Ted Bundy but can manifest in subtler forms among the general population. This is not to suggest that everyone that possesses these traits is inherently cruel and prone to aggression, but rather to emphasise that the boundary between cruelty and moral behaviour is permeable and subject to situational and psychological influences. Understanding these dynamics is thus crucial for addressing and mitigating instances of cruelty in society.

These insights shaped the core of my thesis, guiding my investigation into how traits representing the construct of psychopathy may relate to a risk for interpersonal harm in non-clinical and non-criminal samples. This investigation involves a multidisciplinary approach, mainly integrating insights from cognitive neuroscience and social psychology. This approach aims to provide a nuanced perspective on how the potential for cruelty may manifest in seemingly ordinary people, offering potential implications for both theoretical frameworks and prospective interventions.

Chapter I.

Introduction to the Project

General abstract

Psychopathy, traditionally viewed as a categorical construct, is now increasingly recognised as a spectrum encompassing varying degrees of affective and interpersonal deficits that may contribute to antisocial behaviours like aggression. Building on this perspective, the present thesis investigates how subclinical expressions of psychopathic traits – in particular callous-unemotional traits – relate to empathic responsiveness and aggression proneness in non-clinical, community-based samples. Findings reveal that even at subclinical levels, callous-unemotional traits are linked to reduced affective empathy and more calculated, premeditated aggression, likely driven by disruptions in affective processing. These associations were similarly observed at the neural level, with participants reporting more callous-unemotional traits also exhibiting lower electrophysiological responses to the perception of others' pain. Notably, the study also identifies preliminary evidence that affective states linked to the expression of callous-unemotional traits could be modulated via non-invasive brain stimulation, offering a potential avenue for intervention. Overall, this research supports a dimensional understanding of psychopathy, suggesting that even mild or subclinical manifestations of psychopathic traits can be associated with deficits in the empathic response, and that such deficits may be detectable and modifiable at the neural level. These findings underscore the importance of adopting a multidisciplinary approach in understanding – and potentially mitigating – the risks of psychopathy.

"Don't ask why such things happen. It's just the nature of evil."

– Baron-Cohen in 'The Science of Evil' (2011, p.151)

The problem of psychopathy

A large body of research has shown that psychopathy is not only one of the strongest dispositional predictors of aggression (e.g., Dolan & Doyle, 2000; Heym et al., 2019; Leistico et al., 2008; Neumann & Hare, 2008), but it is particularly associated with aggressive behaviours that are more violent and sadistic[2] (e.g., Holt et al., 1999; Meloy, 1997, 2006; Porter et al., 2003). Although less than 1% of the general population is estimated to be diagnosable as psychopaths, people who meet the criteria for psychopathy are disproportionately represented in prison populations – especially among offenders who commit violent and victim-based crimes (Carré et al., 2013; Kiehl & Hoffman, 2011). Criminal psychopaths not only pose a great societal burden but also signify great financial costs, which calls for the need of intervention efforts aiming to tackle the risks associated with this notorious minority (Kiehl & Buckholtz, 2010; Kiehl & Hoffman, 2011). But how exactly do we identify psychopaths?

Defining psychopathy

The presence of psychopaths in society has been recognised for centuries, with historical references tracing back to the early nineteenth century (e.g., Koch, 1892; Pinel, 1801; Prichard, 1835). However, the characterisation of symptoms that describe these individuals only came about less than a century ago, spearheaded by the pioneering work of Hervey Cleckley. In his book *The Mask of Sanity* (Cleckley, 1941), Cleckley provided one of the first comprehensive descriptions of psychopathy as a disorder, delineating sixteen core personality traits (Table 1.1). He famously coined the term *mask of sanity* to describe how psychopaths can appear normal and even engaging on the surface, while concealing a profound moral and emotional deficit. This depiction echoed earlier conceptualisations of *moral insanity* – described as "a form of moral derangement in which the intellectual faculties are unimpaired, but the moral principles of the mind are depraved or perverted" (Prichard, 1835, p.15). Building on this foundation, Robert Hare further refined the conceptualisation of psychopathy with the development of the Psychopathy Checklist.

Table 1.1. | *Cleckley's traits of psychopathy*

1. Superficial harm	**10.** Lack of insight
2. High intelligence	**11.** Lack of reliability
3. Absence of irrational thinking	**12.** Grotesque behaviour under the influence of alcohol
4. Lack of nervousness/psychoneurotic manifestations	**13.** Impersonal and poorly integrate sex life
5. Low probability of suicide	**14.** Failure to follow any clear life plan
6. Absence of remorse or shame	**15.** Poor judgement and failure to learn from experience
7. Falsehood and insincerity	**16.** Improperly motivated antisocial behaviour
8. Pathological egocentrism	
9. Inability to love and express emotions	

First introduced in 1980 (Hare, 1980), and revised in subsequent editions (Hare, 2003; Hare et al., 1991), the Psychopathy Checklist – Revised (PCL–R) provides a more nuanced and empirically grounded definition of psychopathy as a clinical disorder, highlighting its various facets and offering specific indications for its diagnosis (see Box 1.1). The PCL–R specifically evaluates the prevalence and recurrence of socio-affective deficits and antisocial behaviours that collectively comprise the construct of psychopathy. These features can be broadly characterised into three distinct categories, including: a) affective deficits such as callousness, unemotionality, and remorselessness; b) a grandiose-manipulative interpersonal style, comprising traits such as dishonest charm, lying, and grandiosity; and c) socially deviant behaviours mainly characterised by impulsivity, irresponsibility, and thrill seeking (Hare, 2003). This conceptualisation allows to identify distinct psychopathic subtypes, broadening our understanding of the disorder beyond its antisocial manifestations.

Box 1.1. | *The PCL-R as a diagnostic tool for psychopathy*

To date, the assessment of psychopathy in both clinical and forensic settings heavily rely on the PCL–R. It is considered that an individual meets the criteria for psychopathy when they reach a score of over 25 out of 40 (although the cutoff score is set at 30 in the United States; Hare, 2003). These scores are derived through a meticulous and structured process involving a semi-structured interview conducted by a trained clinician – supplemented by a comprehensive review of collateral information, such as mental health evaluations, criminal records, behavioural observations, and developmental history. This approach ensures that the assessment captures both self-presentation and corroborated behavioural patterns over time.

Although the PCL–R is highly regarded for its empirical validity and clinical relevance, it is important to emphasise that psychopathy is not formally recognised as a psychiatric diagnosis in the primary diagnostic manuals currently in use – i.e., the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) and the International Classification of Diseases (ICD-11). This lack of formal recognition reflects ongoing debates within the psychiatric and psychological communities about how best to define and classify psychopathy. In fact, psychopathy is most often discussed in relation to other diagnosable conditions – particularly antisocial personality disorder – rather than as an isolated disorder. As such, the PCL–R is often used to guide decision-making in forensic and clinical contexts even in the absence of formal diagnostic status. For instance, clinicians may use PCL–R-informed profiles to tailor intervention strategies, placing greater emphasis on behavioural management in cases of high affective detachment. On the other hand, this tool may be used for evaluations of criminal responsibility, parole decisions, treatment amenability, and risk of reoffending as research shows that PCL–R scores can predict future crime and violence (Sohn et al., 2020), even among non-incarcerated samples (e.g., Colins et al., 2017; Hecht et al., 2016).

For instance, one commonly referenced differentiation of psychopathy subtypes has been through the exploration of its *primary* and *secondary* variants (e.g., Del Gaizo & Falkenbach, 2008; Jakobwitz & Egan, 2006; Levenson et al., 1995; Skeem et al., 2003; Vaughn et al., 2009; Yildirim & Derksen, 2015). Individuals described as *primary psychopaths* typically exhibit low anxiety, shallow affect, and minimal capacity for guilt or empathy, coupled with a socially manipulative interpersonal style. In contrast, *secondary psychopaths* tend to display heightened emotional reactivity, increased impulsivity, and a greater propensity for anxiety (Krstic et al., 2018; Yildirim & Derksen, 2015). These distinctions have fuelled debates regarding aetiology and prognosis evaluations in psychopathy. The characterisation of primary psychopathy, in particular, describes individuals who have the capacity to be emotionally detached without engaging in overt aggression – attributes that can be functionally adaptive in certain contexts (e.g., high-status roles in corporate, political, or legal domains) (Hall & Benning, 2006). Such cases suggest that psychopathic traits may be expressed along a continuum of social adaptation.

Psychopathy as a spectrum: 'the good and the bad psychopath'

The portrayal of psychopathic individuals as devoid of empathy, inherently cruel, and predisposed to criminal behaviour, has long shaped conceptualisations of psychopathy (Cleckley, 1941; Hare & Hart, 1993; Hare & Neumann, 2008). Nonetheless, while the construct of psychopathy originated within forensic and clinical contexts, psychopathic traits are not exclusive to violent offenders or psychiatric inpatients but instead exist along a continuum within the general population. In fact, studies have demonstrated that psychopathy does not equate to aggression (Corrado et al., 2004; Lilienfeld et al., 2005), nor is it synonymous with criminal conduct (Campos et al., 2022). This has led to an ongoing debate regarding the role of antisocial behaviour (e.g., criminality, aggression) within the broader framework of psychopathy. Some even argue that affective and interpersonal deficits are more central to the construct of psychopathy (e.g., Frick et al., 2003, 2014; Muñoz & Frick, 2012; Seibert et al., 2011), while antisocial behaviour should be viewed as probable and context-dependent outcomes of these core personality traits (Skeem & Cooke, 2010). As a result, there is a growing consensus that psychopathy is best conceptualised as existing on a spectrum, rather than as a categorical disorder (Edens et al., 2006; Patrick et al., 2009; Skeem et al., 2003).

The research presented in this thesis adopts this approach, using the term *psychopathy* to refer to a constellation of personality characteristics – primarily affective and interpersonal –, which can augur different behavioural outcomes. This more nuanced view allows for the identification of subclinical or functional variants of psychopathy that may reflect differences in prognosis. From this perspective, it is argued that affective deficits such as shallow affect, lack of empathy, and emotional detachment can still manifest in the absence of overt antisocial and/or aggressive behaviour. The notion of *successful psychopathy* emerges within this framework, referring to individuals who exhibit core features of psychopathy yet manage to function effectively within society, and even excel in competitive, high-stakes roles (Babiak & O'Toole, 2012; Benning et al., 2018; Hall & Benning, 2006). In such contexts, traits like emotional detachment and stress tolerance might function as protective factors, enabling individuals to navigate adversity, make rapid decisions under pressure, and be more resistant to internalising disorders like anxiety or depression (e.g., Barr & Quinsey, 2004; Douglas et al., 2006, 2008; Ricarte et al., 2022; Verona et al., 2001). This may be particularly advantageous in careers where high emotional resilience and calculated risk-taking are essential – such as corporate leadership, or law enforcement (Babiak et al., 2010; Babiak & O'Toole, 2012). Importantly, this distinction is not solely a function of personality traits, but can also be consequence of social opportunities, intelligence, socioeconomic status, and environmental circumstances.

For instance, *unsuccessful psychopaths* are more likely to originate from disadvantaged backgrounds, with limited access to education and fewer prosocial pathways for achieving goals (Benning et al., 2018). In contrast, *successful psychopaths* often benefit from more favourable environmental conditions – such as higher socioeconomic status and better educational opportunities –, and often present greater intelligence and cognitive functioning, which may enable them to better regulate impulsivity and strategically channel their affective deficits into more socially adaptive behaviours (Benning et al., 2018; Hall & Benning, 2006;

Mahmut et al., 2008). Nevertheless, even among those classified as *successful*, psychopathic traits may still predict morally dubious behaviours – such as manipulation, deception, or emotional exploitation – that are simply less visible or less likely to result in legal sanctions. This distinction is at the heart of debates contrasting antisocial behaviour in successful *vs* unsuccessful psychopathy. While unsuccessful psychopaths are more frequently involved in overt, violent, and victim-based crimes – such as physical assault, predatory violence, or repeated lawbreaking –, successful psychopaths may commit subtler, white-collar offenses – such as fraud, insider trading, or unethical business practices (see Mahmut et al., 2008 for a discussion). Therefore, while the notion of successful psychopathy offers important insights into the adaptive potential of certain psychopathic traits, it remains essential to recognise that such traits often lay the foundation for exploitative and harmful behaviour. Accordingly, affective deficits associated with psychopathy must be considered as latent risk factors for interpersonal harm – even when expressed without overt externalising behaviours.

Traits of psychopathic violence

Aggression can be described as "any behaviour that is intended to harm another person who does not want to be harmed" (Bushman & Huesmann, 2013, p.833). Depending on its function, aggression can either be *cold-blooded* (proactive), or *hot-blooded* (reactive). While both types of aggression involve some degree of intent to harm, the main distinction between the underlying behaviours is the nature of their intent and the level of premeditation involved (Raine et al., 2006). Reactive aggression is driven by immediate emotional responses, often provoked by external stimuli, whereas proactive aggression is typically unprovoked, emotionless and strategic – serving as a means to an end (Bushman & Anderson, 2001; Raine et al., 2006; Reidy et al., 2011). Generally, individuals displaying high levels of proactive aggression exhibit corresponding levels of reactive aggression, but not the other way around (Euler et al., 2017). This pattern has led some researchers to posit that proactive aggression could be a marker of aggression severity (Brugman et al., 2017). Notably, psychopathy represents the only condition that presents a selective risk for proactive aggression (Camp et al., 2013; Hecht et al., 2016; Levenson et al., 1995; Meloy, 2006). This aligns with the concept of *predatory violence* – a form of aggression that is cold, unemotional, and instrumentally motivated, often employed as a strategic means to achieve dominance, control, or material gain. This form of aggression reflects a capacity to deliberately harm others without concern or remorse, which can be linked to the expression of callous-unemotional traits.

Callous-unemotional traits denote a cluster of personality characteristics marked by a lack of empathy, remorse, and guilt, alongside shallow affect and a diminished sensitivity to others' emotional experiences (Frick et al., 2003). These traits capture affective deficits thought to underlie more severe and persistent patterns of antisocial behaviour (Muñoz & Frick, 2012). Compelling evidence shows that callous-unemotional traits define a subset of children with conduct disorder[3], who exhibit more instrumental, proactive aggression, greater sensation-seeking, and blunted emotional reactivity compared to their peers (Frick et al., 2003; Jones et al., 2009; Salekin et al., 2012). These children often show a more persistent and severe trajectory of antisocial behaviour across development, including increased risk for violence and criminal behaviour in adolescence and adulthood (Fragkaki et al., 2016; Frick et al., 2014; Hawes & Dadds, 2007). Indeed, the clinical relevance of these traits is such that the DSM-5 has incorporated a *Limited Prosocial Emotions* specifier for conduct disorder, reflecting deficits characteristic of individuals with pronounced callous-unemotional traits (American Psychiatric Association, 2013).

Moreover, callous-unemotional traits also manifest at subclinical levels within the general population, where even moderate expressions of these traits are associated with a higher risk of aggression. For example, a study by Essau *et al* (2006) reported that callous-unemotional traits were significantly linked to conduct problems, psychosocial impairment, and elevated sensation seeking in a community sample of adolescent boys and girls. Similarly, a study by Moran *et al* (2009) demonstrated that these traits remained consistently associated with conduct issues, hyperactivity, and emotional difficulties over a three-year period in a large sample of adolescents. These findings are paralleled in adult populations, with research showing that young men with higher levels of callous-unemotional traits engage in more antisocial behaviours, reporting higher numbers of arrests, charges, and incidents violence (Byrd et al., 2013; Kahn et al., 2013). This evidence suggests that callous-unemotional traits signal a predisposition toward maladaptive behaviours even outside clinical contexts. As such, investigating the expression of these traits in normative populations may not only give insights into the dimensional nature of psychopathy but also aid in identifying potential risk and protective factors that influence prognosis in at-risk individuals.

A further consequence of callous-unemotional traits is their resistance to rehabilitation. Offenders with elevated levels of these traits – such as criminal psychopaths – are typically less responsive to conventional crime deterrence strategies like punishment, often resulting in higher rates of recidivism (Hare & Neumann, 2008; Harris & Rice, 2006; Kiehl & Hoffman, 2011; Weaver et al., 2022). This resistance is attributed to their inherent lack of emotional connectedness, which hinders the internalisation of social norms and reduce the effectiveness of strategies that rely on guilt, remorse, or fear of consequences (Felthous, 2011; Frick et al., 2014; Polaschek, 2014). Ultimately, the lack of emotional connectedness in these cohorts reflects a fundamental deficit in empathy, which is hypothesised to be one of the leading reasons why psychopaths – and individuals with pronounced callous-unemotional traits – are more capable of cruelty (Baron-Cohen, 2011).

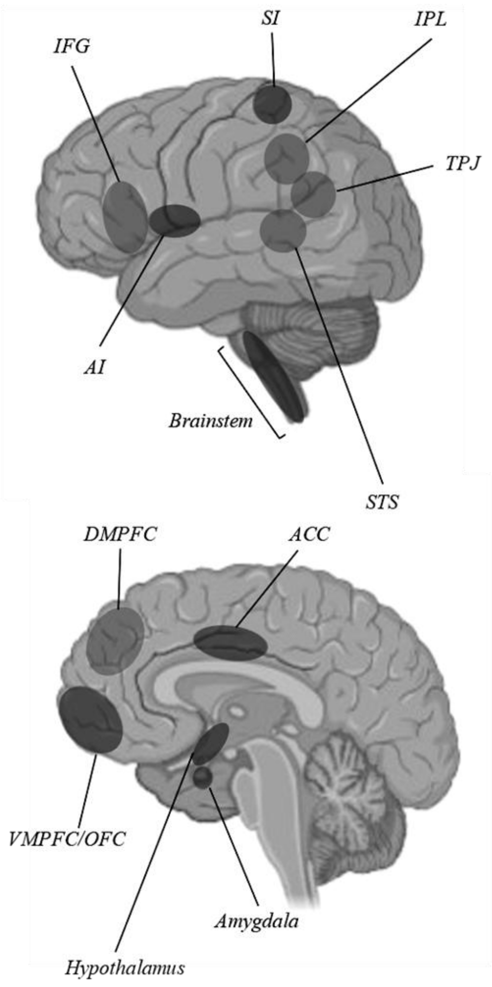
The empathy factor

Empathy reflects a complex interplay between emotional, cognitive, and motivational processes that resist simplistic categorisation. As such, the exploration of empathy as a construct has resulted in different perspectives and theories regarding its nature and definition, which have been thoroughly discussed elsewhere (e.g., Baron-Cohen & Wheelwright, 2004; Decety & Jackson, 2004; Preston & De Waal, 2002). In this thesis, I embrace a consensus-driven framework that recognises empathy as a multifaceted construct that can be dichotomised into its cognitive and affective components (Blair, 2005). Broadly, these elements respectively refer to the ability to understand others' emotions and experiences, as well as to resonate with and care about them (Decety & Jackson, 2004). Both facets are essential for meaningful social interactions, enabling individuals not only to connect with one another but also to respond in ways that are socially appropriate and morally guided (Decety & Jackson, 2004; Decety & Yoder, 2016; Preston & De Waal, 2002). By contrast, impairments in the experience of empathy may lead to maladaptive behaviours that facilitate interpersonal harm, such as those observed in criminal psychopaths. Advances in neuroscience show that such impairments may result from the disconnect of brain regions that enable adaptive behaviour over the course of development (Blair et al., 2006).

A look into the empathic brain

Empathy relies on a distributed network of brain regions that work together to support various affective and cognitive processes – hereby referred to as the "empathy network" (Figure 1.1). Within this network, brain regions associated with affective empathy develop earlier than those linked to cognitive empathy (Shamay-Tsoory, 2011; Van Dongen, 2020). This early development primarily occurs in limbic and paralimbic brain regions responsible for automatic and visceral processes, such as the amygdala, anterior insula (AI), and the anterior cingulate cortex (ACC). In contrast, the cognitive and regulatory components of empathy – such as perspective-taking and emotional control – depend on more specialised neurocircuitry that matures later, typically during late adolescence and early adulthood (Arain et al., 2013; Blakemore, 2008; Caballero et al., 2016). These functions are largely supported by the prefrontal cortex (PFC) – critical for interpreting emotional cues, guiding decision-making, and enabling mentalising – although earlier-developing structures like the temporo-parietal junction (TPJ) also contribute by supporting mental state attribution[4] (Decety, 2010; Marsh, 2018). This developmental trajectory underscores the foundational role of affective processes in empathy, with the gradual integration of cognitive faculties indicating that empathy becomes more refined and complex with increasing age (Decety, 2010; Decety et al., 2012).

Figure 1.1 | *The empathy network*



ABBREVIATIONS

ACC – Anterior Cingulate Cortex

AI – Anterior Insula

Amygdala

Brainstem

DMPFC – Dorsomedial Prefrontal Cortex

Hypothalamus

IFG – Inferior Frontal Gyrus

IPL – Inferior Parietal Lobule

SI – Primary Somatosensory Cortex

STS – Superior Temporal Sulcus

TPJ – Temporal Parietal Junction

VMPFC/OFC – Ventromedial Prefrontal Cortex/ Orbitofrontal Cortex

Studies have revealed that the activation and interconnection of brain regions involved in affective empathy closely resemble those engaged during one's own emotional experiences (Bird & Viding, 2014; Preston & De Waal, 2002). This overlap suggests a shared neural network for first-hand and second-hand affective processing, a phenomenon particularly evident in responses to witnessing others' pain (Marsh, 2018). Indeed, there is a large body of research showing that witnessing or imagining others in pain activates cortical and subcortical structures that are involved in experiencing pain directly (e.g., Decety, 2010, 2015; Lamm & Majdandžić, 2015; Zaki et al., 2016; see Lamm et al., 2011 for a meta-analysis) – with pain-related brain regions collectively referred to as the *pain matrix* (see Iannetti & Mouraux, 2010 for a review; further discussion in Supplement 1.1). For example, a study by Zaki *et al* (2007) showed that painful thermal stimulation elicited similar activation in the limbic system than watching videos of others experiencing pain, particularly in the ACC and AI (see also Morrison et al., 2004). Similar activation patterns in these brain regions have been found in studies using other stimuli such as facial expressions or sounds (Duerden et al., 2013; Keysers et al., 2010; Singer et al., 2004). Notably, the activation of empathy-related brain regions within the pain matrix appears to be specific to the nature of the task. While simply witnessing others in pain leads to increased activation in emotion-related areas, tasks that involve making inferences about others' pain additionally recruit brain regions associated with cognitive processing, such as the TPJ, precuneus, and medial PFC (Bruneau et al., 2015; Tusche et al., 2016; Yao et al., 2016). This indicates that empathising with others' pain requires not only shared affective experiences but also the perception of such pain as external – which is commonly known as *self-other differentiation* (Bird & Viding, 2014a).

Deficits in the ability to empathise with others' pain could lead to maladaptive behaviours. This is because pain empathy[5] is theorised to serve an important evolutionary function by facilitating threat detection and encouraging behaviours that maintain species preservation – such as social bonding and group cohesion (Lamm et al., 2011). Individuals who fail to exhibit these adaptive responses, by contrast, are more likely to develop antisocial tendencies and engage in harmful behaviours at the expense of others' integrity and wellbeing. Criminal psychopaths represent the epitome of these consequences, which may be explained by a widespread dysfunction across the neural networks involved in empathy and emotional processing (Bird & Viding, 2014a; Van Dongen, 2020). A growing body of research demonstrates that psychopathy is associated with hypoactivity in key areas within the empathy network, particularly during tasks requiring emotional engagement or processing others' distress (Anderson & Kiehl, 2013; Ermer et al., 2013; see Kiehl & Hoffman, 2011 for a review).

For instance, studies have linked psychopathy to aberrant activation of the AI and ACC during pain observation (Seara-Cardoso et al., 2015), as well as structural anomalies and hypoactivation in the amygdala (Birbaumer et al., 2005; Blair, 2003; Dolan & Fullam, 2009; Vieira et al., 2015). Interestingly, research suggests that this hypoactivation might be specific to psychopaths who engage in predatory violence (Kiehl et al., 2001), as it is not typically observed in psychopathic offenders who engage in more reactive aggression – who often display the opposite pattern (Brower & Price, 2001). These findings align with the notion of psychopathy consisting of primary and secondary variants; they also imply that impaired affective processing may underlie aggression in psychopathy regardless of subtype (Heym et al., 2019; Van Dongen, 2020). Additionally, studies have identified reduced efficiency in neural communication among both local and distal brain regions in individuals with higher levels of psychopathy, suggesting a more generalised disruption in network connectivity that may hinder the dynamic integration of cognitive and affective processes (Espinoza et al., 2018; Tillem et al., 2018). These impairments extend to frontal regions associated with evaluative and regulatory functions, with research showing abnormal activation of the PFC during moral and emotional processing. For instance, Decety *et al* (2013) found that criminal psychopaths exhibited reduced activity in the ventromedial prefrontal cortex (VMPFC) and lateral orbitofrontal cortex (OFC) – involved in moral decision-making and emotional regulation – when viewing others in pain. Complementing these findings, research using brain stimulation has further established causal links between these regions and empathic processing. For instance, excitatory stimulation to the dorsolateral prefrontal cortex (DLPFC) has been found to increase sensitivity to others' pain (Wang et al., 2014) and reduce aggressive behaviour in offenders with high levels of psychopathy (Molero-Chamizo et al., 2019). These findings not only validate the involvement of specific brain regions in the experience of empathy but also suggest the potential of neuromodulation to target empathy-related brain regions as a means of mitigating aggression, offering a promising direction for psychopathy treatment (see Sergiou et al., 2020 for a review). In exploring this potential avenue, however, we must question whether the experience of empathy itself influences aggression.

Exploring the empathy-aggression link

The presumed relationship between empathy and aggression is rooted in the notion that prosocial and antisocial behaviours are two opposing ends of the same spectrum, implying that an increase in one should correspond to a decrease in the other (Miller & Eisenberg, 1988a). This stems from early research suggesting that individuals with higher levels of empathy tend to engage in charitable and altruistic behaviours (e.g., Batson et al., 1983; Underwood & Moore, 1982), whereas those with low empathy are more likely to display unrestrained aggression (Marshall & Marshall, 2011; Miller & Eisenberg, 1988). However, the empirical picture is surprisingly mixed (Baron-Cohen & Wheelwright, 2004; Vachon et al., 2014a). While meta-analytic evidence shows slight to moderate reductions in empathy among aggressive cohorts (Jolliffe & Farrington, 2004; Miller & Eisenberg, 1988; Morrow, 2020), some studies fail to establish significant correlations between empathy and aggression (e.g., Gantiva et al., 2021; see Vachon et al., 2014 for a discussion). These discrepancies may be partially attributed to simplistic theoretical assumptions.

As mentioned above, the debate about whether empathy and aggression are related often relies on the idea that traits that encourage prosocial behaviours should naturally decrease tendencies for harmful and antisocial behaviour, and *vice versa*. However, assuming that a characteristic found in two seemingly opposing phenomena must inherently drive these outcomes is, in essence, a misleading and spurious correlation. In fact, while empathy equips individuals with the tools to understand and appropriately respond to others' emotions, it does not inherently guarantee a motivation for prosocial behaviour. In some instances, individuals may be emotionally attuned to others yet remain unmotivated to help, particularly when doing so would require significant effort or self-sacrifice (Decety et al., 2016). Additionally, it is proposed that while compassion and genuine concern for others serve as vital catalysts for prosocial behaviour, empathic responses such as emotionally resonating with others' pain can sometimes escalate into personal distress[6], thereby prompting a negative (reactive) behavioural response (Miller & Eisenberg, 1988a). As such, empathy can also lead to aggression under specific circumstances, as seen for instance in cases where empathising with victims of aggression paradoxically fosters aggressive attitudes towards the aggressor (see e.g., Vitaglione & Barnett, 2003). Furthermore, experiences of personal distress may encourage prosocial outcomes but for self-serving purposes – like alleviating one's own internal discomfort (Miller & Eisenberg, 1988a). This further highlights the need to differentiate between the broader concept of empathy and the more targeted notions of sympathy or compassion, which specifically relate to the motivational drive behind prosocial actions (Marsh, 2016; Van Dongen, 2020).

On the other hand, individuals who lack empathy may not engage in aggression unless other contributing factors – such as, impulsivity, personal gain, or social reinforcement – are also present. This is nicely exemplified in the case of the so-called *successful* psychopaths, as these individuals do not typically engage in overt antisocial behaviour despite their affective deficits (Hall & Benning, 2006). This suggests that, in the case of psychopathy, it might not be the absence of empathy what enables aggressive behaviour, but rather the specific configuration of empathic abilities – and, critically, how they are used. Therefore, the nature of the association between empathy and aggression appears to be contingent on the underlying motive of the aggressive behaviour, which further emphasises the need to move away from one-size-fits-all explanations when it comes to understanding the empathy-aggression link.

Summary

Overall, the reviewed evidence supports the understanding of psychopathy as a spectrum, meaning that not all individuals with psychopathic traits are dangerous offenders, and many do not meet the clinical criteria for psychopathy. However, research consistently points to callous-unemotional traits as particularly significant risk factors for the development of psychopathy and associated aggressive behaviours. These traits are notably more stable over time than other psychopathic features and are marked by a deficit in the empathic experience, reflected in atypical functioning and disrupted connectivity within the brain's empathy network. Individuals exhibiting these deficits are less likely to resonate with others' suffering and, as a result, may be more prone to engage in severe forms of aggression, including premeditated or predatory violence. Taken together, these findings suggest that variations across the psychopathy spectrum can be better understood through the interplay between empathy and aggression, which is the premise I followed in the development of my research.

Overview of the project

This research investigates the interplay between empathy, aggression, and the subclinical expression of callous-unemotional traits in community samples. By adopting a subclinical perspective, the thesis aims to enhance our understanding of how socio-affective traits can predict the risk of cruelty among non-clinical and non-criminal populations. The first objective was to determine whether different facets of empathy mediate the relationship between interpersonal callousness and aggressive behaviour in young adults. This investigation was conducted through two online surveys within community-based samples, reported in ***Chapter II***. Next, I sought to investigate individual differences in empathy and aggression proneness within a moral framework, building on discussions that suggest that aggressive behaviours resulting from empathy (or the lack thereof) may be construed as either adaptive or maladaptive depending on their underlying motivations. For example, going back to Ted Bundy's case; while the public's response to his execution can be described as hostile and aggressive, it can also be understood through the lens of morality[7], illustrating how selective empathy shapes biases in moral reasoning. In ***Chapter III***, I sought to investigate the effects of selective empathy for victims *vs* aggressors on moral decision-making through two online experiments, additionally analysing the potential effects of callous-unemotional traits in reducing these biases in the second experiment. In ***Chapter IV***, I further examine whether the expression of callous-unemotional traits can predict individual differences in the vicarious experience of pain, building on research suggesting that aggression in psychopathy may result from a diminished ability to empathise with others' pain. Using electroencephalography (EEG) and machine learning, the study aimed to identify predictable patterns in EEG responses to pain and examine how these patterns vary with callous-unemotional traits. While most studies on pain empathy use functional magnetic resonance imaging (fMRI) for its high spatial resolution, EEG offers unique insights into the temporal dynamics of neural responses to observed pain. In addition, the use of machine learning allows to further explore these dynamics on a trial-by-trial basis. In my last empirical chapter (***Chapter V***), I examine the potential of improving callous-unemotional symptoms via non-invasive brain stimulation (NIBS). This investigation was conducted through a systematic review and meta-analysis examining the effects of NIBS on socio-affective states typically impaired in individuals with pronounced callous-unemotional traits – focusing on empathy-, prosocial- and guilt-related emotions. Lastly, ***Chapter VI*** provides a comprehensive summary and broad discussion of the described studies. In this final chapter, I contextualise the work included in my thesis within the broader landscape of research in the field, elucidating how it builds upon and expands previous findings. Moreover, I address the limitations inherent in the current studies and propose potential avenues for future research.

Authorship and contributions

Studies 1 and 2

Main contributions – *Célia F Camara*: Project planning, experiment design, data collection, analyses and visualisation, manuscript writing; *Paul HP Hanel*: Project planning, experiment design, supervision of data analyses and manuscript revision; *Alejandra Sel*: Project planning and supervision, funding acquisition, manuscript revision; *Carina CJM de Klerk*: Project supervision, manuscript revision.

Additional contributions – *Phoebe E Harris, Jelena Devcic, Imay Soysal, Emily Rice*: Data collection (Study 1); *Alana Dale-Flowers, Dominik Buric, Amy C Wright, Millie R Snook, Vinisha V Chauhan, Laura J Wyatt*: Data collection (Study 2).

Publication – Camara C. F., Sel A., de Klerk C. C., & Hanel P. H. (2025). Direct and indirect effects of interpersonal callousness on aggression through empathy and moral disengagement. *Personality and Individual Differences*, *232*, 112836. https://doi.org/10.1016/j.paid.2024.112836

Studies 3 and 4

Main contributions – *Célia F Camara*: Project planning, experiment design, data collection, data analyses and visualisation, manuscript writing; *Paul HP Hanel*: Project planning, experiment design, supervision of data analyses and manuscript revision; *Alejandra Sel*: Project planning and supervision, funding acquisition, manuscript revision; *Carina CJM de Klerk*: Project supervision, manuscript revision.

Additional contributions – *Phoebe E Harris, Jelena Devcic, Imay Soysal, Emily Rice*: Data collection (Study 3); *Alana Dale-Flowers, Dominik Buric, Amy C Wright, Millie R Snook, Vinisha V Chauhan, Laura J Wyatt*: Data collection (Study 4); *Dr Rael Dawtry*: Feedback on experiment design.

Publication – Camara C. F., Sel A., & Hanel P. H. (*in press*). When empathy leads to aggression: The effects of empathy on punitive attitudes towards aggressors. *British Journal of Social Psychology*.

Study 5

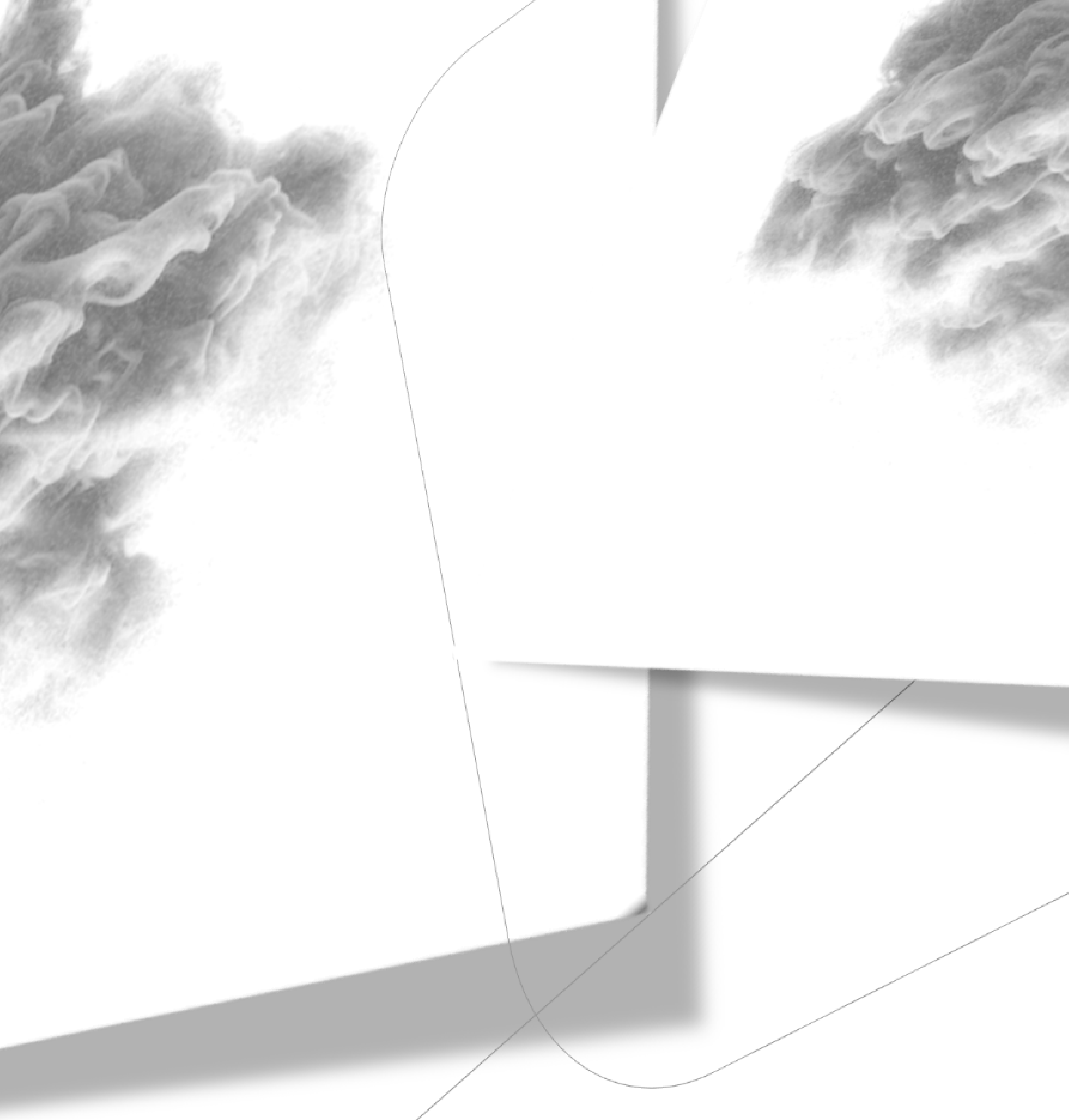
Main contributions – *Célia F Camara*: Project planning, experiment design, data collection, editing of EEG scripts, EEG and behavioural analyses and visualisation, manuscript writing; *Sebastian Halder*: Experiment design, machine learning analyses and visualisation, manuscript revision; *Alejandra Sel*: Project planning, experiment design, writing of EEG scripts, supervision of analyses, funding acquisition, manuscript revision; *Carina CJM de Klerk*: Project planning, supervision of EEG analyses, manuscript revision..

Additional contributions – *Arkadij Lobov*: development of EEG scripts*; Karolina Kudzia, Daniel Chandra Sehar*, *Andrea Espinosa Solis*: Data collection.

Study 6

Main contributions – *Célia F Camara*: Project planning, experiment design, data screening and extraction, writing and editing of scripts for meta-analyses, data analyses and visualisation, manuscript writing; *Carmen Sergiou*: Project planning, experiment design, full-text screening, data extraction, manuscript revision; *Andrés Molero Chamizo*: Project planning, experiment design, abstract and title screening, manuscript revision; *Nathzidy Rivera Urbina*: Project planning, experiment design, abstract and title screening, manuscript revision; *Paul HP Hanel:* Project planning, experiment design, supervision of analyses, manuscript revision. *Michael A Nitsche*: Project planning, experiment design, manuscript revision; *Alejandra Sel*: Manuscript revision.

Additional contributions – *Sara Ascensao*: Initial abstract and title screening; *Khetam Alfaraj*: Provision of script for meta-analysis.



Chapter II.

The Influence of Empathy and Moral Disengagement in Callousness and Aggression

Abstract

The expression of interpersonal callousness in psychopathy has been often associated with antisocial and aggressive behaviour. While empathy deficits are often implicated in this link, few studies have examined how distinct components of empathy may differentially contribute to these outcomes. The studies presented in this chapter seek to address this gap. In Study 1, we predicted and found that lower levels of affective empathy were associated with higher interpersonal callousness. Importantly, affective empathy deficits mediated the relationship between interpersonal callousness and both proactive aggression and social deviance. Study 2 replicated these findings and further revealed that the mediating role of affective empathy on proactive aggression was moderated by individuals' tendencies toward moral disengagement. Taken together, these findings suggest that affective empathy, more so than cognitive empathy, plays a critical role in predicting antagonistic behaviour among individuals exhibiting interpersonal callousness. Additionally, they indicate that moral disengagement may exacerbate these risks, highlighting the interplay between emotional deficits and moral reasoning in aggression proneness.

"…the predictability of individuals' behaviors can be augmented by aggregating these behaviours into a composite profile."

– Loeber & Dishion (1983, pp.81–82)

Introduction

While often conceptualised under the antisocial scope, research suggests that psychopathy is fundamentally rooted in deficits in the affective experience of the psychopath (Cleckley, 1941; Hare & Hart, 1993; Seibert et al., 2011). These deficits are reflected in the expression of callous-unemotional traits, which characterise individuals with diminished empathy, shallow emotions, and a disregard for the feelings and rights of others (Frick et al., 2003). As previously discussed, both psychopathy and callous-unemotional traits can be understood as existing along a continuum – ranging from subclinical manifestations in the general population to more severe (clinically significant) presentations, such as those observed in criminal psychopaths (Frick et al., 2003; Frick & White, 2008; Kimonis et al., 2008). At the more severe end of this spectrum, individuals show pronounced empathy deficits and a greater likelihood of causing interpersonal harm (Neumann & Hare, 2008). While existing research suggests that these patterns are primarily linked to impairments in affective empathy, the specific roles of different empathy components in driving these outcomes – especially when expressed at subclinical levels – have not been sufficiently explored.

In this chapter, I delve into how the subclinical expression of interpersonal callousness may predict the likelihood of engaging in aggressive and antisocial behaviours within community samples, examining the roles of affective and cognitive empathy facets in mediating these effects. Additionally, I draw on findings from research on moral disengagement to further explore nuances in these associations.

Links between callousness and aggression in psychopathy

Psychopathy is often marked by aggressive tendencies that may escalate into predatory violence – wherein individuals engage in calculated and controlled acts of physical harm inflicted on others for personal gain or gratification (Glenn & Raine, 2009). While this type of aggression is more common among serious offenders (Brugman et al., 2017), some scholars argue that general expressions of proactive aggression may also reflect the severity of aggressive tendencies in non-criminal populations (Bushman & Anderson, 2001; Frick et al., 2003). Theoretically, this association is attributed to the affective and interpersonal deficits characterising the construct of psychopathy – as described in the PCL–R (Hare, 2003) –, given that calculated and predatory-like behaviours involve some level of callousness, manipulation and reduced responses to distress (Glenn & Raine, 2009; Hare & Neumann, 2008; Meloy et al., 2018; Saladino et al., 2021; Woodworth & Porter, 2002). Recognising that these features are not exclusive to offenders or psychiatric inpatients, efforts have been made to develop assessments of psychopathy suitable for non-forensic and subclinical populations. One prominent example is the Self-Report Psychopathy Scale (Williams & Paulhus, 2004), with its most recent version – Self-Report Psychopathy Scale – 4th Edition (SRP–4; Paulhus, 2016) – now widely used to assess psychopathic traits in community samples.

The SRP–4 is designed to align with the PCL–R's factor model, delineating distinct facets of psychopathy across two factors. The first factor encompasses affective and interpersonal traits that collectively capture the essence of interpersonal callousness. Closely related to the concept of callous-unemotional traits, interpersonal callousness taps into a broader manifestation of affective deficits observed in psychopathic cohorts – including shallow affect, superficial charm, manipulativeness, lack of remorse, and failure to accept responsibility for one's actions (Paulhus, 2016; Williams & Paulhus, 2004). These traits denote a marked inability (or unwillingness) to resonate with and care about others' emotions, which enables manipulative and exploitative behaviours with little remorse or guilt (Cleckley, 1941; Neumann & Hare, 2008). This pattern gives rise to what is termed *callous aggression*, wherein acts of aggression are carried out with little emotional engagement or concern for others. In contrast, the second factor pertains to the externalising dimension of psychopathy, typified by socially deviant and antisocial tendencies such as substance use, impulsivity and aggression. While these behaviours are relevant to the clinical expression of psychopathy, they are less central in subclinical populations and do not necessarily predict callous aggression (Seibert et al., 2011).

The role of empathy in callous aggression

In psychopathy, empathy deficits can manifest in both cognitive and affective aspects, yet deficits in affective empathy are more characteristic of interpersonal callousness. For example, studies have found that higher levels of callousness correlate with reduced reactivity to others' pain, reflecting lower affective resonance (Frick & White, 2008; Jones et al., 2009; Marsh et al., 2013). It is hypothesised that this reduced affective response contributes to their propensity for proactive aggression, as it may desensitise individuals who exhibit these traits from the suffering of others (Decety & Yoder, 2016; Lovett & Sheffield, 2007; Marshall & Marshall, 2011; Miller & Eisenberg, 1988b). Empirical studies reinforce this hypothesis. For example, both proactive aggression and interpersonal callousness have shown negative associations with dimensions of affective empathy such as affective resonance and empathic concern (see Vachon & Lynam, 2016). Nonetheless, beyond merely indicating a lack of emotional connection, the premeditation of harmful behaviours towards others specifically denotes an element of malicious intent (Glenn & Raine, 2009; Woodworth & Porter, 2002). Considering this, Vachon and Lynam sought to provide a conceptualisation of empathy that could also be more consistent with descriptions of clinical disorders like psychopathy. In 2016, they developed the *Affective and Cognitive Measure of Empathy* (ACME; Vachon & Lynam, 2016), which offers a more differentiated approach to affective empathy.

The ACME offers a novel and more nuanced approach to measuring affective empathy by distinguishing between *affective resonance*, which refers to an individual's capacity to emotionally respond to others' experiences, and *affective dissonance,* which captures the extent to which an individual experiences aversive or antagonistic emotional reactions to others' emotions. The concept of affective dissonance, in particular, allows for a more precise understanding of how empathy deficits manifest in the context of interpersonal callousness and instrumental aggression. Unlike traditional measures of empathy – which assume a unidirectional relationship between empathic capacity and aggression –, affective dissonance captures the presence of inversely aligned emotional responses, making it particularly suited to identifying behaviours potentially driven by a sense of gratification derived from others' suffering. Vachon and Lynam (2016) found that affective dissonance scores were more strongly associated with various forms of aggression – including proactive and reactive aggression – than traditional measures of affective empathy. These findings have since been replicated in both forensic and community samples (e.g., Eman et al., 2022; Levitan & Vachon, 2021). Notably, a study conducted by Preston and Anestis (2020) further showed that the ACME total scores significantly mediated the relationship between proactive aggression and callous traits, further suggesting the utility of this scale in addressing the role of empathy in callous aggression.

On the other hand, research suggests that interpersonal callousness does not consistently involve impairments in cognitive empathy (Vachon & Lynam, 2016). In fact, individuals with these traits can be relatively good at understanding others' emotions (Harris & Rice, 2006; Kiehl & Hoffman, 2011; Polaschek, 2014; Rice et al., 1992; Sutton et al., 1999). In this regard, cognitive empathy may paradoxically facilitate planned, instrumental aggression among those with elevated interpersonal callousness, as their ability to discern others' emotions enhances their capacity to exploit weaknesses without being emotionally moved by them (Harris & Rice, 2006; Kiehl & Hoffman, 2014; Polaschek, 2014; Rice et al., 1992; Sutton et al., 1999). Conversely, deficits in perspective-taking – a fundamental component of cognitive empathy – appear more strongly associated with reactive forms of aggression (Blair, 2013; Chang et al., 2021; Jolliffe & Farrington, 2011; Vachon et al., 2014b; B. A. White et al., 2015). These findings suggest that individual differences in perspective-taking may underlie distinct aggression profiles, such as those differentiating primary and secondary psychopathy. However, the precise role of other facets of cognitive empathy in aggression warrants further investigation.

Notably, cognitive empathy also involves subcomponents that are discernible from the concept of *perspective-taking,* which specifically refers to the ability to understand and adopt another person's point of view – often used interchangeably with terms like *mentalising* or *theory of mind (ToM)* (Baron-Cohen et al., 1995; Tager-Flusberg et al., 2000). Another important facet of cognitive empathy is *empathic accuracy*, which refers to the ability to accurately perceive and understand another person's emotional state or intentions (Ickes, 1993, 2001; Zaki et al., 2008; Zaki & Ochsner, 2012). Additionally, *emotional understanding* represents a further dimension of cognitive empathy that extends beyond merely recognising an emotion, involving both the identification and interpretation of others' emotions in different contexts (Vachon et al., 2014b; Vachon & Lynam, 2016; Zaki & Ochsner, 2012). Differentiating between these subcomponents is essential when examining the distinct roles that facets of empathy may play in the link between interpersonal callousness and aggression.

Beyond empathy: insights from moral disengagement

Beyond highlighting empathy deficits, callous aggression also reveals a pronounced lack of moral sensibility (Erzi, 2020; Porter et al., 2003; Shulman et al., 2011). Those who pursue personal goals at the expense of others – employing tactics like manipulation – may justify or dismiss the consequences of their immoral actions, even when they can recognise such actions as wrong (Gini et al., 2014). This phenomenon – known as *moral disengagement* – is typically facilitated by cognitive strategies that reduce the psychological discomfort associated with violating one's moral standards (Bandura, 1990; Bandura et al., 1996). Common mechanisms include displacing responsibility, minimising the harm caused, and dehumanising victims. In this way, moral disengagement functions both as a psychological buffer and as a tool that enables aggression without the accompanying guilt or self-condemnation (Caprara et al., 2014; Gini et al., 2015). For instance, in emotionally charged situations involving provoked aggression, individuals may use moral disengagement to preserve a positive self-image by attributing blame to the provocateur or downplaying the severity of their actions. However, this is assuming that the individual is predisposed to such feelings of guilt in the first place.

In fact, Bandura's theory of moral disengagement (1990; 1996) also suggests that the link between moral disengagement and diminished guilt may not simply reflect a coping mechanism, but rather a more fundamental dispositional tendency. That is, individuals who are naturally inclined toward moral disengagement may experience less guilt, rather than using cognitive strategies to alleviate guilt after transgressions. This distinction is particularly relevant when considering individuals with callous traits. For example, while moral disengagement is less likely to occur if the perceived harm inflicted outweighs the personal benefit (Miller & Eisenberg, 1988a), callous individuals may still justify their immoral actions by disregarding or downplaying the potential harm they cause (Shulman et al., 2011). Research also shows that individuals exhibiting callous traits are typically more tolerant to moral violations (Erzi, 2020; Gini et al., 2014, 2015; Risser & Eckert, 2016; Shulman et al., 2011; X. Wang et al., 2017). Over time, this tolerance may evolve into an indifference or habituation towards immoral actions, thereby reinforcing callous tendencies and potentially escalating them towards more severe forms of aggression (Hyde et al., 2010; Shulman et al., 2011).

Supporting this premise, studies consistently show that moral disengagement predicts the engagement in bullying and other types of violent behaviours in adolescents (Falla et al., 2021; Paciello et al., 2008; Sticca & Perren, 2015). Additionally, moral disengagement in adolescence has been found to predict criminal behaviour in early adulthood (Fontaine et al., 2014). Evidence from meta-analytic findings further indicates moderate relationships between moral disengagement and aggression (Gini et al., 2014; Luo & Bussey, 2023; Wang et al., 2014). In particular, a meta-analysis conducted by Gini *et al*’s (2014) found that moral disengagement moderates the relationship between grandiose-manipulative traits and proactive aggression. This suggests that moral disengagement may specifically facilitate calculated, goal-oriented aggression when paired with emotional detachment and self-serving interpersonal styles, indicating potential implications of empathic processing. Indeed, studies also show that the link between empathy and aggression can be influenced by the expression of moral disengagement (Chowdhury & Fernando, 2014; Hyde et al., 2010; Paciello et al., 2013; Wang et al., 2017). Consequently, the interplay between moral disengagement and poor empathy may be a key factor in understanding why individuals with callous traits might be more likely to engage in aggression. However, no study to date has explicitly explored how empathy and moral disengagement interact to influence aggressive behaviour in the context of interpersonal callousness. Investigating this could provide further insights into the role of empathy in callous aggression.

Project rationale

In sum, previous research indicates that the emotionally detached disposition characteristic of callous individuals elevates their risk for engaging in severe and persistent forms of aggression, suggesting greater importance of deficits in affective empathy than cognitive empathy in the link between interpersonal callousness and antagonistic behaviour (e.g., Mullins-Nelson et al., 2006; Neumann & Hare, 2008). Despite this, there is little evidence regarding how distinct subcomponents within each empathy facet influence the link between interpersonal callousness and externalising behaviour (see Preston & Anestis, 2020 for an example). Additionally, research shows that moral disengagement plays a key role in facilitating aggression (Gini et al., 2014), with evidence even suggesting that moral disengagement may exacerbate the influence of empathy deficits on aggression (Wang et al., 2017). However, no study to date has directly tested how moral disengagement interacts with empathy to affect callousness and aggression.

To address these gaps, our study investigates the links between interpersonal callousness and aggression within a community sample of young adults through the mediation of empathy facets. Recognising the nuances in affective and cognitive empathy, we used two complementary scales to assess different subcomponents related to each facet of empathy. Additionally, we extend previous similar research by not only distinguishing between reactive and proactive aggression (Preston & Anestis, 2020), but also including a wider spectrum of socially deviant behaviours linked to psychopathy – ranging from minor rule-breaking to more overt aggression. In a follow-up study, we aimed to replicate these effects and further explored the mediation effects of both affective and cognitive empathy subcomponents through the mediation of moral disengagement.

Building on previous findings, we hypothesised that the link between interpersonal callousness and proactive aggression would be mediated by reduced affective empathy (Preston & Anestis, 2020). Conversely, we predicted that deficits in perspective-taking would mediate the association between interpersonal callousness and reactive aggression (Chang et al., 2021; Vachon et al., 2014). Finally, we predicted that the indirect pathways from callousness to aggression via empathy would be stronger among individuals who show higher levels of moral disengagement (Wang et al., 2017). Data and scripts have been made available and can be accessed at: https://osf.io/c28n9/?view_only=f545e5192af940d69e693ab4b31afcd6

Study 1

Methods

Participants and procedure

This study was approved by our local Ethics Committee and advertised on a departmental undergraduate research portal, where students participate in for class credit. Outreach efforts also included email invitations and word-of-mouth referrals, involving recruitment of participants not affiliated with the university. External participants joined the survey as part of a broader online study with the incentive of entering a cash prize raffle upon survey completion. Study completion typically took around 25 minutes.

A power analysis revealed that, to detect a medium effect size of $r = 0.34$ with a power of 0.95, we needed 106 participants. Initially, we recruited 124 participants, although 11 participants did not complete the survey, and three failed at least 2 of the 4 attention checks. Excluding an additional participant aged 17 due to the study's exclusive focus on adults, the final analysis involved 109 participants aged 18 to 31, $M = 21.58$, $SD = 2.55$. This sample comprised 33 men, 75 women, and one non-binary/third-gender participant. The majority were university students ($N = 84$), and only 12 participants disclosed having a diagnosed mental health condition, including depression, anxiety, attention deficit and hyperactivity disorder/attention deficit disorder (ADHD/ADD), or borderline personality disorder.

Survey questionnaires

Self-Report Psychopathy Scale – Short Form (SRP–SF; Paulhus et al., 2016). The SRP–SF is a 29-item instrument adapted from the full SRP–4 (Paulhus et al., 2016). The SRP–SF was developed to reflect the two-factor model of psychopathy outlined by the PCL–R (Hare, 2003) making it well-suited for use in non-forensic and community samples (see Dotterer et al., 2017). The SRP–SF comprises two higher-order factors: Factor 1, consisting of the

Interpersonal Manipulation and *Callous Affect* subscales (α = 0.83), and Factor 2, composed of the *Erratic Lifestyle* and *Antisocial Behaviour* subscales (α = 0.74). Factor 1 captures the core affective and interpersonal traits of psychopathy – such as superficial charm, lack of empathy, and manipulativeness –, whereas Factor 2 is mostly reflective of externalising behaviours such as aggression and rule-breaking tendencies (Hare & Neumann, 2008; Patrick et al., 2009). Studies have discussed that the inclusion of criminal and antisocial traits as integral parts of psychopathy might limit the applicability of the SRP–SF in non-forensic settings (Williams & Paulhus, 2004), which is consistent with the understanding that psychopathic traits within community samples are less defined by its antisocial facet. This justifies the focus on Factor 2 as a more consequential outcome rather than a required component within the psychopathy spectrum. Building on this theoretical framework, the present study conceptualises Factor 2 as an outcome variable representing social deviance, examining its predictability through the interpersonal and affective traits defined by Factor 1.

Buss and Perry Aggression Questionnaire (BPAQ; Buss & Perry, 1992). The BPAQ is a 29-item scale often used as a generalised assessment of aggression, including measures of physical aggression, verbal aggression, anger, and hostility. In this study, we used the *Physical Aggression* subscale (e.g., “I have physically threatened people I know”, $\alpha = 0.81$) as it directly reflects overt expressions of violence.

Reactive and Proactive Aggression Questionnaire (RPQ; Raine et al., 2006). The RPQ includes 23 items to assess the functions of aggression, including a *Reactive Aggression* subscale (e.g., “Yelled at others when they have annoyed you”, $\alpha = 0.81$), and a *Proactive Aggression* subscale (e.g., “Yelled at others so they would do things for you”, $\alpha = 0.69$). Notably, the RPQ includes items reflecting violent behaviour (e.g., “Used physical force to get others to do what you want”).

Interpersonal Reactivity Index (IRI; Davis, 1980). The IRI is a widely used 28-item self-report questionnaire that measures cognitive empathy via perspective-taking and affective empathy via empathic concern. Additionally, the IRI includes two supplementary constructs (i.e., *Fantasy* and *Personal Distress* subscales) assumed to be associated with dispositional levels of empathy. However, given the study's focus on affective and cognitive empathy facets, we only considered participants' responses in the *Perspective-taking* (e.g., "When I'm upset at someone, I usually try to put myself in their shoes", α = 0.81) and *Empathic Concern* (e.g., "I am often quite touched by things that I see happen.", α = 0.77) subscales for our analysis. Items for each measure were reverse-coded, such that higher scores represent lower levels of empathy – an approach followed in previous similar research (Levitan & Vachon, 2021).

Affective and Cognitive Measure of Empathy (ACME; Vachon & Lynam, 2016). The ACME scale also covers traditional forms of empathy, including items related to empathic accuracy and emotion understanding (e.g., "I can usually tell how people are feeling", α = .90), and items representing both affective resonance (e.g., "I feel awful when I hurt someone's feelings", α = 0.83) and affective dissonance (e.g., "It's funny to see people get humiliated", α = 0.91). All items in this scale were also reverse-coded, with the exception of the *Affective Dissonance* subscale and other pre-reverse-coded items (Levitan & Vachon, 2021).

Analysis

Descriptive statistics were first calculated for all study variables using IBM SPSS Statistics (Version 28.0). Next, we conducted bivariate correlations to explore the relationships between interpersonal callousness, empathy, and aggression. To ensure robust inference, 95% bias-corrected confidence intervals ($BCCI_{95\%}$) were estimated from 1000 bootstrap samples, and significance was determined at $p < 0.05$ using the Holm-Bonferroni correction to control for multiple comparisons.

Finally, we examined the mediating effects of empathy in the relationship between interpersonal callousness and aggression via Structural Equation Modelling (SEM), using the *lavaan* package in *R* programming language (Rosseel, 2012). For this analysis, we employed a multivariate approach, including empathy variables as mediators, interpersonal callousness as a predictor and aggression variables as outcomes, while also accounting for the covariance among empathy variables and aggression variables (see Figure 2.1). Mediation was assessed by examining the direct and indirect effects of interpersonal callousness on aggression through each empathy mediator. The total effect was the sum of the indirect effects through the mediators and the direct effect of the predictor. Significant interactions were probed using 1000 bootstrap samples from the original dataset, along with $BCCI_{95\%}$ where mediation was considered significant when the confidence interval did not include zero (MacKinnon et al., 2007). Multicollinearity among empathy mediators was also evaluated using variance inflation factors (VIFs), with values exceeding 2.71 used as a threshold to assess multicollinearity (Richter & Bixler, 2022).

Results

Sample characteristics

Reported in Table 2.1, participants levels of interpersonal callousness, empathy, and aggression fell within the normative ranges described in previous adult samples (e.g., Levitan & Vachon, 2021). This suggests that the sample was broadly representative of a non-clinical population in terms of socio-emotional functioning. No extreme outliers were identified, although we observed skewed distributions for aggression and interpersonal callousness (see results in Supplement 2.1).

Correlations

Bivariate correlations are reported in Table 2.1 below. As anticipated, interpersonal callousness positively correlated with all aggression outcomes (reactive, proactive, physical, and social deviance), with correlations ranging from moderate to strong – r values between 0.48 and 0.64. Additionally, participants with higher levels of interpersonal callousness reported lower empathy overall. In particular, interpersonal callousness positively correlated with more affective dissonance, $r = 0.71$, $BCCI_{95\%}$ [0.58, 0.81], $p < .001$, as well as more deficits in affective resonance, $r = 0.62$, $BCCI_{95\%}$ [0.46, 0.73], $p < .001$, and empathic concern, $r = 0.54$, $BCCI_{95\%}$ [0.37, 0.67], $p < .001$. A weaker but still significant association was observed with reduced perspective-taking, $r = 0.39$, $BCCI_{95\%}$ [0.22, 0.54], $p < .001$. However, there was no correlation with emotion recognition, $r = -0.08$, $BCCI_{95\%}$ [-0.27, 0.09], $p = .381$.

Table 2.1. | *Mean, standard deviation, and bivariate correlations for main variables in Study 1*

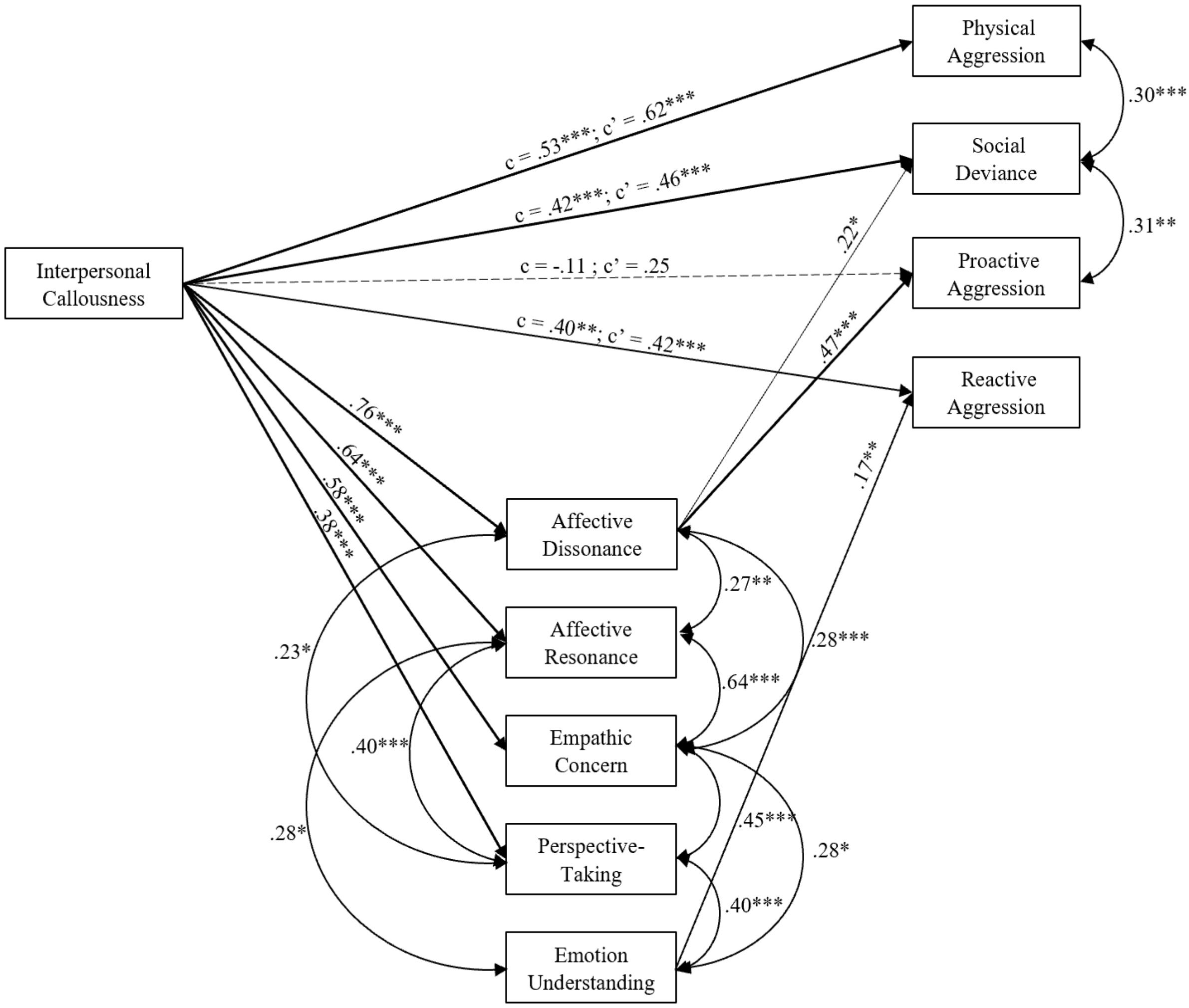
Variables	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
1.F1	2.30	.68	–									
2.F2	1.88	.49	**.63**	–								
3.React	.61	.35	**.62**	**.55**	–							
4.Proact	.09	.14	**.48**	**.53**	**.52**	–						
5.Phys	2.46	.77	**.64**	**.60**	**.57**	**.51**	–					
6.PT	1.43	.76	**.39**	.25	**.36**	.23	**.37**	–				
7.COG	2.25	.64	-.08	-.02	.14	-.01	.02	.28	–			
8.EC	1.25	.69	**.54**	.28	.29	**.35**	**.36**	**.55**	.17	–		
9.RES	1.89	.59	**.62**	**.37**	**.38**	**.37**	**.44**	**.53**	.13	**.74**	–	
10.DIS	1.68	.69	**.71**	**.54**	**.59**	**.59**	**.56**	**.50**	.01	**.53**	**.57**	–

Note. F1 = Interpersonal Callousness; F2 = Social Deviance; React = Reactive Aggression; Proact = Proactive Aggression; Phys = Physical Aggression; PT = Perspective Taking; COG = Emotion Understanding; EC = Empathic Concern; RES = Affective Resonance; DIS = Affective Dissonance. Higher scores in empathy scales denote larger deficits.

Significant values after Holm-Bonferroni correction in **bold**.

Moreover, affective empathy deficits were consistently correlated with aggression outcomes – *r* values between 0.29 and 0.59. In contrast, correlations with measures of cognitive empathy were more inconsistent. Specifically, we found significant correlations between reduced perspective-taking with reactive, $r = 0.36$, $BCCI_{95\%}$ [0.16, 0.54], $p < .001$, and physical aggression, $r = 0.37$, $BCCI_{95\%}$ [0.18, 0.53], $p < .001$, but no correlation with emotion understanding overall. All significant correlations remained after Bonferroni-Holm correction.

Figure 2.1. | *Empathy mediation on the links between callousness on aggression*



Note. Standardised path coefficients are displayed. Single-headed arrows represent significant direct effects, and curved double-headed arrows represent estimated covariances among empathy variables. For clarity, lines between empathy and aggression variables are omitted if they were non-significant.
Significance estimated at $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Discussion

Study 1 supports the notion that deficits in affective empathy play a more relevant role than cognitive empathy in the expression of callous psychopathic traits and their association with aggression (Mullins-Nelson et al., 2006). Our findings specifically highlight the importance of affective dissonance (over other aspects of affective empathy) in fostering socially deviant behaviours and premeditated aggression among individuals exhibiting these traits, which is consistent with prior investigations (see Vachon & Lynam, 2016).

However, a note of caution is warranted in interpreting these findings due to the overrepresentation of women in the study sample, particularly considering documented gender differences in empathy (e.g., Lui et al., 2016; Van Hazebroek et al., 2017), psychopathy (e.g., Ciucci & Baroncelli, 2014; Colins et al., 2017), and aggression (e.g., Berkout et al., 2011; Knight et al., 2002). This raises questions about the generalisability of the results. Moreover, the reliance on a predominantly student-based sample further restricts the applicability of our conclusions to broader demographics (Hanel & Vione, 2016).

Study 2

Method

Participants and procedure

In Study 2, we aimed to enhance the reliability and generalisability of our findings in Study 1 by recruiting a larger and more diverse sample. To this end, participants were recruited via Prolific (www.prolific.co) as part of a broader online study and were compensated £7 for their participation. To maintain consistency with Study 1 and reduce age-related variability, we focused exclusively on young adults aged 18 to 25. In total, data were collected from 319 respondents. However, nine participants were excluded due to timeouts or failed attention checks, and one was excluded due to missing age data. Study completion among the remaining participants took around 27 minutes on average.

Our final sample included 310 participants aged 18 to 25, $M = 22.75$, $SD = 1.75$, including 161 men, 146 women and 3 participants who identified as non-binary/third gender. Employment status varied, with 140 participants being employed, 87 students, and 55 unemployed. Other categories included 18 self-employed, 3 homemakers, and 1 participant both studying and working. Additionally, 68 participants reported a mental health diagnosis, encompassing various disorders including anxiety, depression, ADHD/ADD, bipolar disorder, post-traumatic stress disorder, Tourettes syndrome, derealization disorder, emotionally unstable personality disorder, obsessive compulsive disorder, and anorexia.

All procedures in this study have been approved by the university Ethics Committee.

Survey questionnaires

In this study, we employed the same survey questionnaires used previously to assess interpersonal callousness (SRP–SF Factor 1) and social deviance (SRP–SF Factor 2), aggression (BPAQ and RPQ), and empathy (IRI and ACME). Overall, the scales demonstrated strong internal consistency, with Cronbach's alpha coefficients ranging from 0.79 to 0.89 (specific reliability coefficients reported in Supplement 2.2).

Additionally, we assessed moral disengagement using the *Moral Disengagement Scale* (MDS; (Bandura et al., 1996). This is a 32-item self-report questionnaire that includes various facets of moral disengagement, presenting participants with statements such as "Damaging some property is no big deal when you consider that others do worse" or "People who get mistreated usually do things that deserve it" ($\alpha = 0.87$). The cumulative average of all items was calculated to yield a composite score, with higher scores indicating higher levels of moral disengagement.

Analysis

This study replicated our analysis in Study 1. We first calculated descriptive statistics and conducted bivariate correlations in SPSS, using 1000 bootstrap samples to estimate 95% bias-corrected confidence intervals and applying the Holm-Bonferroni correction at the $p < 0.05$ level to address multiple comparisons. We then employed SEM using the *lavaan* in *R* (Rosseel, 2012) to examine the mediating effects of empathy on the relationship between interpersonal callousness and aggression, assessing both direct and indirect effects.

In addition, we extended our analysis via a moderated mediation approach to test our hypothesis that moral disengagement influences the links between empathy and aggression. For this analysis, we used PROCESS macro in *R*, employing a customised script obtained from: https://www.processmacro.org (Hayes, 2012). This script offers models for different types of moderated mediation tests. For our investigation, we selected model 14 to assess the interaction between the mediator and the moderator in predicting the outcome. Interaction significance was assessed through an interaction effect of $\alpha < 0.05$, and a $BCCI_{95\%}$ that did not include zero. For significant effects, we report the conditional indirect effects of the mediator using the Johnson-Neyman method. This approach provides a range of values of the moderator at which the slope of the predictor goes from non-significant to significant at $p < 0.05$. Age, gender, mental diagnosis, and occupation were entered as covariates in these assessments.

Results

Sample characteristics

As in Study 1, participants' levels of interpersonal callousness, empathy, aggression, and moral disengagement fell within the normative range (see Table 2.2). Normality distributions for each variable are reported in Supplement 2.2.

Correlations

Interpersonal callousness positively correlated with all measures of aggression, ranging from moderate to high – r values between 0.45 and 0.65. We also found positive correlations between interpersonal callousness and empathy deficits overall, with the most robust correlations found for affective empathy measures – r values between 0.49 and 0.65 –, and weaker correlations with perspective-taking, $r = 0.30$, $BCCI_{95\%}$ [0.19, 0.40], $p < .001$, and emotion understanding, $r = 0.14$, $BCCI_{95\%}$ [0.03, 0.24], $p = .017$ – although the latter was not significant after applying the Holm-Bonferroni correction.

Correlations between empathy and aggression where only significant with affective facets of empathy, with values indicating that lower affective empathy correlated with higher aggression – r values between 0.26 and 0.49. Additionally, moral disengagement was positively correlated with interpersonal callousness, aggression, and lower empathy – r values between 0.23 and 0.53. In contrast, its correlation with emotional understanding did not remain significant after correcting for multiple comparisons, $r = 0.11$, $BCCI_{95\%}$ [0.01, 0.23], $p = .044$. Correlation coefficients with their corresponding confidence intervals are reported in Table 2.2.

Table 2.2. | *Mean, standard deviation, and bivariate correlations for main variables in Study 2*

Variables	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
1.F1	2.03	.64	–									
2.F2	1.70	.50	**.65***	–								
3.React	.50	.33	**.45***	**.42***	–							
4.Proact	.06	.16	**.46***	**.45***	**.49***	–						
5.Phys	2.33	.84	**.63***	**.54***	**.49***	**.36**	–					
6.PT	1.28	.69	**.30***	.19	.21	.13	.20	–				
7.COG	2.51	.67	.14	.03	.05	.03	.05	**.39***	–			
8.EC	1.06	.66	**.49***	**.37***	.16	**.29***	**.38***	**.49***	**.33***	–		
9.RES	1.74	.56	**.55***	**.40***	**.26***	**.36***	**.43***	**.51***	**.40***	**.78***	–	
10.DIS	1.45	.48	**.65***	**.48***	**.36***	**.47***	**.49***	**.37***	.16	**.57***	**.57***	–
11. MD	2.06	.49	**.55***	**.47***	**.34***	**.36***	**.52***	**.23***	.11	**.35***	**.36***	**.56***

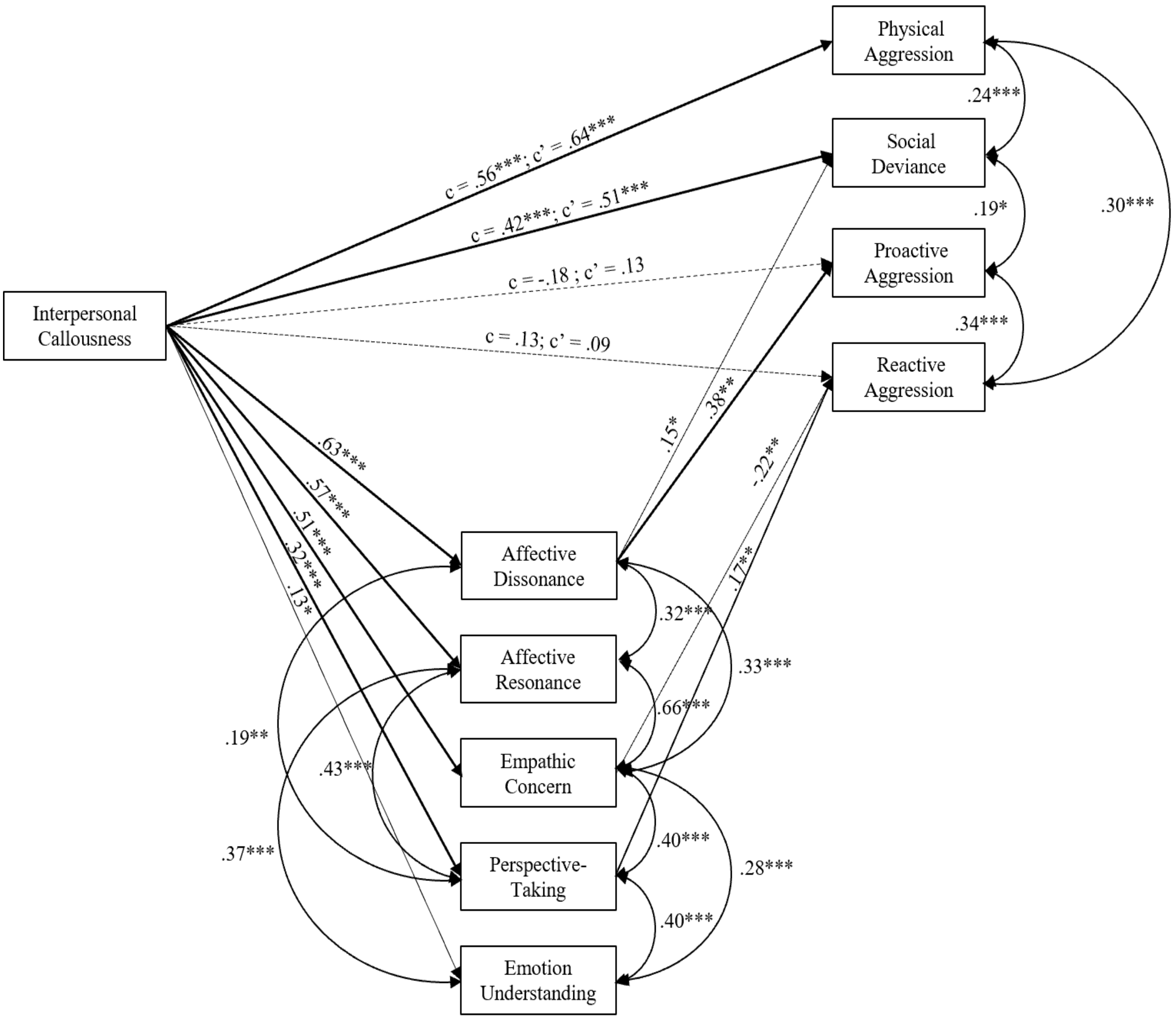
Note. F1 = Interpersonal Callousness; F2 = Social Deviance; React = Reactive Aggression; Proact = Proactive Aggression; Phys = Physical Aggression; PT = Perspective Taking; COG = Emotion Understanding; EC = Empathic Concern; RES = Affective Resonance; DIS = Affective Dissonance; MD = Moral Disengagement. Higher scores in empathy scales denote larger deficits.

Significant values after Holm-Bonferroni correction are presented in **bold**.

Mediation of empathy

Path analyses (illustrated in Figure 2.2) revealed that affective dissonance exhibited a partial indirect effect on link between interpersonal callousness and social deviance, β = 0.09, $BCCI_{95\%}$ [0.01, 0.17], and a full indirect effect on proactive aggression, β = 0.24, $BCCI_{95\%}$ [0.09, 0.36]. Additionally, the effect of interpersonal callousness on reactive aggression was partially mediated by empathic concern, β = -0.11, $BCCI_{95\%}$ [-0.19, -0.03], and perspective-taking, β = 0.05, $BCCI_{95\%}$ [0.02, 0.10]. No other mediating effects were observed and VIF remained below 2.71, indicating no issues of multicollinearity (Richter & Bixler, 2022).

Figure 2.2. | *Empathy mediation on the links between callousness on aggression*

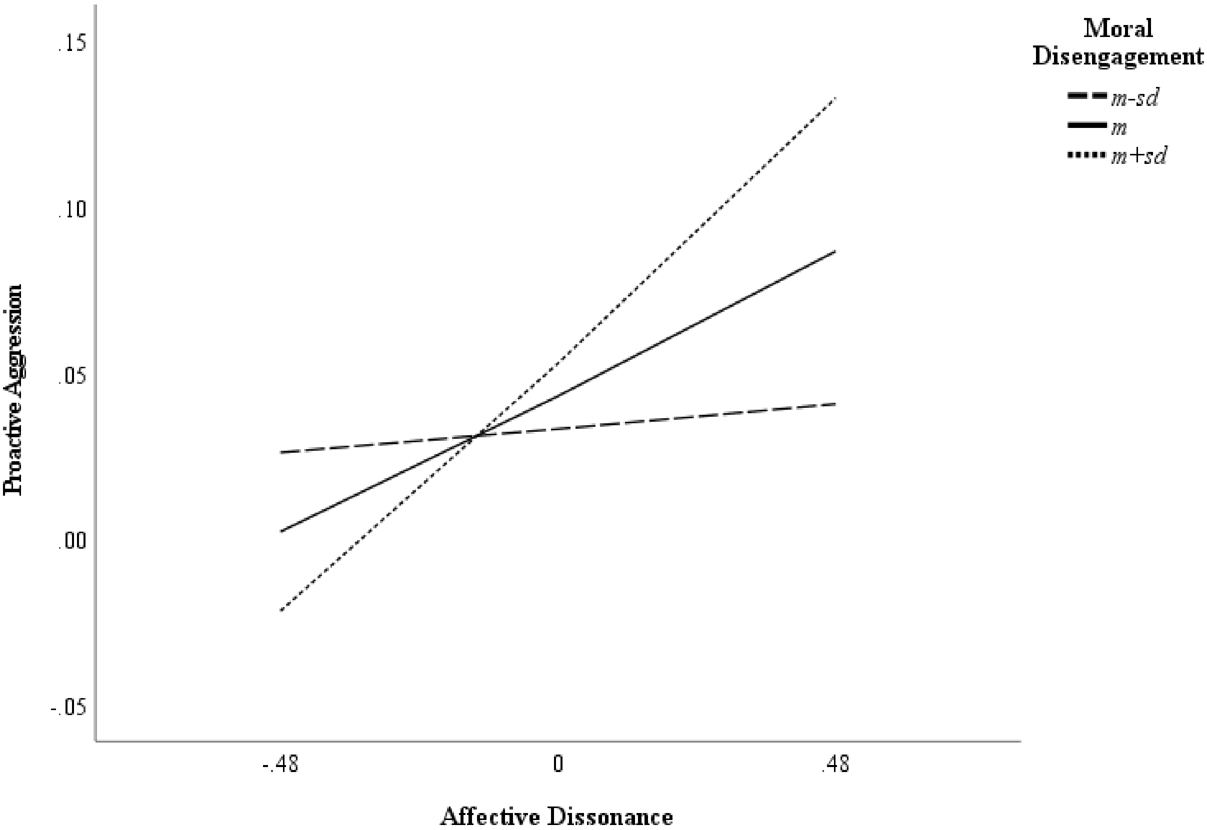


Note. Standardised path coefficients are shown, indicating significant direct effects of interpersonal callousness and empathy, as well as estimated covariances among empathy and aggression variables. Non-significant paths are omitted for simplicity.
Significance estimated at $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Indirect effects of empathy through moral disengagement

Moderated mediation was tested for all significant mediators. Our first model explored empathic concern and perspective-taking as mediators of interpersonal callousness and reactive aggression. The analysis revealed that moral disengagement did not significantly moderate the indirect effects of neither empathic concern, p = .972, nor perspective-taking, p = .590, on reactive aggression. However, a noteworthy finding was the positive effect of mental diagnosis on reactive aggression, β = 0.09, p = .038. Yet, subsequent tests showed that mental diagnosis did not significantly alter the indirect effects of empathic concern, $p = .755$, and perspective-taking, $p = .877$, on reactive aggression either. Next, we examined the influence of moral disengagement on the mediation effect of affective dissonance on proactive aggression. The model revealed that moral disengagement positively influenced the relationship between affective dissonance and proactive aggression, $\beta = 0.15$, $BCCI_{95\%}$ [0.05, 0.27], $p < .001$. As illustrated in Figure 2.2, this moderating effect was mostly significant at higher values of moral disengagement – e.g., at 1.10 standard deviations above the mean, $\beta = 0.25$, $p < .001$, and at average levels of moral disengagement, $\beta = 0.09$, $p < .001$, but lost significance at less than 0.25 standard deviations below the mean, $\beta = 0.05$, $p = .079$, as indicated by Johnson-Neyman estimates (refer to Supplement 2.3 for more details).

Figure 2.3. | *Interaction effects of affective dissonance and moral disengagement on proactive aggression*



Note. The analysis revealed a significantly positive slope in the mean score of proactive aggression in relation to affective dissonance at higher values of moral disengagement (*m+sd*), as well as at mean levels of moral disengagement (*m*), but not for below average moral disengagement (*m-sd*).

Discussion

Study 2 replicated our initial findings, confirming the significance of affective dissonance over other empathy components in proactive aggression, with moral disengagement seemingly amplifying these effects. These results suggest that individuals who are less attuned to others' emotions may be particularly prone to aggression, especially if they also have a tendency to morally disengage.

In contrast, perspective-taking emerged as a partial mediator between interpersonal callousness and reactive aggression, consistent with our hypothesis that perspective-taking negatively influences reactive aggression. Interestingly, we also observed that higher levels of empathic concern were associated with lower levels of reactive aggression. This aligns with previous research indicating that reactive aggression is often associated with difficulties in understanding others' perspectives (Blair, 2013; Chang et al., 2021). The negative relationship between empathic concern and reactive aggression further suggests that empathy may serve as a protective factor against impulsive, reactive aggression, highlighting the nuanced role of empathy in aggression dynamics.

General discussion

Broadly, our findings indicate that deficits in affective empathy, rather than cognitive empathy, play a key role in the relationship between interpersonal callousness and aggression. Specifically, affective dissonance was the only measure of affective empathy to show a full mediating effect on the relationship between interpersonal callousness and proactive aggression. This result aligns with theories that suggest that exploitative and premeditated aggression in psychopaths might relate to the emotional disconnect these individuals experience with others' emotions (Cleckley, 1941; Hare & Neumann, 2008). Notably, given that our observed effects were tested within a community sample, our results specifically suggest that these patterns can be replicated at subclinical levels of psychopathy. This highlights not only the relevance of interpersonal callousness but also affective dissonance in increasing risks for predatory behaviour in psychopathy (see Paulhus, 2014 for a discussion). However, whether affective dissonance is a feature or a consequence of interpersonal callousness warrants further investigation.

Furthermore, in Study 2, we also found that deficits in perspective-taking partially contributed to reactive forms of aggression among participants reporting more callous interpersonal traits. Reactive aggression typically arises in response to perceived threats or provocations, wherein the ability to accurately perceive others' perspectives becomes crucial (Mohr et al., 2007). Therefore, individuals lacking in perspective-taking skills are more likely to misinterpret social situations, which can lead to hostile reactions in self-defence or retaliation (Blair, 2013; Chang et al., 2021). The absence of this effect in Study 1 could be attributed to the smaller sample size, which might have limited the variability in perspective-taking skills and, consequently, the ability to detect these small effects. The larger sample size in Study 2 likely provided a more accurate picture of the relationship between perspective-taking deficits and reactive aggression, which suggests that a more powered sample could be needed to observe the influence of perspective-taking on aggression. We propose that future research should consider these factors and possibly incorporate larger sample sizes to further elucidate the role of cognitive empathy in aggression.

On the other hand, the moderating effects of moral disengagement on each function of aggression suggest differences in the cognitive processes underlying proactive and reactive aggression, with the latter being more closely tied to immediate emotional responses rather than moral considerations. As previously discussed, moral disengagement involves the justification of unethical behaviours through cognitive restructuring, allowing individuals to distance themselves from the moral implications of their actions (Bandura, 1990; Gini et al., 2015). In the context of proactive aggression, moral disengagement may exacerbate the effects of affective empathy deficits by rationalising and justifying harmful behaviours towards others (Gini et al., 2014). In this sense, individuals with more tendencies to morally disengage may perceive acts of aggression as acceptable – or even desirable if they are prone to affective dissonance –, which could ultimately facilitate future engagement in aggression (Hyde et al., 2010; Shulman et al., 2011). It is worth noting, however, that recent longitudinal research suggests that empathy deficits and morally disengaged attitudes may stem from repeated engagement in aggression, rather than the other way around (Falla et al., 2021). This underscores the complexity of the interactions between empathy deficits, moral disengagement, and aggression. Conducting longitudinal research is therefore crucial to better understand the interplay between these constructs.

Despite the valuable insights gained from this research, several limitations must be acknowledged. First, the sample consisted predominantly of non-criminal individuals with higher education backgrounds – namely, university students. This selective and relatively homogenous population may constrain the generalisability of our findings to more diverse or forensic groups, where interpersonal callousness, empathy deficits, and aggressive tendencies may manifest more prominently and differently. Future studies should prioritise the inclusion of high-risk or criminal populations to better evaluate the external validity and broader applicability of these associations.

Furthermore, it is important to recognise that deficits in empathy and moral reasoning represent only part of the broader socio-affective profile associated with interpersonal callousness (Shulman et al., 2011). In fact, in our analyses interpersonal callousness continued to exert unique effects on aggression, even after accounting for empathy deficits and moral disengagement. This suggests that additional factors (e.g., guilt insensitivity, manipulativeness, or emotional detachment) may also play a significant role. Future research would benefit from incorporating a wider range of trait measures related to interpersonal callousness to better elucidate these underlying mechanisms.

Finally, aggressive and antisocial behaviours are not determined solely by individual traits. A substantial body of research has highlighted the predictive power of psychosocial influences such as family dynamics, socioeconomic adversity, and mental health status in shaping future engagement in criminal behaviour (e.g., Booth et al., 2016; Caspi et al., 2002; Dallaire, 2007; Farrington, 2000; Kolivoski & Shook, 2016). Although our samples included some individuals with clinical conditions, they represented a small minority. We propose that future research should adopt more ecologically valid sampling strategies by including participants from diverse socioeconomic backgrounds and with clinically relevant histories. Such approaches would provide a more comprehensive understanding of how individual vulnerabilities interact with contextual risk factors to influence antagonistic behaviour.

Conclusion

Despite the acknowledged limitations, this research shows that even within educated populations with no criminal records, the subclinical expression of interpersonal callousness involves affective and moral deficits that are key for understanding its link with aggression (Brugman et al., 2017). These patterns align with the notion that low empathy might facilitate the expression of cruelty and criminal behaviour in individuals exhibiting psychopathic traits but does not inevitably lead to such outcomes, hence confirming that psychopathy does indeed exist within a spectrum. Consequently, it is important to study the expression of psychopathic traits like interpersonal callousness across diverse sample types. We posit that understanding this spectrum can inform the development of tailored interventions to target specific traits and behaviours before they escalate into criminal actions. Specifically, our findings underscore the importance of addressing emotional incongruence (i.e., affective dissonance) as a potential means to prevent or mitigate aggressive outcomes among individuals exhibiting interpersonal callousness.

Chapter III.

When Empathy Leads to Aggression: Exploring Biases in Moral Judgement and Punishment

Abstract

When witnessing aggression towards others, we tend to sympathise with the victim and condemn the aggressor. This natural tendency to favour the victim can influence how we perceive and interpret such incidents, potentially leading to biases in decisions about punishment. However, the mechanisms underlying these biases remain poorly understood. In this chapter, I sought to investigate whether people determine their decisions to condemn aggressors based on their predisposition to sympathise with the victim, exploring how negative sentiments towards the aggressor may influence these decisions. The study additionally explores these effects through the expression of callous-unemotional traits, hypothesising that moral judgements and decisions to punish may differ among individuals who are less emotionally responsive, as they are less likely to sympathise with victims. Our findings revealed that greater concern for victims intensified punitive attitudes towards aggressors, primarily mediated by participants' negative evaluations of the aggressor. Notably, such empathic inclinations were less prevalent among participants with higher levels of callous-unemotional traits, as reflected by their lower concern for victims and greater inclination towards harsh punishments. These results offer insights into how justice-related attitudes may be shaped and potentially biased by individual differences in emotional responsiveness.

"(P)unishment must be exactly proportional to the "inner viciousness" (inneren Bösartigkeit) of the offender and (…) to achieve this the criminal should (…) be punished by having the exact thing done to [them] that [they have] done to [their] victim – a death for a death, for example."

— quote from Immanuel Kant in 'Metaphysics of Morals'

(as cited by Jeffrie G. Murphy, 2019)

Introduction

The concept of *punishment* unfolds a complex discourse that has engaged scholars from different disciplines. Central to this discourse is the idea that "justice is served" when individuals receive what they are entitled to or deserve based on their actions (Buchanan & Mathieu, 1986; Mikula, 2003; Murphy, 2019; von Grundherr et al., 2021). According to Kant's principles of justice, the crime must be judged proportionately to the gravity of the transgression, recognising the moral agency of wrongdoers, as well as the importance of avoiding extremes of harshness and/or leniency (see von Grundherr et al., 2021 for a discussion). However, as illustrated by the case of Ted Bundy and many other cold-blooded murderers, brutal aggressions towards innocent victims often elicit a shared sense of moral outrage that demands for extreme punitive measures (Carlsmith & Darley, 2008). These responses are often rooted in the belief that punishments should mirror the severity of the suffering inflicted on victims (Bastian et al., 2011; Carlsmith et al., 2002), suggesting a bias in the extent to which individuals empathise with victims over perpetrators.

In this chapter, we investigate how such biases might be determined by the perceived intentionality of the transgression, while also exploring how judgements of the transgressor might further influence punitive attitudes. The goal is to gain insights into how individual's predispositions to selectively apply empathy towards victims over perpetrators might shape their reactions to perceived injustice.

Empathic biases in moral judgement

Empathy has been broadly defined by its affective and cognitive facets, comprising the capacity to share and understand the subjective experience of others in reference to oneself (Decety, 2010). However, empathy can also refer to the capacity to feel concern about the welfare of others, which plays a key role in guiding moral behaviour (Decety, Michalska, et al., 2012;

Decety & Jackson, 2004). Theoretically, it has been proposed that empathic concern prompts individuals to engage in charitable and altruistic behaviours (e.g., Batson et al., 1983; Underwood & Moore, 1982), as they are more prone to consider the well-being of others. Empathic concern has also been proposed to guide individuals in their moral decisions, helping them internalise right from wrong by considering the perspectives, feelings, and needs of others (Kohlberg, 1981; Piaget, 1932). Nevertheless, under specific circumstances, directing empathy towards some individuals may paradoxically diminish our empathy towards others (Simas et al., 2020). For example, people tend to show stronger empathy and helping intentions for ingroup members, which has been correlated with less empathetic responses towards those perceived as outgroup members (e.g., Cikara & Fiske, 2013; Tarrant et al., 2009, 2012; Vanman, 2016). Notably, this bias also seems to contribute to prejudice and discrimination against outgroups (Lalonde & Silverman, 1994; Tyler & Blader, 2003; Vanman, 2016).

In the context of aggression, empathic biases can look like taking the victim's side over the perpetrator's. By adopting the victim's perspective, individuals might engage in self-referential cognitive processes that amplify their sensitivity to the victim's suffering, thereby eliciting a sense of injustice on the victim's behalf (Ames et al., 2008; Decety & Yoder, 2016; Ruby & Decety, 2004). These reactions may in turn foster negative perceptions of wrongdoers, thereby influencing decisions to punish them (e.g., Lin et al., 2024). As put in the words of Michelle Brown:

> "(b)ecause empathy often implies choosing a side, a favoring of one who is more closely like one's self or whose feelings seem more urgent and immediate than others, it may culminate in a narrow perspective that is partisan and inconsistent with the equality of law" – (Brown, 2012, p. 385)

A recent study by Lu and McKeown (2018) provides some evidence for this. Using a Dictator Game task, the authors found that participants with higher levels of empathic concern were more predisposed to punish unfair outgroup members than they were to compensate ingroup members. Similarly, in an earlier study, participants' expressions of empathic anger on behalf of a victim of aggression significantly predicted their intentions to punish the perpetrator (Vitaglione & Barnett, 2003). These findings provide important insights as to how empathy, in some contexts, might indeed lead to aggression (Batson, 1997; see also Neuberg et al., 1997)

It is important to note, however, that the application of punishment as penalty for an immoral action can be seen as a social response seeking to maintain societal harmony, and thus can be described as *aggression with prosocial motives* (Levy, 2022). At the same time, these prosocial motivations differ across individuals (Shichman & Weiss, 2022), meaning that the underlying nature as to why someone decides to punish a transgression might be subject to their inherent biases and perceptions.

The psychology of punishment

Research in psychology shows that people's evaluation of a transgression is often shaped by the extent to which blame is ascribed to the transgressor (e.g., Feather, 1994; Feather & Dawson, 1998; Weiner, 1995). As such, the same act may be judged differently depending on the perceived intentions of the perpetrator, as demonstrated in studies using *moral sensitivity* tasks (e.g., Baez et al., 2014, 2017; Decety et al., 2012; Santamaría-García et al., 2017; Young & Saxe, 2008). In these tasks, participants are typically presented with scenarios where a transgression occurs either accidentally or intentionally. Findings consistently reveal that participants are more likely to morally condemn intentional transgressions compared to accidental ones (Baez et al., 2014; Bastian et al., 2011). Conversely, judgements of aggression may be more lenient when aggressors are perceived as not responsible for their actions or when the aggression is justified for a "greater good" (Mikula, 2003).

This aligns with the *Attribution of Blame* model described by Shaver and Shaver (1985), which posits that assessments of injustice hinge on the degree of responsibility attributed to the perpetrator for violating someone else's entitlements without adequate justification. For example, while intentional harm to innocent victims is condemned and penalised (Mikula, 2003), aggressions towards violent offenders – often enforced through extreme forms of punishment such as capital penalties or castration – tend to be more justified and endorsed by the public (e.g., Bastian et al., 2013; Viki et al., 2012). This justification hinges on the idea that justice is implemented when punishments are proportional to the harm inflicted, with harsher punishments often perceived as effective strategies to both prevent the offender from reoffending as well as to discourage others from committing similar crimes. However, it should not be ignored that punishment may also be an emotional response to perceived injustice, which may drive individuals to advocate for (and potentially engage in) more severe measures for retaliation.

In fact, negative emotions arising from perceived injustice towards victims can lead individuals to dehumanise perpetrators (Carlsmith et al., 2002; Darley & Pittman, 2003). This process may occur when perpetrators are seen as violating moral principles or social norms, justifying perceptions of them as less human and undeserving of basic rights (Haslam et al., 2005). Such judgements gain special relevance in legal contexts, where the severity of punishment for the wrongdoer is often based on the perceived intentionality of the committed crime (Bastian et al., 2011). This can lead to harsher penalties and reduced advocacy for the reform and reintegration of violent offenders (Viki et al., 2012), especially when they are perceived as irredeemable and *evil* (Bastian et al., 2013; Harris & Rice, 2006; Osofsky et al., 2005).

Individual differences in perceptions of injustice

Nevertheless, not all people will exhibit the same drives for retribution. While some react strongly to perceived injustices, others may minimise or entirely disregard the negative consequences of moral violations. Research suggests that people's responses to perceived injustice may vary depending on their emotional involvement when making moral judgments

(Brown, 2012; Decety & Cowell, 2015; Mikula, 2003). Therefore, an individual's capacity (or willingness) to emotionally engage with others' experiences is key when examining their drive to retaliate on the behalf of others. Compelling examples can be found in the context of psychopathy.

People with psychopathic tendencies are typically less sensitive to perceived injustice (Decety & Yoder, 2016). For example, individuals with psychopathy often engage in calculated and goal-oriented aggression (e.g., using manipulation or deceit to achieve personal gain) even when fully aware that their actions cause harm to others (Gini et al., 2015). Similarly, studies have shown that psychopaths can correctly identify morally wrong actions but fail to show guilt or remorse for their wrongdoings (Koenigs et al., 2012). To put it simply, "(p)sychopaths know what is right or wrong, but simply don't care" (Cima et al, 2010, p. 66). Such disregard for others has been attributed to the expression of callous-unemotional traits, which broadly reflects a lack of empathy (Lockwood et al., 2013). Individuals exhibiting these traits – even at subclinical levels – express less concern for others' suffering (e.g., Blair, 2005; Cheng et al., 2012; Lockwood et al., 2013; Marsh et al., 2008, 2013), are more likely dismiss the consequences of moral violations (Barchia & Bussey, 2011; Caravita et al., 2012; Gini et al., 2015; Koenigs, 2012), and have higher risks for aggression (Camara et al., 2025).

That is not to say, however, that all individuals exhibiting callous-unemotional traits are inherently immoral and aggressive (Campos et al., 2022). In fact, research on successful psychopathy indicates that some psychopathic traits may be advantageous in certain circumstances (Hall & Benning, 2006). For example, individuals with callous-unemotional traits may be less influenced by emotional appeals when making moral judgements (Fragkaki et al., 2016; Vasconcelos et al., 2021). This reduced susceptibility to emotional distress, in turn, can be beneficial in professions that require emotional detachment, such as law enforcement (Dutton, 2012). From this perspective, one could argue that callous-unemotional traits might reduce affective biases during punishment decision-making, although no previous study has investigated this.

Project rationale

In the present research, we aimed to build upon the current understanding of individual biases in punitive decision-making by examining how increased concern for victims influence punitive attitudes towards perpetrators, exploring how the presence of callous-unemotional traits might modulate this association. We conducted two experiments to address these questions. In the first experiment, participants evaluated scenarios depicting interpersonal harm, both intentional and accidental. They were asked to express their concerns for both the victim and the perpetrator assess the perceived "meanness" of the perpetrator and determine appropriate levels of punishment. The second experiment replicated this task and additionally measured participants' reported levels of callous-unemotional traits to explore how responses in the task could be influenced by the expression of these traits.

We hypothesised that participants expressing more concern for victims would adopt harsher punitive attitudes towards perpetrators, especially when the transgression was intentional (Baez et al., 2014; Bastian et al., 2011). We additionally proposed that these effects would be more pronounced among participants expressing lower concern for perpetrators and if the perpetrator was perceived as "mean" (Bastian et al., 2013; Decety & Cowell, 2015; Viki et al., 2012). Lastly, we anticipated that the expression of callous-unemotional traits would in turn diminish the link between concern for victims and punishment severity, given that individuals with these traits are less likely to feel concern for victims (Decety & Yoder, 2016).

Data and scripts for each experiment have been made available and can be accessed at: https://osf.io/bfmnt/?view_only=8108167bd8a540e98282b78fe2d407f1.

Study 3

Method

Participants and procedure

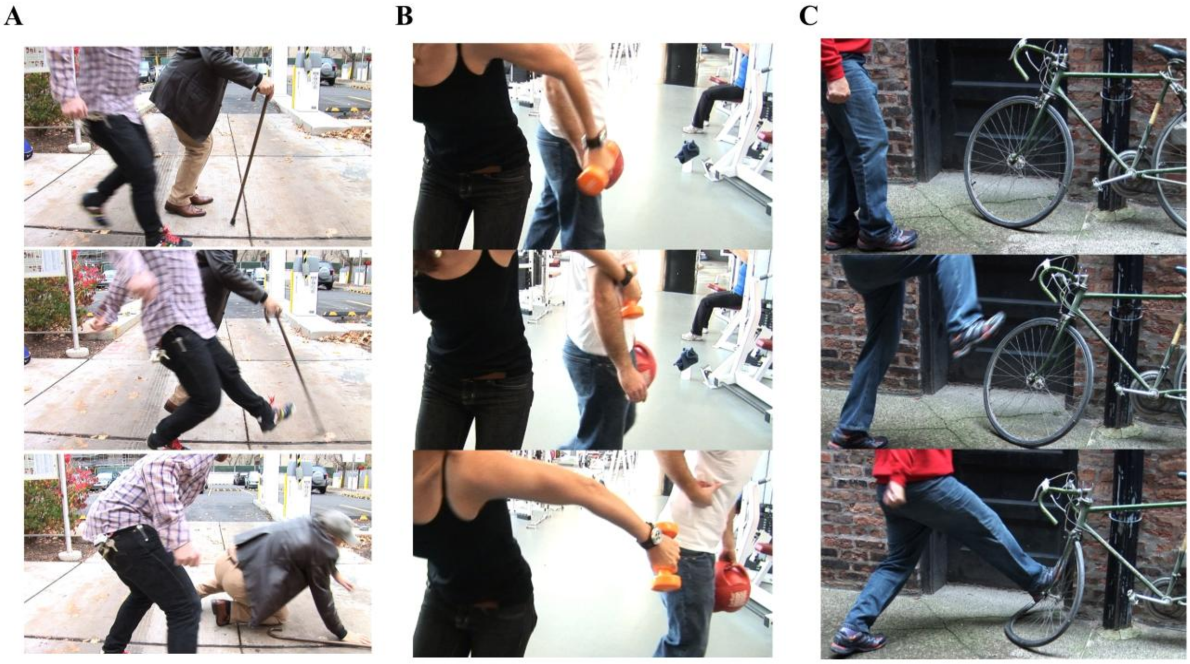
A power analysis indicated that 111 participants were required to detect a small-to-medium effect size ($f = 0.175$) in a 3x2-within-subjects design with 95% power. Of the 126 participants initially recruited, 15 did not complete the task, and one participant was excluded for failing attention checks. This resulted in a final sample of 110 participants. The sample primarily consisted of young adults aged 18 to 33, $M = 21.69$, $SD = 2.77$, including 76 women, 34 men, and one participant who identified as non-binary/third gender. The majority identified as White ($N = 68$), while 23 participants identified as Asian/Asian British, 8 as Black, and 11 belonged to multiple or other ethnic groups. Most participants completed the experiment within 10 to 25 minutes.

This experiment was part of a larger online study administered through Qualtrics, approved by the University of Essex Ethics Committee.

Experimental design and stimuli

The experiment used a *Moral Sensitivity Task* adapted from Decety *et al* (2012). Participants were presented with three different types of transgressions, where: a) one person deliberately injures another; b) one person unintentionally injures another; and c) intentional harm was directed at an object rather than a person (see Figure 3.1 for examples). After each scenario, participants selected 'Yes' or 'No' to indicate whether the transgression was deliberate or unintentional ("was the action done on purpose?"). Subsequently, they used visual analogue scales ranging from 0 to 100 to rate their concern for both the victim ("how sorry do you feel for the injured person/damaged object?") and the perpetrator ("how sorry do you feel for the perpetrator?"), as well as their judgement of the perpetrator's meanness ("how mean was the perpetrator?") and appropriate punishment ("how harshly would you punish the perpetrator?").

Figure 3.1. | *Examples of animated images in moral sensitivity task*



Note. Panel a illustrates a scenario depicting intentional harm to another person, representing the intentional condition. Panel b depicts a scenario portraying accidental harm to another person, representing the accidental condition. Lastly, panel c showcases a scenario demonstrating intentional harm to an object, serving as the control condition.

The task comprised 2 initial practice trials to familiarise participants with the instructions, followed by 6 experimental trials (2 per condition, randomised across participants). The experimental trials featured a different set of animated pictures (10 per condition), randomly presenting 2 pictures from each condition to participants. The stimuli were created using 3 digital colour images, shown sequentially at durations of 500, 200 and 1000 ms to imply motion. Participants were blinded to the protagonists' faces to avoid facial bias, but the victim's emotional reaction and the perpetrator's intent remained inferable through body language. To reduce gender bias, the conditions included interactions across male-to-male, female-to-female, male-to-female, and female-to-male pairings.

Analysis

We examined participants' responses across conditions in *SPSS* (version 29.0). First, we assessed the success of our manipulation by evaluating participants' accuracy in distinguishing between intentional and accidental conditions, with successful manipulation defined as correctly identifying intentionality in more than 50% of the trials. Next, we tested our hypothesis that participants would express greater concern for victims than for perpetrators by conducting a 2 (victim *vs* perpetrator) x 3 (control *vs* intentional *vs* accidental) repeated-measures ANOVA. Significant interactions were further explored using simple main effects analyses. Additionally, we examined differences in perceived perpetrator meanness and punishment ratings across conditions using one-way repeated-measures ANOVAs. For all analyses, follow-up pairwise comparisons were conducted using a Holm-Bonferroni correction, with an alpha level set at 0.05 and partial eta-squared (η_p^2) reported as the effect size. We also reported 95% confidence intervals ($CI_{95\%}$) where applicable.

To test our hypotheses about how perceived perpetrator meanness and concern for perpetrators influence the relationship between concern for victims and punishment, we used SEM with the *lavaan* package in *R* (Rosseel, 2012). We initially introduced the variables measuring concern for perpetrators and perceived meanness as mediators into separate models to investigate their mediation effects individually. Next, we examined the relative effects of these variables into a combined model. Significant interactions were probed using 1000 bootstrap samples from the original dataset and computed $BCCI_{95\%}$. Interactions were considered significant if the confidence intervals did not include zero (MacKinnon et al., 2007). Model fit was assessed using conventional indices, following recommended guidelines (L. Hu & Bentler, 1998; Schreiber et al., 2006). These included the chi-square statistic (χ^2), Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), Comparative Fit Index (CFI), and Tucker-Lewis Index (TLI). Acceptable fit was defined as RMSEA and

SRMR values ≤ .08, and CFI and TLI values ≥ .90. A non-significant χ^2 ($p > .05$) was also considered indicative of good model fit, though interpreted with caution due to its sensitivity to sample size. Additionally, we report relative fit indices – including the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and the Sample-Size Adjusted BIC (SABIC), with lower values indicating better fit – for comparisons.

Results

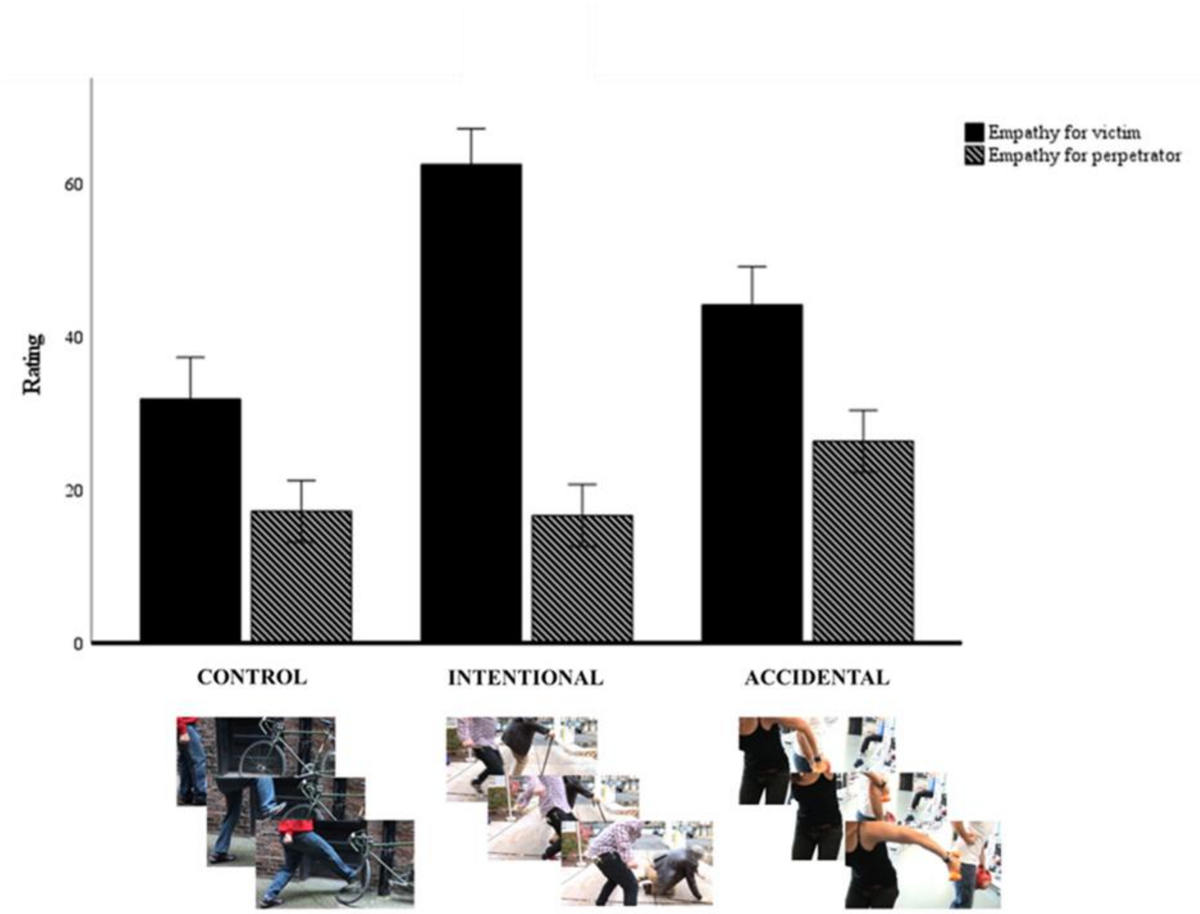
Perceived intentionality

Participants accurately identified perpetrators' intentions in ~90-96% of cases involving intentional harm and in ~75% of accidental harm scenarios, indicating that the experimental manipulation was successful.

Concern for victims versus perpetrators

There was a significant interaction between expressed concern for victims and perpetrators across conditions, $F_{(1,109)} = 54.92$, $p < .001$, $\eta_p^2 = 0.33$. As anticipated, participants expressed significantly more concern for victims in the intentional condition, $M = 62.44$, $SE = 2.35$, $CI_{95\%}$ [57.78, 67.10], compared to both the accidental, $M = 44.13$, $SE = 2.51$, $CI_{95\%}$ [39.15, 49.11], and the control, $M = 31.82$, $SE = 2.76$, $CI_{95\%}$ [26.35, 37.29], conditions. In contrast, participants expressed significantly less concern for perpetrators in both the intentional, $M = 16.66$, $SE = 2.04$, $CI_{95\%}$ [12.61, 20.70], and control, $M = 17.21$, $SE = 2.02$, $CI_{95\%}$ [13.20, 21.22], conditions compared to the accidental condition, $M = 26.34$, $SE = 2.04$, $CI_{95\%}$ [22.30, 30.38], whereas concern for perpetrators did not differ between intentional and control conditions, $p = .793$. Additional main effect analyses showed that expressed concern for victims was significantly higher than concern for perpetrators in each condition (see Figure 3.2 for an illustration of these contrasts). All differences were significant at $p < .001$ (uncorrected), hence remaining significant at $p < .05$ after Holm-Bonferroni correction.

Figure 3.2. | *Empathic concern for victim vs perpetrator in Study 3*



Note. Error bars represent 95% confidence intervals. In all conditions, concern for victims was significantly greater than concern for perpetrators (control: $M = 31.82$, $SE = 2.76$, $CI_{95\%}$ [26.35, 37.29] *vs* $M = 17.21$, $SE = 2.01$, $CI_{95\%}$ [13.20, 21.22]; intentional: $M = 62.44$, $SE = 2.35$, $CI_{95\%}$ [57.78, 67.10] *vs* $M = 16.66$, $SE = 2.04$, $CI_{95\%}$ [12.61, 20.70]; accidental: $M = 44.13$, $SE = 2.51$, $CI_{95\%}$ [39.15, 49.11] *vs* $M = 26.34$, $SE = 2.04$, $CI_{95\%}$ [22.30, 30.38]. All comparisons significant at $^{***}p < .001$.

Judgements of perpetrator meanness and punishment

There was a significant difference in participants' ratings of both perceived perpetrator meanness, $F_{(1,109)} = 140.40$, $p < .001$, $\eta_p^2 = 0.56$, and punishment severity, $F_{(1,109)} = 90.86$, $p < .001$, $\eta_p^2 = 0.45$, across conditions. Participants rated the perpetrator as significantly meaner when the aggression was intentionally directed at another person, $M = 68.96$, $SE = 2.03$, $CI_{95\%}$ [64.94, 72.99], or at an object, $M = 42.25$, SE = 2.54, $CI_{95\%}$ [37.23, 47.28], compared to accidental harm to another person, $M = 26.43$, $SE = 2.02$, $CI_{95\%}$ [22.43, 30.44]. Similarly, participants supported significantly harsher punishments for intentional harm to both a person, $M = 54.53$, $SE = 2.27$, $CI_{95\%}$ [50.03, 59.04], and an object, $M = 37.06$, $SE = 2.52$, $CI_{95\%}$ [32.08, 42.05], than for unintentional harm, $M = 21.38$, SE =1.97, $CI_{95\%}$ [17.46, 25.29]. Post-hoc pairwise comparisons revealed that these differences were significant at $p < .05$ after Holm-Bonferroni correction.

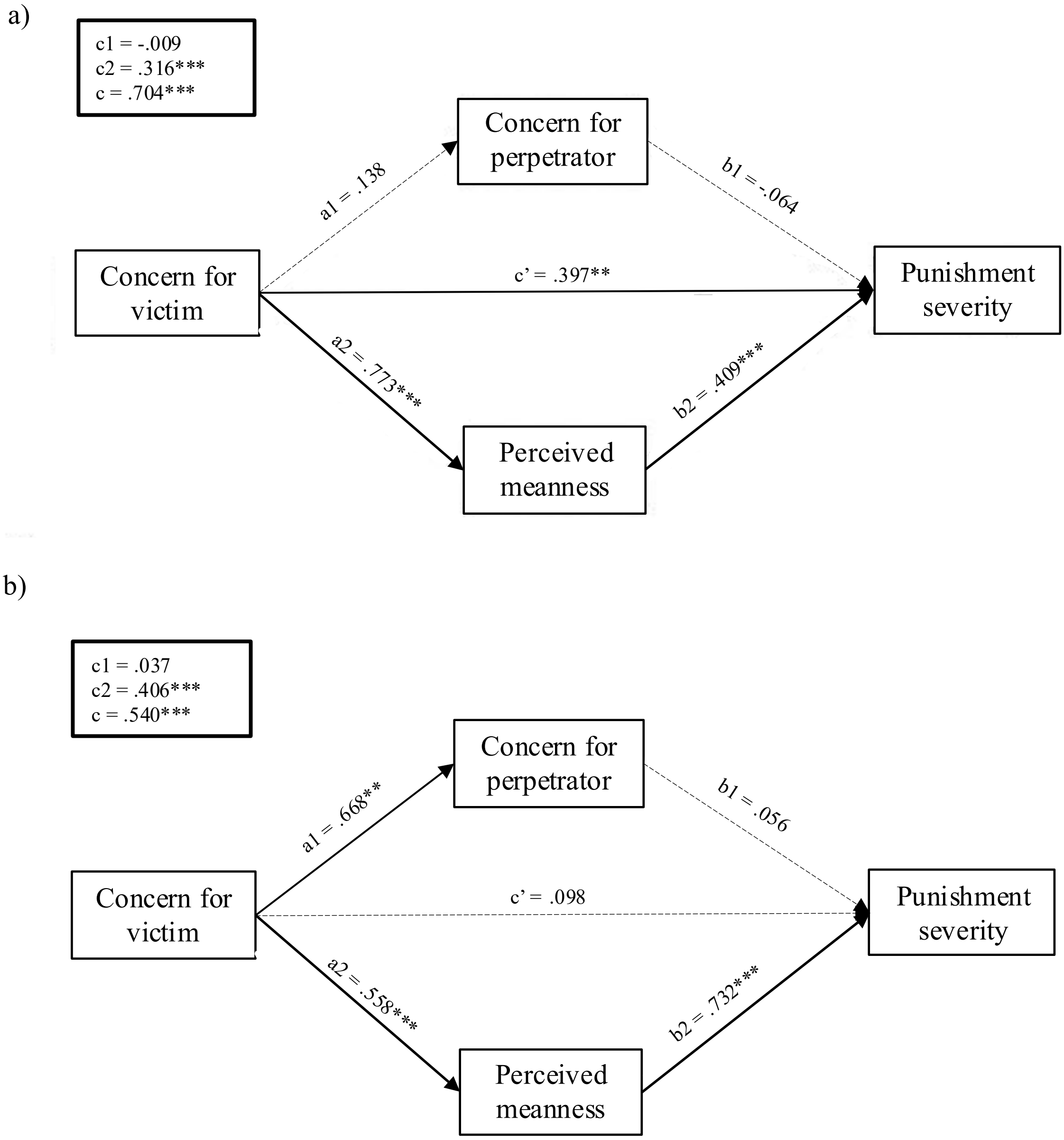
Table 3.1. | *Direct and indirect effect estimates of mediation models in Study 3*

	Intentional			**Accidental**		
	β	***SE***	***BCCI$_{95\%}$***	***β***	***SE***	***BCCI$_{95\%}$***
Model 1						
Total effect	.704	.052	**.593, .799**	.544	.065	**.416, .667**
Direct effect	.713	.053	**.605, .808**	.653	.128	**.410, .894**
Indirect effect	-.010	.013	-.040, .017	-.109	.094	-.294, .072
Model 2						
Total effect	.704	.055	**.591, .804**	.544	.066	**.409, .667**
Direct effect	.386	.118	**.158, .640**	.143	.082	-.020, .300
Indirect effect	.317	.094	**.135, .508**	.401	.064	**.277, .536**
Model 3						
Total effect	.704	.056	**.595, .810**	.540	.066	**.411, .672**
Direct effect	.397	.121	**.169, .637**	.097	.141	-.152, .395
Indirect effect						
via concern for perpetrator	-.009	.014	-.035, .024	.037	.075	-.123, .176
via perceived meanness	.316	.095	**.132, .505**	.406	.067	**.278, .535**

Note. Model 1 = Estimates through individual mediation of concern for perpetrator; Model 2 = Estimates through individual mediation of perceived perpetrator meanness; Model 3 = Estimates through the combined mediation of concern for perpetrator and perceived perpetrator meanness. Reported direct effects are estimated after accounting for indirect effects.

Significant estimates (BCCI$_{95\%}$ not including 0) are in **bold**.

Figure 3.3. | *Path estimates in a) intentional and b) accidental conditions in Study 3*



Note. Mediation was assessed by examining the direct (c') and indirect effects of interpersonal callousness on aggression through concern for the perpetrator (c1) and perceived perpetrator meanness (c2). The total effect (c) was the sum of the indirect effects through the mediators and the direct effect of the predictor. Significance estimated at $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Direct and indirect effects of concern for victims on punishment

Table 3.1 provides direct and indirect effect estimates for each mediation model. Examination of individual mediation models revealed that concern for perpetrators did not significantly influence the relationship between concern for victims and punishment in neither condition. In contrast, perceived perpetrator meanness partially mediated the link between concern for victims and punishment in intentional scenarios, $\beta = 0.32$, $SE = 0.09$, $BCCI_{95\%}$ [0.13, 0.51], and fully mediated this relationship in accidental scenarios, $\beta = 0.40$, $SE = 0.06$, $BCCI_{95\%}$ [0.28, 0.54]. These effects remained when assessing concern for perpetrators and perceived meanness as mediators into combined models for each condition (see Figure 3.3). Fit indices indicated that both models were a good fit to our data, with relative fit indices suggesting better fit for the *Accidental* model (see Supplement 3.1).

Discussion

Our findings align with previous research, reinforcing the idea that people's judgments of perceived transgressions are significantly influenced by their perceptions of the perpetrator's intentions (Decety, Michalska, et al., 2012; Decety & Cowell, 2015; Mikula, 2003). Specifically, when harm is perceived as intentional, individuals tend to adopt harsher punitive attitudes towards the perpetrator (Decety et al., 2012). Furthermore, selectively applying empathy towards victims of aggression increased participants' tendency to impose harsher punishments on the aggressors. Notably, this effect was primarily mediated by participants' negative perceptions of the aggressor, aligning with existing literature that suggests that retaliatory responses to aggression are influenced by negative assessments of the aggressor (Carlsmith et al., 2002). In our study, this effect was particularly strong in scenarios of accidental harm, where the relationship between concern for victims and punishment severity was fully mediated by perceived perpetrator meanness. This finding implies that people's punitive attitudes against aggressors might be more influenced by their negative perceptions of the aggressor rather than their judgement of the aggression itself or their concern for the victim. In contrast, while participants also expressed lower concern for perpetrators, this did not seem to influence their decisions to punish, contradicting our initial expectations. These effects were further examined in Study 4.

Study 4

Method

Participants and procedure

Even though the sample in Study 3 was large enough to detect the postulated effects, in this study we decided to almost triple our sample size to allow us to detect potentially small moderating effects of callous-unemotional traits. The final sample included 310 participants recruited via Prolific. Participants were given a maximum of 3 hours to complete the experiment to ensure timely completion while minimising fatigue or distraction effects that could compromise data quality. Upon successful completion, each participant received a £7 reimbursement. All participants gave their informed consent and provided their demographic data before proceeding to the experiment. This experiment was approved by the university's ethics committee as part of a larger study.

Participants were young adults aged 18 to 25 (M = 22.75, SD = 1.75), including 161 men, 146 women, and 3 participants who identified as non-binary/third gender. More than half of the sample identified as White (N = 188), with 51 participants identifying as Asian/Asian British, 43 as Black, and 28 reporting multiple/other ethnic groups. After providing their demographic data, participants performed the same experimental task described in Study 3, followed by survey questionnaires. Most participants completed the study in around 27 minutes.

Assessment of callous-unemotional traits

We used the *Inventory of Callous Unemotional Traits* (ICU; Frick et al., 2003) to assess callous-unemotional traits. This questionnaire includes items denoting callousness (e.g., "The feelings of others are unimportant to me"), disregard for the consequences of one's actions on others (e.g., "I do not care who I hurt to get what I want"), and shallow affect (e.g., "I do not show my emotions to others"). Responses are provided on a 4-point Likert scale from 0 (not at all true) to 3 (definitely true), and the summed score was computed for follow-up analyses – excluding items 2 ("What I think is right and wrong is different from what other people think") and 10 ("I do not let my feelings control me") due to prior research indicating higher internal consistency after removing these items, along with their low item-total correlations with the total ICU scale (see Kimonis et al., 2008). After excluding these items, the remaining ICU items yielded a reliability score of $\alpha = 0.82$, indicating high internal consistency.

Analysis

Initial analyses replicated the procedures described in Study 3. First, we examined whether participants correctly differentiated between intentional and accidental conditions as a manipulation check. We then analysed participants' responses across conditions using repeated-measures ANOVAs in SPSS and further explored condition-specific effects via main effects analyses, reporting $CI_{95\%}$ for all estimated means. Effect sizes were assessed using η_p^2, and Holm-Bonferroni corrections were applied to adjust for multiple comparisons.

Furthermore, we examined whether the relationship between concern for victims and punishment was moderated by levels of callous-unemotional traits in both intentional and accidental conditions. To test these effects, we conducted multiple linear regression analyses that included the main effects of concern for victims and callous-unemotional traits, as well as their interaction. Where the interaction term was significant, we used the interactions package in *R* to perform simple slopes analyses (J. A. Long & Long, 2021) estimating the effect of concern for victims on punishment at low, average, and high levels of callous-unemotional traits. We also used the Johnson-Neyman technique to identify the range of callous-unemotional trait values for which the relationship between concern for victims and punishment was statistically significant.

Lastly, mediation analyses were conducted using SEM with the *R* package *lavaan* (Rosseel, 2012), testing the direct and indirect effects of concern for victims on punishment decisions through concern for perpetrators and perceived meanness. These analyses included callous-unemotional traits as an additional mediator. As in Study 1, mediators were first assessed in individual models and then combined into a single model for both intentional and accidental conditions. Significant interactions were probed using 1000 bootstrap samples and $BCCI_{95\%}$ excluding zero (MacKinnon et al., 2007). Normative fit indices ($\chi 2$, RMSEA, SRMR, CFI, TLI) were used to assess model fit in the combined models, with relative fit indices (AIC, BIC, SABIC) additionally reported for model comparisons (Ullman & Bentler, 2012).

Results

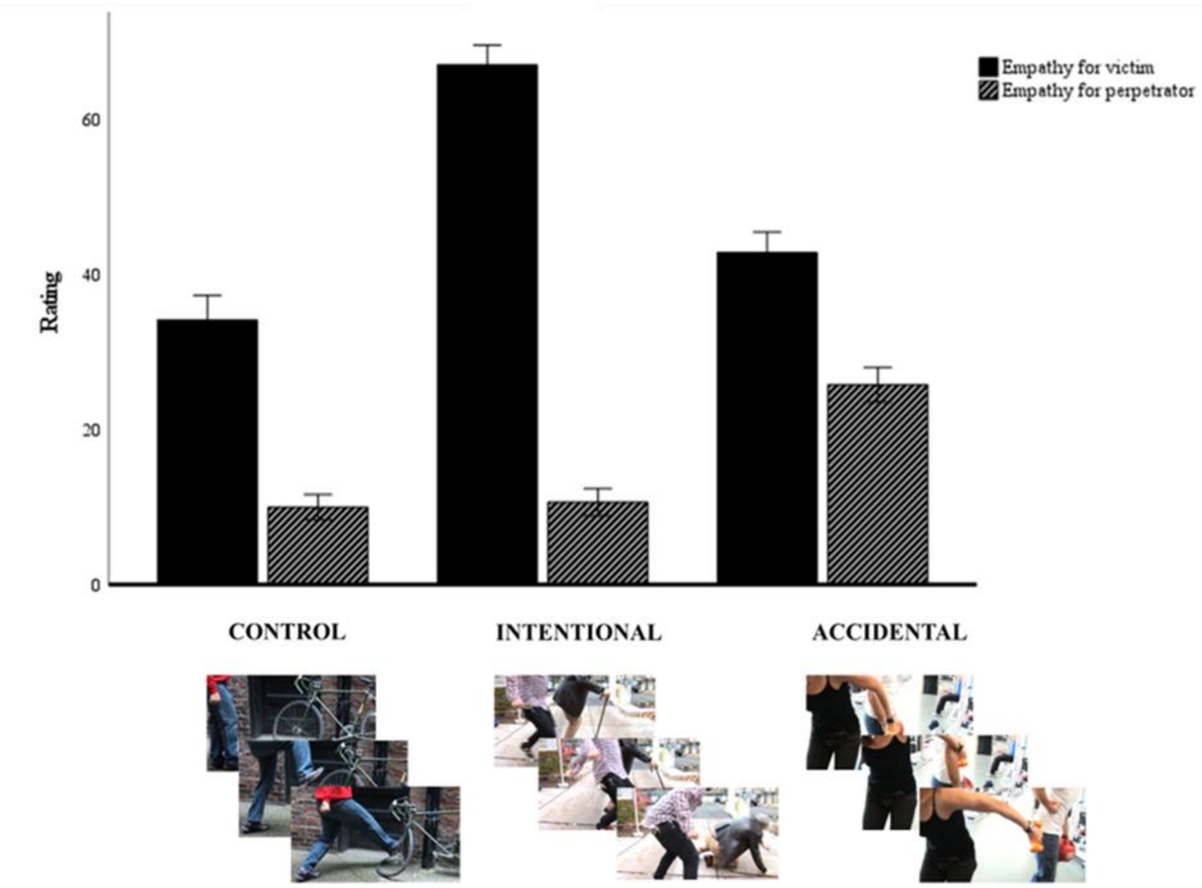
Perceived intentionality

Participants accurately identified intentional harm with a high success rate of 93.9%. In contrast, their accuracy in identifying unintentional harm was notably lower, at 55.5%. This discrepancy highlights the challenge participants faced in discerning the perpetrator's intentions in accidental scenarios. However, this accuracy still exceeded the pre-established 50% threshold, indicating that (for the most part) participants were still able to identify intentionality in these conditions.

Concern for victims versus perpetrators

Consistent with our previous findings, there was a significant difference in participants' expressed concern for victims and perpetrators across conditions, $F_{(1,309)} = 246.46$, $p < .001$, $\eta_p^2 = 0.44$. However, simple main effects analyses indicated that participants expressed greater concern for victims when the transgression was intentional, $M = 66.92$, $SE = 1.28$, $CI_{95\%}$ [64.04, 69.44], than when it was accidental, $M = 42.76$, $SE = 1.35$, $CI_{95\%}$ [40.10, 45.42], or directed at an object, $M = 34.09$, $SE = 1.61$, $CI_{95\%}$ [30.93, 37.25].

Figure 3.4. | *Empathic concern for victim vs perpetrator in Study 4*



Note. Error bars represent 95% confidence intervals. Across all conditions, concern for victims was significantly higher than concern for perpetrators (Control: M = 34.09, SE = 1.61, $CI_{95\%}$ [30.93, 37.25] *vs* M = 10.01, SE = 0.83, $CI_{95\%}$ [8.38, 11.65]; Intentional: M = 66.92, SE = 1.28, $CI_{95\%}$ [64.40, 69.44] *vs* M = 10.66, SE = 0.89, $CI_{95\%}$ [8.90, 12.41]; Accidental: M = 42.76, SE = 1.35, $CI_{95\%}$ [40.10, 45.42] *vs* M = 25.76, SE = 1.14, $CI_{95\%}$ [23.52, 28.00]). All comparisons significant at $^{***}p < .001$.

In turn, concern for perpetrators was higher in the accidental harm condition, M = 24.76, SE = 1.14, $CI_{95\%}$ [23.52, 28.00], than in both the intentional, M = 10.66, SE = 0.89, $CI_{95\%}$ [8.90, 12.41], and object harm, M = 10.01, SE = 0.83, $CI_{95\%}$ [8.38, 11.65], conditions. Further contrasts additionally showed that participants expressed significantly more concern for victims than for perpetrators in all conditions (main effects in Figure 3.4). All contrasts were significant at $p < .05$ after applying the Holm-Bonferroni correction, except for the difference between concern for perpetrators in the intentional versus control conditions, $p = .472$.

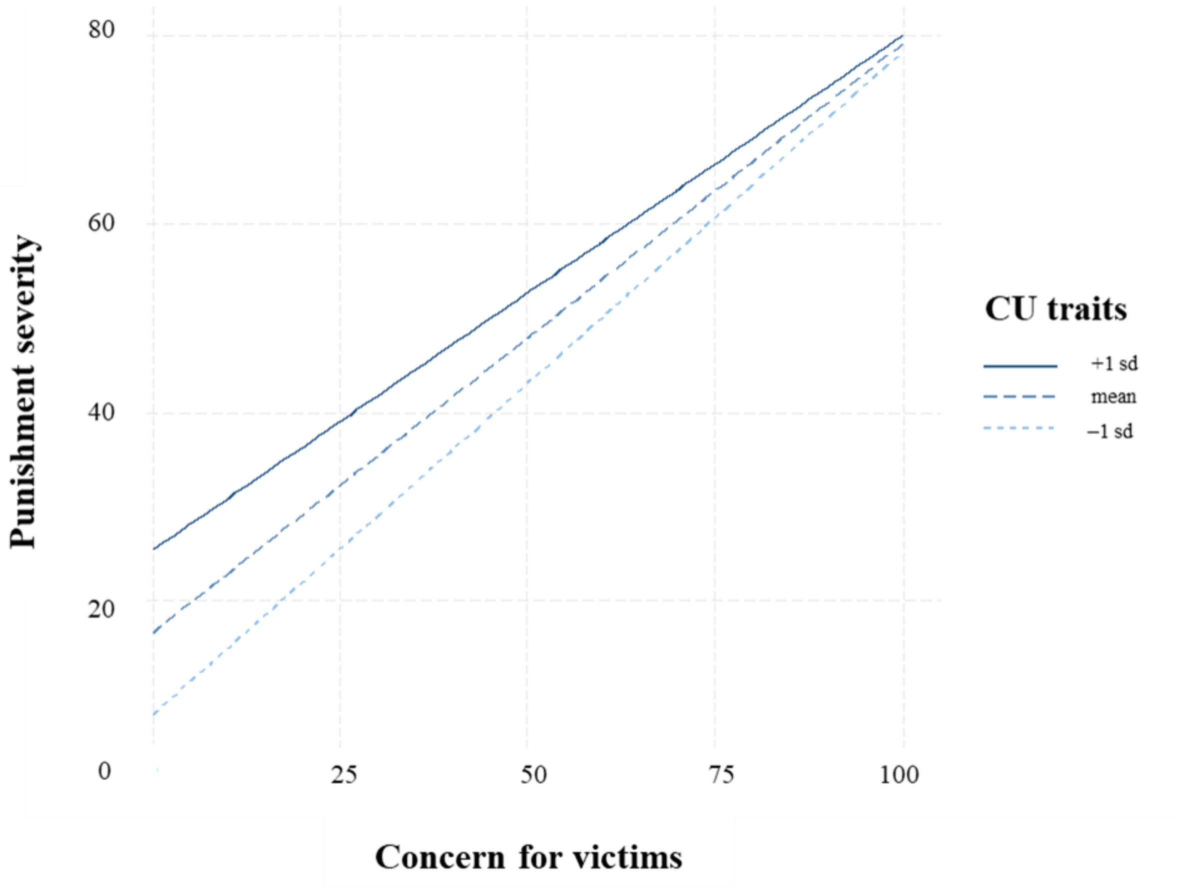
Moral judgements

There was a significant main effect of transgression type on both perceived meanness of the perpetrator, $F_{(1,309)} = 511.47$, $p < .001$, $\eta_p^2 = 0.62$, and punishment severity, $F_{(1,309)} = 371.06$, $p < .001$, $\eta_p^2 = 0.55$. Participants rated the perpetrator as significantly meaner when the harm was intentional, *M* = 73.61, *SE* = 1.05, $CI_{95\%}$ [71.54, 75.67]), compared to harm directed at an object, *M* = 46.76, *SE* = 1.57, $CI_{95\%}$ [43.68, 49.84], or accidental harm to a person, *M* = 21.41, *SE* = 1.13, $CI_{95\%}$ [19.19, 23.63]. Similarly, witnessing intentional harm to another person led to harsher punishment decisions, *M* = 58.88, *SE* = 1.13, $CI_{95\%}$ [56.65, 61.11], than when harm accidentally caused to another person, *M* = 18.34, *SE* = 1.00, $CI_{95\%}$ [16.38, 20.31], or directed at an objected, *M* = 37.64, *SE* = 1.41, $CI_{95\%}$ [34.87, 40.42]. All comparisons were significant at $p < .001$ (uncorrected), with significance retained after Holm-Bonferroni correction.

Effects of callous-unemotional traits

Lastly, we examined callous-unemotional traits' moderation and mediation effects on the links between concern for victims and punishment of perpetrators. As anticipated, concern for victims led to harsher punishment in both intentional, $\beta = 0.80$, *SE* = 0.11, $p < .001$, and accidental, $\beta = 0.27$, *SE* = 0.10, $p = .005$, conditions. However, in the intentional harm condition, the effect of victim concern on punishment severity weakened at higher levels of callous-unemotional traits, $\beta = -0.01$, *SE* = 0.01, $p = .033$. Interestingly, participants with more callous-unemotional traits also supported more severe punishment overall, $\beta = 1.09$, *SE* = 0.32, $p = .001$. Furthermore, simple slopes analyses (Figure 3.5) indicated that the relationship between victim concern and punishment remained significant across different levels of callous-unemotional traits (all *p*s < .001). In contrast, in the accidental harm condition, callous-unemotional traits did not significantly predict punishment severity, $\beta = 0.22$, *SE* = 0.21, $p = .314$, nor did they moderate the effect of concern for victims, $\beta = 0.00$, *SE* = 0.00, $p = .291$.

Figure 3.5. | *Link between punishment severity and victim bias in the intentional condition*



Note. Effects of empathy on punishment severity at higher (+1 SD), mean, and lower (-1 SD) levels of callous-unemotional traits (CU traits).

Mediation analyses revealed further nuances in the role of callous-unemotional traits in relation to concern for perpetrators and perceived perpetrator meanness in punitive decision-making (Table 3.2). When evaluating each mediator into individual models, we found that perceived meanness partially mediated the relationship between concern for victims and punishment severity in both the intentional, $\beta = 0.45$, $SE = 0.05$, $BCCI_{95\%}$ [0.36, 0.57], and accidental conditions, $\beta = 0.35$, $SE = 0.04$, $BCCI_{95\%}$ [0.27, 0.43], indicating a positive effect. In contrast, callous-unemotional traits had a negative indirect effect on this association – although only for the intentional condition, $\beta = -0.04$, $SE = 0.01$, $BCCI_{95\%}$ [-0.08, -0.02], which is consistent with our results in the moderation analyses. Moreover, concern for perpetrators emerged as significant mediator in the accidental condition, β = -0.08, *SE* = 0.04, $BCCI_{95\%}$ [-0.15, -0.01], with data suggesting a partial negative effect.

Table 3.2. | *Direct and indirect effect estimates of mediation models in Study 4*

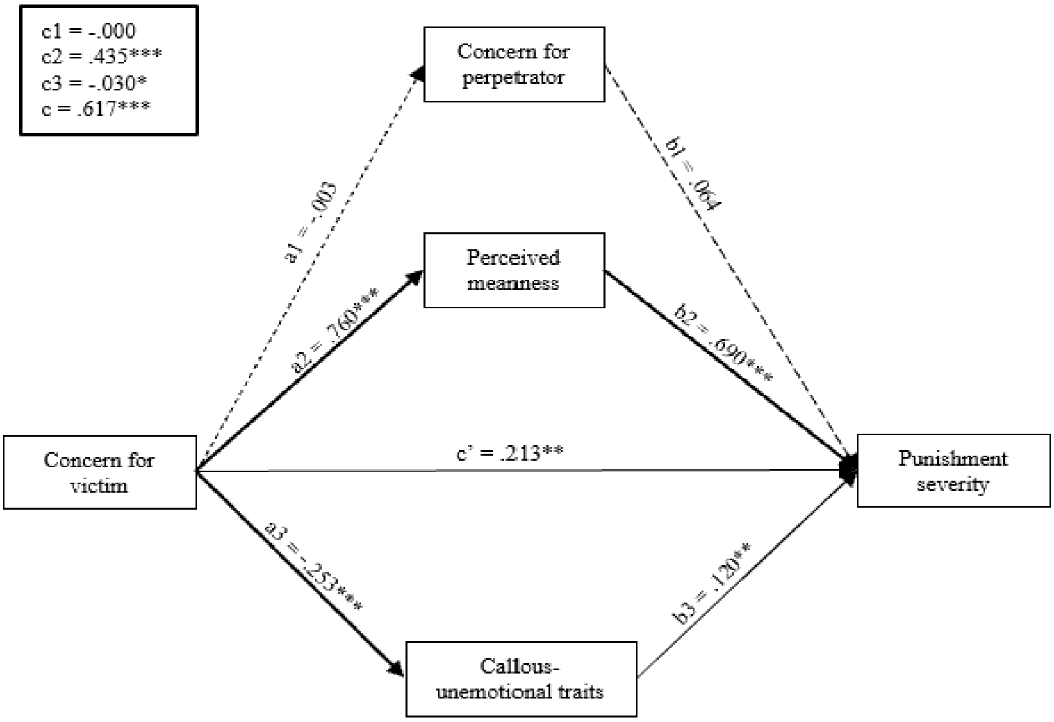
	Intentional			**Accidental**		
	β	***SE***	***$BCCI_{95\%}$***	***β***	***SE***	***$BCCI_{95\%}$***
Model 1						
Total effect	.614	.049	**.513, .704**	.460	.045	**.368, .549**
Direct effect	.614	.049	**.513, .704**	.542	.060	**.420, .652**
Indirect effect	-.000	.003	-.005, .005	-.082	.037	**-.154, -.007**
Model 2						
Total effect	.614	.048	**.510, .702**	.460	.046	**.372, .551**
Direct effect	.166	.075	**.022, .317**	.113	.040	**.033, .190**
Indirect effect	.447	.051	**.353, .557**	.347	.041	**.266, .426**
Model 3						
Total effect	.614	.050	**.510, .710**	.460	.045	**.369, .547**
Direct effect	.659	.046	**.567, 747**	.487	.043	**.404, .571**
Indirect effect	-.045	.015	**-.079, -.018**	-.026	.014	-.056, -.000
Model 4						
Total effect	.617	.048	**.523, .707**	.466	.046	**.373, .557**
Direct effect	.213	.072	**.074, .362**	.166	.056	**.059, .274**
Indirect effect						
via concern for perpetrator	-.000	.003	-.007, .007	-.029	.025	-.085, .022
via perceived meanness	.435	.050	**.343, .540**	.342	.040	**.268, .427**
via callous-unemotional traits	-.030	.014	**-.061, -.008**	-.013	.008	-.032, -.000

Note. Model 1 = Estimates through individual mediation of concern for perpetrator; Model 2 = Estimates through individual mediation of perceived perpetrator meanness; Model 3 = Estimates through individual mediation of callous-unemotional traits; Model 4 = Estimates through the combined mediation of concern for perpetrator, perceived perpetrator meanness and callous-unemotional traits. Reported direct effects are estimated after accounting for indirect effects.

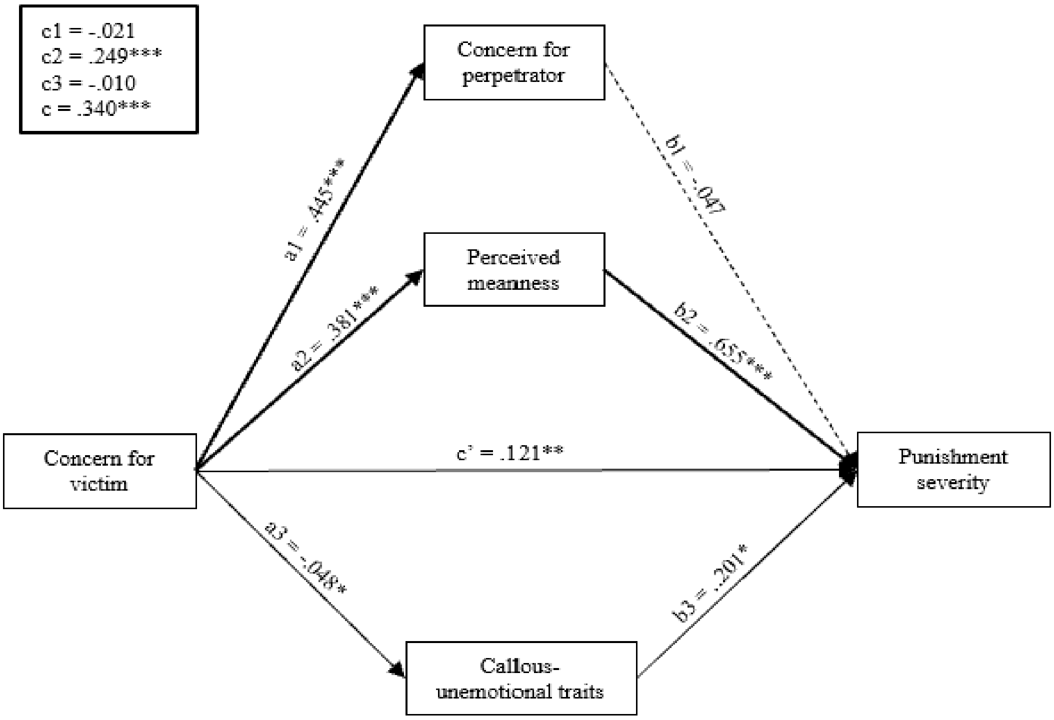
Significant estimates ($BCCI_{95\%}$ not including 0) are in **bold**.

Figure 3.6. | *Path estimates in a) intentional and b) accidental conditions in Study 4*

a)



b)



Note. Mediation was assessed by examining the direct (c') and indirect effects of interpersonal callousness on aggression through concern for perpetrator (c1), perceived perpetrator meanness (c2), and callous-unemotional traits (c3). The total effect (c) was the sum of the indirect and direct effects.

Significance estimated at $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

When introducing all mediators into a combined model, the indirect effects of both perceived meanness, $\beta = 0.43$, $SE = 0.05$, $BCCI_{95\%}$ [0.34, 0.54], and callous-unemotional traits, $\beta = -0.03$, $SE = 0.01$, $BCCI_{95\%}$ [-0.06, -0.01], remained significant when the transgression was intentional, and only perceived meanness emerged as a significant mediator in the accidental condition, $\beta = 0.34$, $SE = 0.04$, $BCCI_{95\%}$ [0.27, 0.43]. Fit indices indicated good fit overall (Supplement 3.2), although in this case the *Intentional* model had the better fit. See models in Figure 3.6.

Discussion

The outcomes of this study largely replicated Study 3, demonstrating the robustness of our observations. Expanding on this, the current study further indicates that the impact of empathy on participants' punishment decisions for perceived aggression diminishes at higher levels of callous-unemotional traits. This observation supports our hypothesis that callous-unemotional traits attenuate the empathic bias in punitive decision-making.

On the other hand, however, participants with higher levels of callous-unemotional traits were also more prone to advocating for harsher penalties against aggressors. This suggests that, at higher levels of callous-unemotional traits, individuals are more predisposed to endorsing harsher punitive measures – which is consistent with links between callous-unemotional traits and aggression highlighted in previous research (Brugman et al., 2017; Camara et al., 2025; Frick et al., 2003; Ritchie et al., 2022). The broader implications of these findings are discussed in greater detail in the following section.

General discussion

Our findings align with existing research on how evaluations of transgressions are largely influenced by perceptions of the perpetrator's intentions. Consistent with the Attribution of Blame model (Shaver & Shaver, 1985), which proposes that perceived responsibility for norm violations drives moral evaluations, our results confirmed that when harm is perceived as intentional – a condition under which blame is more readily assigned – participants were more likely to endorse harsher punishments. This reflects the inherent subjectivity in moral judgements, showing that the same act can be punished differently depending on the perceived intent behind it. Research additionally shows that intentional aggressions are more strongly condemned than accidental ones (Baez et al., 2014; Decety, Michalska, et al., 2012; Young & Saxe, 2008). Our study extends this literature by providing preliminary insights into how such evaluations relate to evaluative judgments and empathic responses towards the perpetrator.

Partially in line with our initial hypotheses, we found that participants expressed greater concern for victims than for aggressors, although their expressed concern for perpetrators did not seem to directly influence decisions to punish. Rather, perceived meanness of the perpetrator emerged as a significant mediator in the relationship between concern for victims and punishment severity, suggesting that negative appraisals may play a more central role in justice-related decisions than empathic concern alone. This is consistent with previous research indicating that biases in punitive decision-making can be largely influenced by people's perceptions of the parties involved (e.g., Bastian et al., 2013; Viki et al., 2012). In fact, it is plausible that perceived perpetrator meanness influences empathic responses to the aggressor as well. That is, seeing the perpetrator as "mean" may serve to justify the withholding of empathy and reinforce punishment (Bastian et al., 2013; Harris & Rice, 2006; Osofsky et al., 2005). In this view, one potential avenue for follow-up research could be to further examine whether the link between concern for perpetrators and punishment severity might be mediated by perceived perpetrator meanness.

Furthermore, the current work contributes to the literature by examining how individual differences in callous-unemotional traits influence these empathic processes and their relationship to punitive attitudes, opening new avenues for further exploration. Our findings indicate that individuals with more callous-unemotional traits exhibit lower concern for victims and a weaker link between victim-focused empathy and punishment, yet still tend to favour harsher punishments overall. Interestingly, this pattern aligns with findings from studies of patients with the behavioural variant of frontotemporal dementia (bvFTD) – a neurodegenerative condition marked by poor affect, diminished empathy and impaired social cognition (Piguet et al., 2011; Rascovsky et al., 2011). Patients with bvFTD have been shown to rate punishments as more severe, regardless of whether harm was inflicted intentionally or accidentally (Baez et al., 2014), a pattern that mirrors the reduced concern and elevated punishment ratings observed at higher levels of callous-unemotional traits in our study. These convergences support the view that traits associated with emotional detachment and reduced affective resonance may reduce the impact of perceived intentionality on moral evaluations.

Supporting this interpretation, our data further showed that in scenarios involving accidental harm, callous-unemotional traits did not moderate the relationship between concern for victims and punishment, even though they were correlated with increased punishment severity in these cases. This could suggest that punitive decisions among individuals with higher levels of callous-unemotional traits may be less guided by empathic sensitivity or moral evaluation, and more by a rigid or indiscriminate orientation towards punishment – consistent with prior research indicating that callous-unemotional traits confer unique risks for more instrumental or severe forms of aggression (e.g., Brugman et al., 2017; Camara et al., 2025; Frick et al., 2003). These possibilities gain special relevance in legal and forensic contexts, where individual traits may predispose certain decision-makers (e.g., jurors) to punitive bias, regardless of contextual factors like intent. They also underscore the importance of considering personality-driven variations in justice motivation, which may help inform targeted interventions or training programs aimed at promoting fairness and reducing bias in evaluative judgements.

Altogether, these findings highlight the nuanced role of perceptions of both aggressions and aggressors in shaping punishment decisions, suggesting that punitive attitudes are influenced by a combination of empathy, cognitive evaluation and context. However, these interpretations should be approached with caution due to the study limitations. One key issue lies in the underlying assumption that emotion and partiality are intrinsically linked, suggesting that impartiality is more unemotional or rational. While the ideal of judicial dispassion is deeply ingrained in legal theory and popular opinion, recent arguments (e.g., Maroney, 2024) suggest that emotions can be effectively managed and do not inevitably lead to bias or partiality. Relatedly, while our analysis has primarily focused on emotional influences, elements such as social identity, familiarity, and other contextual variables might also predispose individuals to certain biases in their moral judgments (e.g., Abbink & Harris, 2019; Schiller et al., 2014). Consequently, the scope of our analysis is limited, and future research should consider these interactive factors to provide a more comprehensive understanding of the dynamics underlying punitive decision-making.

In addition, the lack of a consistent definition of empathy poses challenges in generalising our findings to its various dimensions. As noted by Brown (2012, p. 386): "Because the definitions and understandings of empathy we work with are generally positioned in the realm of the abstract, then empathy 'in the field' and 'on the ground' may require alternative dimensions in its articulation." In other words, given the multifaceted nature of empathy, reliance on one conceptualisation may risk oversimplifying the role empathy plays on behavioural outcomes, as it neglects the contextual and situational factors that influence how empathy is expressed in everyday social behaviour. As such, it is essential to clarify which aspect of empathy is under investigation – be it cognitive or affective – and to consider how these facets manifest differently across various social contexts (Vachon & Lynam, 2016). This distinction is crucial because each facet of empathy might interact with moral judgment and punitive decision-making in unique ways. Although fully resolving these definitional ambiguities is beyond the scope of our current work, it highlights an important consideration that should be explored in future studies looking into the role of empathy in punishment.

Moreover, the use of animated images to elicit empathy is another key limitation in the current research. While animated stimuli may offer a more dynamic representation compared to static images (Decety et al., 2012), they cannot fully capture the complexity and nuance of real-life situations. In everyday scenarios, moral judgments are often influenced by the contextual details surrounding the actions of the perpetrator and victim (Jin & Peng, 2021). By omitting such contextual information, our study may have limited the scope of participants' responses, preventing a more comprehensive understanding of the factors that influence punitive decision-making. Future research that incorporates more detailed background context may help capture a broader range of reactions and provide a more accurate reflection of real-world moral judgments.

Finally, while our sample was designed to examine patterns in a normative population, our exploration of callous-unemotional traits would benefit from studying a sample with a broader or more pronounced distribution of these traits. Individuals with higher levels of callous-unemotional traits – such as those classified as *psychopaths* – may exhibit more distinct patterns of empathy and punishment, offering a deeper understanding of their role in justice-related decisions. This approach could provide a more nuanced account of how specific socio-affective traits shape punitive attitudes and behaviours.

Conclusion

Overall, the common thread across our findings is that people's punishment decision-making is influenced by their perceptions of both the aggression and the aggressor, although the latter seems to have a greater effect. While punishment severity significantly increased for intentional transgressions, negative perceptions of the perpetrator influenced punishment regardless of the intentionality of the transgression itself. This suggests that punitive responses are not solely guided by principles of justice – such as proportionality or the need for correction – but are also susceptible to biases rooted in how we perceive those we judge. Notably, the tendency for individuals with more callous-unemotional traits to endorse harsher punishment regardless of perceived intentionality suggests that punishment may also reflect something deeper in an individual's moral compass, such as a greater predisposition to engage in interpersonal harm (especially if there is no motivation for retribution or vengeance). These reflections provide a basis for future research to explore the broader implications of callous-unemotional traits in moral judgement and punitive decision-making.

Chapter IV.

On the Topic of Pain Empathy: An EEG and Machine Learning Approach to Vicarious Pain

Abstract

Observing others in pain typically triggers automatic neural and physiological reactions that are often considered markers of empathy. However, the extent to which one's empathic predispositions and sensitivity to pain contribute to these reactions is uncertain. The study included in this chapter examines whether the neurophysiological response underlying pain perception corresponds to traits of empathy, callous-unemotional traits, and aggression. Using EEG, we recorded neural activity from 37 healthy participants as they passively viewed painful and neutral scenarios, while also evaluating their subjective ratings of the target's pain. We analysed EEG data using cluster-based permutation tests and employed classification algorithms to assess pain-related EEG fluctuations on a trial-by-trial basis. Our goal was to determine whether neural responses to observed pain would correspond to participants' subjective pain ratings as well as to their self-reported levels of empathy, callous-unemotional traits, and aggression. The results revealed trends pointing at changes in pain-related EEG activity specific to the expression of callous-unemotional traits, although effect sizes were relatively small. Moreover, our classifiers' accuracy did not achieve significance above chance levels, revealing that pain-related EEG fluctuations were not reliably distinguishable from non-painful stimuli at the individual trial level. The significance of these results is discussed.

"My greatest wish for humanity is not for peace or comfort or joy. It is that we all still die a little inside every time we witness the death of another. For only the pain of empathy will keep us human."

— Neal Shusterman in 'Scythe' (2017, p.338)

Introduction

Picture yourself immersed in a gripping movie plot where the main character loses control of their speeding car, hurtling towards a cliff. Even though you are fully aware that this is just a fictional scenario, you might notice your heart rate increasing, as if you were right there in the driver's seat. Similarly, when witnessing someone accidentally stub their toe, you might instinctively touch your own foot, almost as if it had happened to you. These reactions reflect our ability to mentally immerse ourselves into others' experiences and "feel" their pain without experiencing it first-hand (Bernhardt & Singer, 2012; de Vignemont & Singer, 2006; Lockwood, 2016). For this reason, second-hand pain observation is often seen as an ecologically valid way to investigate empathy (e.g., de Tommaso et al., 2019; Fabi & Leuthold, 2017; Mu et al., 2008; Perry et al., 2010; Whitmarsh et al., 2011). However, as discussed in previous chapters, empathy is a multifaceted construct and "one's own emotional experiences restrict the kind and degree of empathy that can be felt" (Bird & Viding, 2014, p.527).

In this chapter, I draw from the concept of pain empathy to investigate the extent to which individuals' empathic predisposition correlates with their perception of others' pain. I specifically examine whether the expression of different empathy facets, callous-unemotional traits and aggression corresponds to the vicarious response to observing physical pain in others, as measured with EEG.

Investigating pain empathy through EEG

Pain empathy encompasses the ability to detect, recognise, and resonate with another's experience of pain and is thought to be an early-emerging marker of empathic development (Bernhardt & Singer, 2012). Evidence of this can be seen as early as infancy, with newborns exhibiting distress in response to the cries of other infants or mirroring caregivers' emotional expressions (Hoffman, 1975; Sagi & Hoffman, 1976; Zahn-Waxler et al., 1979; see also

Davidov et al., 2021). These initial responses reflect emotional mimicry, engaging sensorimotor processes that are also observed during firsthand experiences of pain (Decety & Michalska, 2010; Lamm et al., 2011). Throughout brain development, these primal reactions become integrated into more complex neural systems within brain areas like the AI, ACC, and OFC, as evidenced by neuroimaging studies (e.g., Decety et al., 2012; Shamay-Tsoory et al., 2009; Zaki et al., 2016). Higher-order affective representations, in turn, enable individuals to better understand others' mental and physical states through their own emotional and bodily experiences, thereby fostering a shared sense of pain believed to reflect empathic processing (Levy et al., 2018).

As such, paradigms that elicit vicarious pain responses have been widely implemented to study pain empathy in experimental contexts. These paradigms typically involve images depicting nociceptive stimuli to others, such as hands or feet in situations of physical harm, or painful facial expressions (Coll, 2018; Jauniaux et al., 2019), and reveal that observing, imagining or merely anticipating pain in others engages emotional and somatosensory neural pathways overlapping with those involved in direct experiences of pain (Banissy & Ward, 2007; Decety & Meyer, 2008; Fallon et al., 2020; Lamm et al., 2011; Lockwood, 2016; Singer et al., 2004; Vachon-Presseau et al., 2012). This overlap in neural activation can be tracked in real time via EEG, making it a valuable tool for exploring the temporal dynamics of pain perception.

EEG reveals larger amplitudes in event-related potentials (ERPs) that are time-locked to the observation of painful stimuli. Early ERPs typically involve frontocentral negative deflections emerging approximately 100–200 ms after stimulus onset, which are thought to reflect the rapid and automatic processing of pain-related cues (Fan & Han, 2008). Pain observation also involves positive deflections, including the P3 component (~300 ms) and the late positive potential (LPP, ~400–1000 ms), with the maximal effect recorded over centroparietal electrodes (Coll, 2018). These later ERPs are thought to reflect more complex cognitive processes, including sustained attention and evaluative processing (de Vignemont & Singer, 2006; Fabi & Leuthold, 2017; Fan & Han, 2008). Notably, both early and late ERPs correlate with the intensity and unpleasantness of the pain perceived in others (Cheng et al., 2012; Decety et al., 2010; Fabi & Leuthold, 2017; Han et al., 2008; Meng et al., 2012), although meta-analytic evidence reveals more robust and reliable effects on late components – particularly the LPP (see Coll, 2018).

Furthermore, studies report that witnessing others in painful situations elicits changes in alpha (8–12 Hz) and beta (13–30 Hz) oscillatory power – mostly related to the desynchronisation of groups of neurons over centroparietal regions (T. Chen et al., 2023; Cheng et al., 2008; Fabi & Leuthold, 2018; Lübke et al., 2020; Motoyama et al., 2017; Mu et al., 2008; Perry et al., 2010; Riečanský et al., 2015; Valentini et al., 2012). These event-related dynamics (ERDs) resemble neural patterns observed during first-hand pain experiences (Ploner et al., 2006; Riečanský et al., 2015; Whitmarsh et al., 2011). In particular, sensorimotor alpha desynchronisation – commonly referred to as *mu suppression* (Kuhlman, 1978) – has been associated with the perceived intensity of observed pain (Babiloni et al., 2006; Hoenen et al., 2015). Additionally, research has shown correlations between self-reported unpleasantness and increases in theta power (3–5 Hz) over parietal regions (Mu et al., 2008) – an effect also seen during direct pain and tactile stimulation (Michail et al., 2016). As such, EEG activity locked to the observation of pain in others is considered a vicarious reaction reflecting an individual's ability for empathy (Bird & Viding, 2014).

Nevertheless, inconsistencies in the literature raise important questions about whether changes in EEG activity during pain observation can be attributed to empathy. While some studies have found significant correlations between pain-related EEG changes and self-reported dispositional empathy (Cheng et al., 2008; Corbera et al., 2014; Fabi & Leuthold, 2017; Gonzalez-Liencres et al., 2016; Lübke et al., 2020; Vaes et al., 2016), others have failed to replicate these effects (Chen et al., 2023; Cogoni et al., 2023; Fabi & Leuthold, 2018; Perry et al., 2010; Van Dongen et al., 2018; Yang et al., 2009). For example, one study found that mu suppression did not correlate with empathy scores but was instead influenced by participants' mood (Li et al., 2017). Similarly, a study comparing fibromyalgia patients with healthy controls found group differences in EEG responses during second-hand pain observation but no significant correlations with trait empathy in either group (de Tommaso et al., 2019). Moreover, while previous research has reported a significant correlation between trait empathy and pain-related ERPs (Fabi & Leuthold, 2017), a recent systematic review reports no significant associations between empathy scores and ERP components in empathy-related tasks (Almeida et al., 2024). These discrepancies underscore the need to re-evaluate the extent to which an individual's capacity for empathy can be examined through their electrophysiological responses to second-hand pain.

Individual differences in pain processing: the role of callousness and aggression

Callous-unemotional traits offer a valuable framework for evaluating the link between empathic predisposition and the EEG response to second-hand pain (Branchadell et al., 2024). Individuals with these traits often show insensitivity to others' feelings and experiences, hence reflecting an inherent lack of empathy. Studies report distinct patterns of pain-related EEG fluctuations in these cohorts. For example, Decety *et al* (2015) found that participants with more callous-unemotional traits exhibited lower P3 and LPP amplitudes during observed pain, particularly when they were prompted to focus on their concern for others. More recent work similarly reports that traits of callousness such as psychopathic meanness[8] predict lower LPP amplitudes in pain perception tasks (Branchadell et al., 2024; Brislin et al., 2022), with both self- and other-perspective ratings of pain intensity being lower among participants exhibiting these traits (Brislin et al., 2022). Additionally, psychopathic meanness has been linked to lower LPP amplitudes in response to perceived harm to others (Van Dongen et al., 2018). These patterns indicate that callous-unemotional traits may predict differences in pain processing (Brislin et al., 2022), which could in turn be related to the tendency to underestimate pain in individuals exhibiting these traits and increase their risk for aggression (Branchadell et al., 2024).

Indeed, research shows that individuals who frequently engage in aggressive behaviour often exhibit lower sensitivity to pain (e.g., Niel et al., 2007; Reidy et al., 2009). While direct investigations into the relationship between trait aggression and EEG activity during pain observation are scarce, some studies reveal distinct patterns of neural activity among individuals with more aggressive traits. For example, trait aggression has been linked to asymmetrical brain activity in the alpha and beta frequencies within the left hemisphere (Hofman & Schutter, 2012; Niv et al., 2018; Rybak et al., 2006). Notably, left-sided over right-sided cortical excitability has been associated with a higher drive for goal-oriented behaviour and a reduced likelihood to avoid potential threats or aversive stimuli (Harmon-Jones & Winkielman, 2007; Schutter et al., 2008). At the same time, however, research also suggests no direct link between aggressive traits and EEG responses to emotional stimuli. A study by ter Harmsel *et al* (2022) compared electrophysiological responses (P3, LPP and mu suppression) between delinquent young adults and controls during passive viewing of aggressive interactions, reporting no significant differences between the groups and no correlations with reactive and/or proactive aggression. Altogether, these findings suggest that individuals more prone to aggression might not have immediate differences in perceiving others' pain but may instead be less inhibited by aversive responses to such pain – although current EEG evidence is scarce to make these conclusions.

Furthermore, studies have found that both criminal psychopaths (Cheng et al., 2012) and community samples with callous-unemotional traits (Decety et al., 2015) exhibit increased mu suppression during pain observation. While previous studies refer to mu suppression as a vicarious response to pain (Mu et al., 2008; Whitmarsh et al., 2011), the finding that individuals with callous-unemotional traits exhibit increased mu suppression to second-hand pain suggests otherwise, given the inherent empathy deficits of these cohorts. This is consistent with research reporting no significant correlations between pain-related mu suppression and empathy (Li et al., 2017; Perry et al., 2010), which further suggests that mu suppression might not directly reflect an individual's empathic response/sensitivity to others' pain. An alternative explanation is that sensorimotor alpha desynchronisation to second-hand pain might instead reflect perceptual processing (Whitmarsh et al., 2011). Nevertheless, it is worth noting that other research has failed to find correlations between spectral power changes during pain observation and callous traits (Van Dongen et al., 2018). These mixed results highlight limitations in current interpretations of EEG activity in pain empathy research.

Insights from machine learning

Recent advancements in machine learning have provided new perspectives on the specificity and predictability of EEG responses to pain. Unlike traditional EEG analyses, which rely on averaging across trials, machine learning allows for single-trial classification. This capability has led to promising initial findings, suggesting that EEG responses to pain are, to some extent, predictable (Dinh et al., 2019; Mari et al., 2022, 2023a, 2023b). Seminal work by Mari *et al* (2022) systematically reviewed the application of machine learning in pain-related EEG analysis, demonstrating that models can differentiate painful from non-painful stimuli based on EEG features. Across two studies, random forest models were trained on EEG data for the classification of trial-by-trial time-frequency features related to perceived intensity during direct pain (Mari et al., 2023a; Mari et al., 2023b). In both studies, the models effectively classified EEG signals during low and high pain trials, with classifiers achieving validation accuracies around 58–73%. In contrast, classifying EEG responses to second-hand pain has proven more challenging. In a later study, Mari *et al* (2023c) tested the ability of machine learning to distinguish pain-related EEG activity during passive observation of painful and non-painful scenarios with human targets. Their analysis revealed that the models' accuracy in classifying pain-specific responses did not exceed chance levels. Instead, the models appeared to be more sensitive to the visual features of the stimuli, such as distinguishing between scenes and faces. Similar patterns were observed in a follow-up study using an active paradigm, wherein machine learning accurately classified visual stimuli – but not pain *vs* no-pain classes – above chance levels (Mari et al., 2025). These outcomes suggest that the EEG response to visual cues in pain perception paradigms is more salient than pain-specific responses (Mari et al., 2025), which emphasises the need for caution when interpreting EEG signals as indicators of vicarious pain and empathy.

Adding a new perspective to this research, however, a recent study conducted by Wang *et al* (2025) has shown that socio-affective factors can enhance classifier accuracy. Using an emotion induction paradigm, the authors examined whether guilt influenced machine learning classification of EEG responses to observed pain. Participants viewed images of hands in painful and non-painful scenarios under two conditions – one designed to elicit guilt and another serving as a neutral control. The results showed that when participants experienced guilt, their EEG responses to observed pain became more distinct, particularly in P3 amplitudes. Under these conditions, classifiers achieved higher accuracy. This suggests that emotional states such as guilt can heighten the neural differentiation of pain-related stimuli, potentially influencing the reliability of EEG-based pain classification. Moreover, the findings support the idea that personality traits, such as callous-unemotional traits – which are associated with deficits in guilt experience (Waller et al., 2020) – may influence neural responses to second-hand pain. However, a critical limitation of Wang *et al*'s (2025) study is that classifier accuracy was not tested against chance levels, raising concerns about whether the observed improvements truly reflect enhanced neural specificity or are simply due to random fluctuations.

Project rationale

Research on pain empathy reiterates that individual differences in empathy may influence neural responses to second-hand pain, though these effects remain inconsistent. In our study, we combine EEG and machine learning to examine how perceiving others' physical pain relates to empathy, callous-unemotional traits, and aggression in a community sample. Participants passively viewed painful and neutral scenarios involving human targets while their EEG activity was recorded and their perceptions of the targets' pain assessed. They also completed self-report measures of empathy, callous-unemotional traits, and aggression. Based on prior work (Branchadell et al., 2024; Brislin et al., 2022; Cheng et al., 2012; Decety, 2015; ter Harmsel et al., 2022; Van Dongen, 2020), we hypothesised that callous-unemotional traits would negatively correlate with P3 and LPP amplitudes (Brislin et al., 2022; Decety et al., 2015). Due to limited prior research on spectral power changes (e.g., Decety et al., 2015; van Dongen et al., 2018), we made no predictions in the time-frequency domain.

In addition to exploring these hypotheses, we employed machine learning techniques to classify pain-related EEG data on a trial-by-trial basis, evaluating the accuracy of classifiers in distinguishing pain *vs* no-pain conditions. While previous studies have primarily focused on ERP features for classifying EEG responses to second-hand pain (Mari et al., 2023c, 2025), we sought to extend this work by also integrating classifiers trained on EEG features at the time-frequency domain. Our goal was to determine whether the robust performance of time-frequency classifiers observed in direct pain studies (Dinh et al., 2019; Mari et al., 2022, 2023a, 2023b) could be replicated for second-hand pain. Furthermore, we aimed to explore whether the performance of these classifiers could be influenced by participants' socio-affective traits, though this analysis depended on the classifiers achieving accuracy above chance levels.

Study 5

Method

Participants and procedure

The study recruited right-handed young adults to align with prior research indicating lateralisation in pain perception (e.g., Hofman & Schutter, 2012; Timmers et al., 2018). Handedness was determined through an online screening survey using the Edinburgh Handedness Inventory (Oldfield, 1971), where a positive laterality quotient (LQ) indicates right-handedness (McMeekan & Lishman, 1975).

A-priori power analyses indicated that 34 participants were required to detect a moderate effect size ($\rho = 0.5$) with a 90% power. Initially, 50 respondents volunteered to take part. However, four withdrew before completing the experiment, eight were excluded due to significant data contamination (e.g., excessive ocular or motion artifacts in the EEG signal), and one was excluded due to a negative LQ. This left us with a final sample of 37 participants, including 20 men and 17 women aged 20–35 years old ($M = 24.49$, $SD = 3.15$). All participants had normal or corrected-to-normal vision and reported no major neurological condition at the time of participation. After the screening phase, eligible participants completed a second online survey (~15 min), followed by a lab session (~2–3 h). However, survey data was missing for one participant, leaving a total of 36 participants (19 men) for our descriptive analyses. Participants were reimbursed with a £10 voucher after completing the experiment.

All study procedures adhered to the ethical guidelines of the Declaration of Helsinki and were approved by the University of Essex Science and Health Ethics Sub-committee.

Online survey

The survey included self-report measures assessing trait empathy, callous-unemotional traits, and aggressive behaviour:

Affective and Cognitive Measure of Empathy (ACME; Vachon & Lynam, 2016). The ACME was selected as it captures aspects related to emotion contagion (e.g., "I feel awful when I hurt someone's feelings") and mentalisation (e.g., "I can usually tell how people are feeling") – both of which are central to the experience of vicarious pain (Bernhardt & Singer, 2012).

Inventory of Callous Unemotional Traits (ICU; Frick et al., 2003). The ICU subscales tap into different facets of socio-affective behaviour that are relevant for the concept of pain empathy, including items related to concern for others (e.g., "The feelings of others are unimportant to me"), caring behaviours (e.g., "I try not to hurt others' feelings", reverse coded), and emotional responsiveness (e.g., "I express my feelings openly") (Decety et al., 2012; Goubert et al., 2005).

Reactive and Proactive Aggression Questionnaire (RPQ; Raine et al., 2006). The RPQ subscales reflect differences in affective sensitivity, considered relevant for pain perception (X. Li et al., 2020). Items for reactive aggression relate to heightened emotionality and impulsivity (e.g., "Damaged things because you felt mad"), whereas items for proactive aggression reflect more emotionless and calculative behaviours (e.g., "Hurt others to win a game").

Lab session

During the lab session, participants performed two experimental tasks presented on *E-Prime* (version 2.0). The experiment began with a *Passive Viewing Task.* In this task, participants were exposed to images either depicting physical pain to a human target or a matched scenario involving no pain (Figure 4.1), with EEG activity being recorded and monitored throughout. Following this, participants completed a *Pain Rating Task*, where they rated the perceived intensity of pain experienced by the target in each scenario.

Experimental design and stimuli

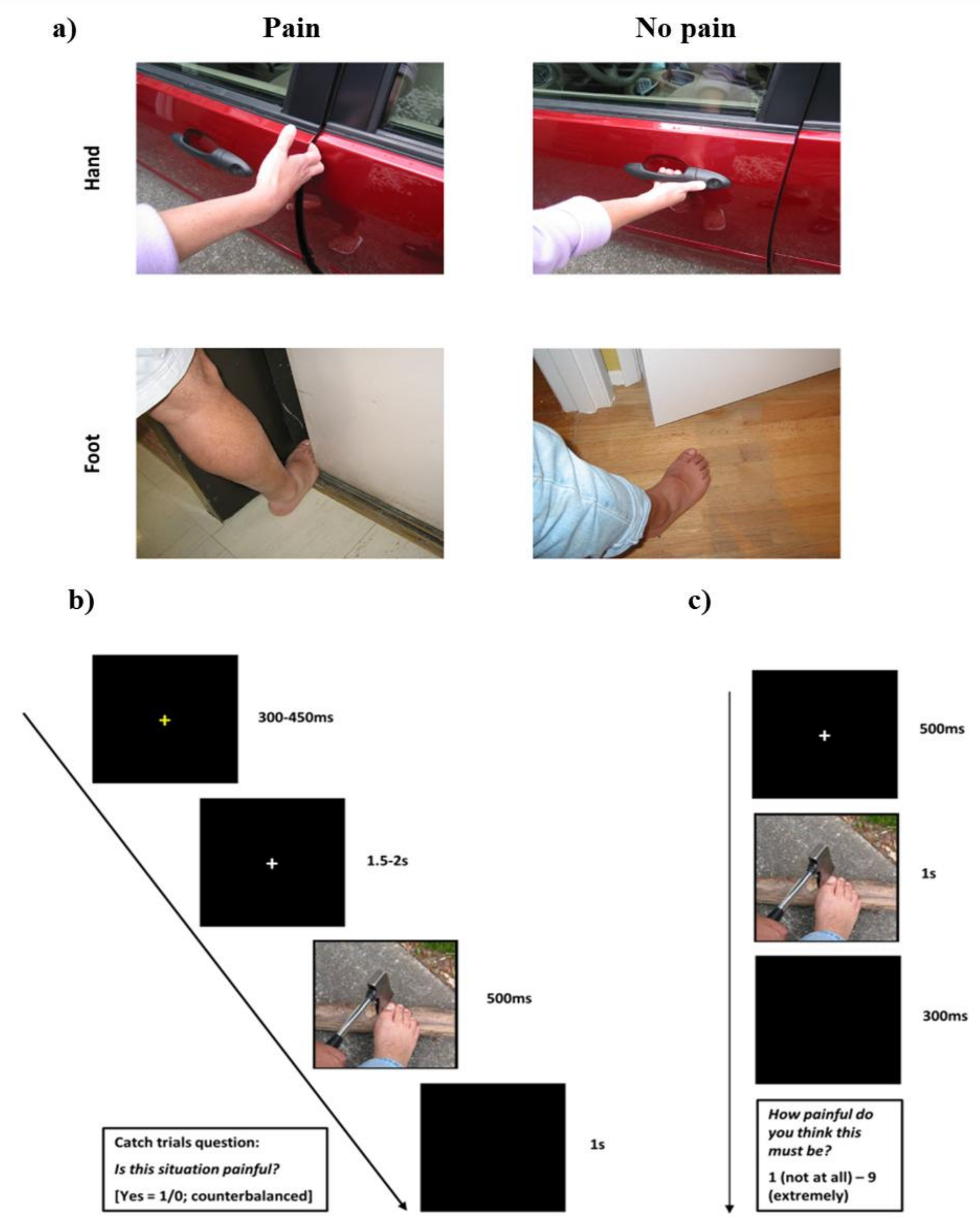
The experiment followed a 2x2 factorial design, manipulating limb type (hands *vs* feet) and condition (painful *vs* neutral). The images were taken from a first-person perspective and were designed to represent identical actions, with the painful conditions including an additional element that suggested pain (Jackson et al., 2006). The *Passive Viewing Task* consisted of four blocks, each containing 50 trials. Each trial began with a yellow fixation cross, which lasted between 1.3–1.45 s, prompting participants to blink and minimise eye movement during EEG recording. This was followed by a white fixation cross for 1.5–2 s, indicating the start of a new trial. The images were presented for 500 ms in a random order. In 20% of the trials, participants were instructed to indicate whether the preceding image was painful by pressing ‘1’ or ‘0’ on a keyboard. These catch trials were used to verify attention to the stimuli but were excluded from EEG analysis. The *Pain Rating Task* consisted of 64 trials, equally distributed across four blocks (32 trials per condition). In each trial, an image was displayed for 1 s, followed by a rating period where participants used the mouse in their right hand to rate the perceived intensity of pain 300 ms after the image’s appearance. A 500 ms pre-stimulus fixation cross was presented before each image to ensure a consistent starting point for each trial. Participants rated the perceived pain intensity on a scale from 1 (*no pain*) to 9 (*extreme pain*).

EEG data acquisition and analysis

EEG recording

EEG activity was recorded using an elastic EEG cap equipped with 60 active scalp electrodes. Electrode placement followed the 10M equidistant layout (Easycap Equidistant Layouts, 2018). To ensure precise positioning, anatomical landmarks (pre-auricular points, nasion, and inion) were used to locate the head centre, referenced to the Cz electrode. The ground electrode was placed at AFz, and electro-oculogram (EOG) electrodes were positioned above and below the left eye to monitor ocular artifacts. All electrodes were online-referenced to the left mastoid. EEG signals were continuously monitored, maintaining electrode-to-skin impedances below 50 kΩ. Data were recorded using BrainAmp amplifiers (BrainProducts, Munich, Germany) with a 0.1 µV analogue-to-digital conversion and a sampling rate of 1000 Hz.

Figure 4.1. | *Experimental tasks*



Note. Illustration of a) images included in each condition; b) the experimental design for the passive viewing task, including the catch trials question; and c) the experimental design for the rating task, including the pain intensity question and rating scale.

Data processing

EEG data were processed using custom MATLAB scripts with functions from Fieldtrip (Oostenveld et al., 2011). All scripts for data processing can be accessed through OSF: https://osf.io/f2p9b/?view_only=076befab2730484ca1257d2eb7939d29.

Preprocessing. The raw signal was filtered (0.1–30 Hz) and downsampled to 500 Hz. Epochs were segmented by pain condition (painful *vs* neutral) and limb (hand *vs* foot), spanning -1.5–2 s relative to onset. While 200 trials per participant were initially planned (40 main trials and 10 catch trials per condition), technical issues reduced this to an average of 159.54 (SD = 4.13). Post-hoc analysis found this reduction non-significant (p = .290), with most participants retaining at least 50 trials per condition (see Supplement 4.1).

Epochs exceeding ±100 µV in any channel were automatically removed, followed by visual inspection. Independent Component Analysis (ICA) eliminated physiological artifacts (e.g., eye blinks, muscle activity), with an average exclusion of 3 components removed per participant (SD = 0.77; Supplement 4.2). Noisy trials were interpolated from neighbouring electrodes, but those requiring interpolation of more than 5 channels were excluded. Data were then re-referenced to the average of all scalp electrodes, excluding non-scalp electrodes, for further analysis.

ERPs. For ERP analyses, baseline correction was performed by subtracting the signal from -500 to -100 ms before stimulus onset. Individual ERPs were averaged across participants per condition. A topographical analysis examined spatial ERP distributions across electrode sites within a 500–1000 ms post-stimulus window. Statistical significance was determined via a nonparametric randomization test with 5000 iterations (Maris & Oostenveld, 2007). This analysis identified significant ERP distributions across frontal, central, central-parietal, and parieto-occipital electrodes, which were grouped into regions of interest (ROI) for further analysis. To estimate significant temporal indices within ROI, we used a cluster-based permutation using repeated measures *t* tests with 5000 iterations (Maris & Oostenveld, 2007). We used Holm-Bonferroni correction to address multiple comparisons.

ERDs. Spectral analyses used a multitaper time-frequency transformation to analyse individual epochs across frequencies from 4–30 Hz, in 1 Hz steps. This method converts time-domain EEG signals into the frequency domain using a Fast Fourier Transform (FFT) algorithm, which helps us understand how signal power changes over time and across different frequencies (Goldfine et al., 2011). To improve computational efficiency, we applied zero-padding to extend the length of the signal and used a Hanning taper with 3 cycles for each time window during the FFT. Additionally, we performed a relative baseline correction within a time window from -1.1–0 s before stimulus onset.

At the group level, time-frequency activations were computed using a nonparametric approach with 5000 iterations (Maris & Oostenveld, 2007). This method involved extracting individual time-frequency data points and identifying statistically significant clusters across time, space, and frequency relative to the baseline. Topography analysis showed significant activation at electrode sites distributed across frontal, centroparietal, and parietooccipital regions. Similar to the ERP approach, 5000-iteration repeated measures *t* tests identified significant estimates in the time-frequency domain, with multiple comparisons corrected using the Holm-Bonferroni method.

Machine learning classification

We developed two models to identify EEG activity related to pain *vs* no-pain conditions at both the time and time-frequency domains. Both models were optimised for our dataset using automatic parameter selection to ensure robust performance and avoid overfitting.

ERP classification

For ERP classification, we employed shrinkage linear discriminant analysis (LDA), an advanced form of traditional LDA that improves classification accuracy by stabilising covariance estimation, particularly in cases with high-dimensional data and limited training samples (Blankertz et al., 2011). This approach enhances feature selection by transforming the EEG data into a lower-dimensional space where the most discriminative information is retained, facilitating the identification of pain- and no-pain-related ERP components.

The raw EEG data was downsampled to 20 Hz to reduce computational complexity while preserving the critical temporal structure of ERPs (Rivet et al., 2009). The data was then epoched within a 0.2–1.2 s window after stimulus onset, selected based on prior research demonstrating that pain-related ERP components typically emerge within this time frame (Coll, 2018). This approach ensured that the classification focused on periods where pain-related neural activity is most prominent.

ERD classification

For ERD classification, we used the Common Spatial Pattern (CSP) algorithm, which identifies spatial patterns that maximise class separation (Müller et al., 2008). CSP is particularly effective in motor and cognitive EEG applications but is known to be sensitive to the number of features included, which can lead to overfitting and reduced generalisation performance (Lemm et al., 2005; Wu et al., 2015). To mitigate this issue, we first applied Principal Component Analysis (PCA) to reduce the dimensionality of the data to 60 principal components. The top four components were selected for classification efficiency.

The time window for ERD classification was set between 0.25–0.75 s post-stimulus, as previous research suggests that pain-related ERD, particularly in the alpha band, is prominent within this interval (e.g., Hu et al., 2013; Whitmarsh et al., 2011; Zebarjadi et al., 2021). A bandpass filter of 9–11 Hz was applied to target the low-alpha to mid-alpha frequency range, where pain-related oscillatory modulations have been reported (e.g., Babiloni et al., 2006; Zebarjadi et al., 2021). This method aimed to enhance classifier performance by isolating pain-relevant oscillatory activity while minimising the influence of non-specific EEG fluctuations.

Assessment of classifier performance

To determine whether the classifiers' performance in identifying pain-related EEG patterns exceeded chance levels, we applied the binomial cumulative distribution thresholding method proposed by Combrisson and Jerbi (2015). We computed the statistical threshold in MATLAB using the following function, as described by Mari *et al*'s (2025):

$$Statistical\ Threshold = binoinv\left(1 - a, n, \frac{1}{c}\right) * \frac{100}{n}$$

where:

- α = significance level (0.05 for a 95% confidence threshold),
- n = total number of classification trials (in our study, n ~ 160 trials),
- c = number of classes in the classification task (for binary classification, c = 2),
- *binoinv* = MATLAB's inverse binomial cumulative distribution function, which determines the accuracy threshold that must be exceeded for significance.

The results indicated that for statistical significance at the 0.05 level, our models' accuracy must exceed 56.25% (Combrisson & Jerbi, 2015).

Results

Sample characteristics

Survey scores were analysed on SPSS (IBM SPSS Statistics for Windows, Version 29.0). On average, participants' levels of empathy, callous-unemotional traits, and aggression fell within estimated normative levels (Baker et al., 2008; Byrd et al., 2013; Dryburgh & Vachon, 2019). Descriptive statistics for total and subscale scores are presented in Table 4.1.

Table 4.1. | *Sample characteristics*

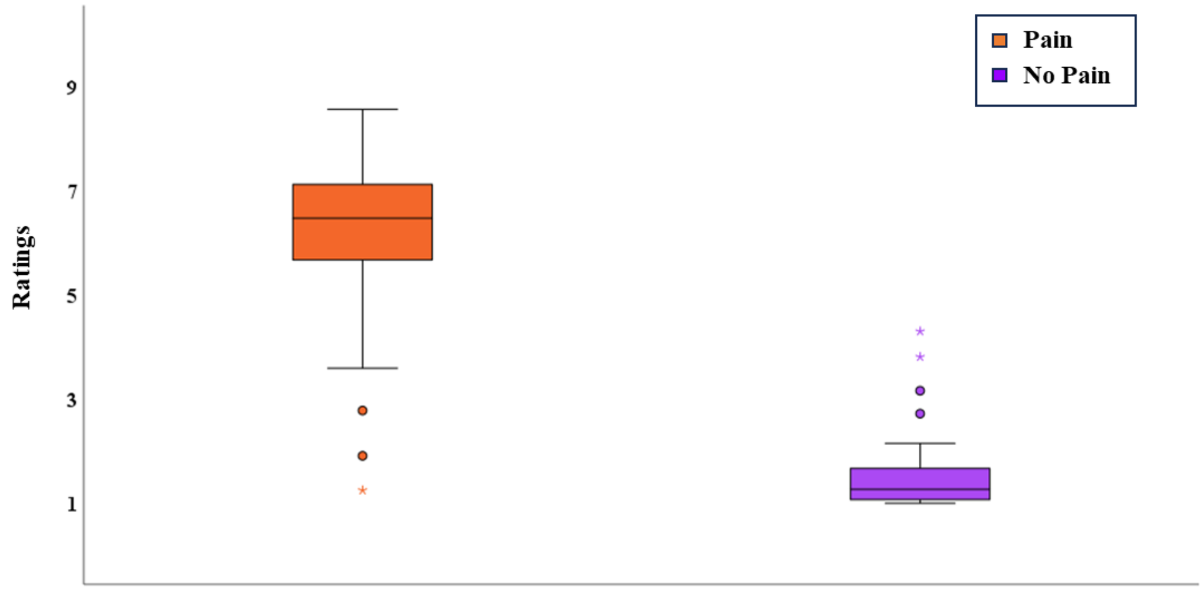
	All			Men			Women		
Variables	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***
Age	37	24.49	3.11	20	25.25	2.34	17	2.59	3.71
Ratings	37	4.59	.94	20	4.2	1.29	17	5.07	.28
RTs	37	585.41	663.22	20	404.09	532.15	17	776.95	757.33
Understanding	36	39.94	5.01	19	39.37	5.29	17	40.59	4.74
Resonance	36	51.94	5.13	19	50.53	5.18	17	53.53	4.72
Dissonance	36	44.86	3.23	19	44.53	3.19	17	45.24	3.35
ACME total	36	136.75	9.29	19	134.42	9.24	17	139.35	8.90
Callousness	36	6.06	2.99	19	7.16	2.67	17	4.82	2.92
Uncaring	36	5.50	2.50	19	6.26	2.18	17	4.65	2.62
Unemotional	36	7.50	2.95	19	7.95	2.55	17	7.00	3.35
ICU total	36	19.06	6.39	19	21.37	4.75	17	16.47	7.11
Reactive	36	6.06	3.44	19	6.53	3.31	17	5.53	3.61
Proactive	36	1.36	1.90	19	1.79	1.75	17	.88	2.00
RPQ total	36	7.42	4.77	19	8.32	4.45	17	6.41	5.05

Note. *n* = number of participants, *M* = Mean, *SD* = Standard Deviation. Participant age is reported in years. Rating times (RTs) are reported in milliseconds. Ratings and RTs indicate the difference in participants' perceived pain intensity and rating times in pain *vs* no-pain conditions. Scores are presented for the RPQ (Reactive and Proactive Aggression Questionnaire), ICU (Inventory of Callous-Unemotional Traits), and ACME (Affective and Cognitive Measure of Empathy) scales and subscales. Comparisons between men and women are provided.

Subjective pain perception

Participants accurately distinguished painful from non-painful scenarios in 79.3% of catch trials (*SD* = 11.6%), confirming the reliability of our experimental design. On average, painful images were rated as moderately painful (M = 6.15, SD = 1.73), whereas neutral images were predominantly rated as not painful (M = 1.70, SD = 0.99). Follow-up t tests confirmed that this difference was significant, $t_{(36)}$ = 16.16, p < .001, Cohen's d = 2.66. Participants also took longer to evaluate painful images than non-painful ones, $t_{(36)}$ = 5.28, p < .001, Cohen's d = 0.87 (average reaction times in Table 4.1). However, as illustrated in Figure 4.2, there was substantial variability in pain ratings, with scores ranging from 1.25 to 8.56 on a 9-point scale.

Figure 4.2. | *Picture ratings in painful and neutral conditions per participant*

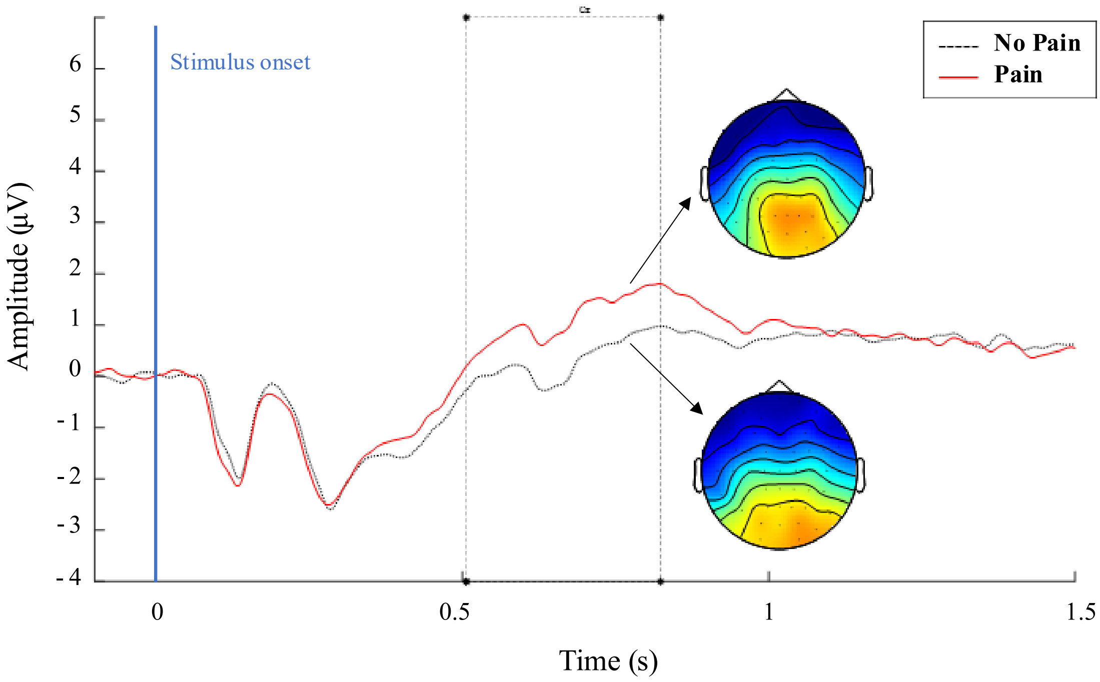


Note. Pain intensity was rated from 1 (no pain) to 9 (extreme pain).

EEG correlates of perceived pain

Pain observation elicited larger LPP amplitude compared to non-painful stimuli, with maximal distribution over central, $t_{(36)}$ = 1.15, p < .001, parietal, $t_{(36)}$ = 2.82, p = .004, and centroparietal, $t_{(36)}$ = 3.10, p = .002, electrode sites (C1, C2, C4, Cz, CP1, CP2, CP3, CP4, CPz, P1, P2, P4, Pz, POz) between 508–918 ms after stimulus onset (see Figure 4.3 for an example). Pain observation was also associated with changes in spectral power within the theta, $t_{(36)}$ = -3.06, p = .004, and alpha, $t_{(36)}$ = -2.68, p = .011, frequency bands – showing maximal distribution over centroparietal electrodes (C5, C3, C1, Cz, C2, C4, C6, TP7, CP5, CP3, CP1, CPz, CP2, CP4, CP6, TP9, T7) between 650–1300 ms after stimulus onset (Figure 4.4). No significant effects were found at the beta band frequency.

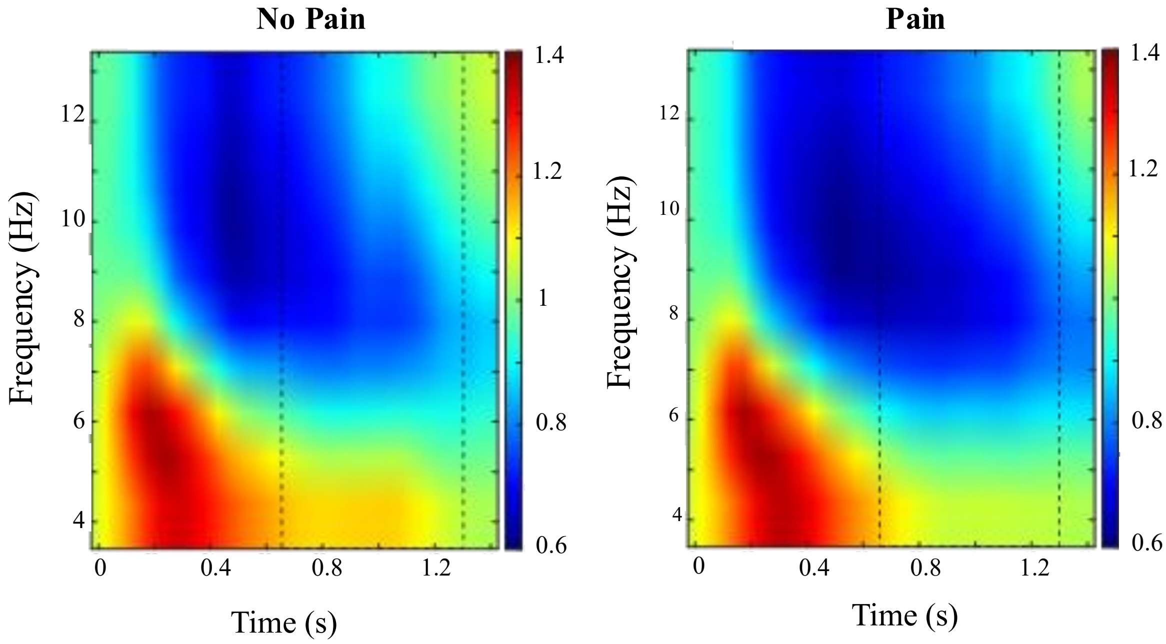
Figure 4.3. | *Grand averaged ERP at Cz for painful vs neutral images*



Note. Significant time range (508–918 ms) is highlighted in the black dotted square**.**

To examine the relationship between pain-related EEG activity and participants' subjective pain ratings, we conducted bivariate correlations in SPSS. Given the relatively small sample size, we used $CI_{95\%}$ to determine statistical significance, as this method provides a more precise estimate of effect sizes and reduces reliance on arbitrary *p* value thresholds (Amrhein et al., 2019). Correlations were considered statistically significant if their $CI_{95\%}$ did not include zero. No significant correlations were found between EEG activity and pain intensity ratings – as indicated by large 95% confidence intervals, all of which included zero (see data in Supplement 4.3).

Figure 4.4. | *Grand averaged ERD per pain condition at the centroparietal region*



Note. The black rectangle represents the time frequency window (650 to 1300 ms) in which changes in spectral power in the alpha and theta frequencies are observed.

Correlations with empathy, callous-unemotional traits, and aggression

To further explore individual differences, we examined correlations between EEG activity and trait measures of empathy, callous-unemotional traits, and aggression – using the ACME, ICU, and RPQ subscale scores (Supplement 4.3). A negative correlation was observed between callousness scores and pain-related ERP amplitude over centroparietal electrodes, r = -0.34, $CI_{95\%}$ [-0.60, -0.02]. Similarly, uncaring behaviour negatively correlated with ERP amplitude changes over both centroparietal, r = -0.36, $CI_{95\%}$ [-0.61, -0.03], and central electrodes, r = -0.34, $CI_{95\%}$ [-0.60, -0.01]. Additionally, unemotional traits positively correlated with increased modulation of theta activity during pain observation compared to no-pain conditions, r = 0.39, $CI_{95\%}$ [0.07, 0.64]. In contrast, pain-related theta modulation negatively correlated with proactive aggression, r = -0.36, $CI_{95\%}$ [-0.61, -0.03]. No significant correlations were found between trait measures and participants' pain intensity ratings ($CI_{95\%}$ including zero).

Classifiers accuracy

Classifier accuracy was evaluated by examining both the variance and the median accuracy in differentiating pain-related EEG data from no-pain data. Comparative analyses revealed that the ERP classifier achieved a median accuracy of approximately 54.18% (*SD* = 6%), whereas the ERD classifier reached a median accuracy of around 53% (*SD* = 10%) (see Supplement 4.4 for an illustration). Despite these differences, neither classifier met the pre-established accuracy threshold of 56.25%, indicating that both models failed to reliably distinguish between pain and no-pain conditions.

Discussion

In this chapter, we investigated the neural correlates of pain observation and the extent to which individual differences in socio-affective traits modulate these responses. Specifically, our study examined how the expression of emotional congruence and understanding corresponded to pain-related EEG fluctuations, building on the assumption that trait empathy influences neural responses to others' pain (Lamm et al., 2011). We additionally explored the potential modulation of these neural responses in relation to callous-unemotional traits – including traits of callousness, unemotionality, and uncaring behaviour –, as well as reactive and proactive forms of aggression. Given the proposed role of callous-unemotional traits in individuals' sensitivity and reactivity to others' pain (e.g., Cheng et al., 2012; Decety et al., 2015), we anticipated a negative relationship between callous-unemotional traits and ERP amplitudes during pain observation. However, we refrained from making specific predictions regarding spectral changes or associations with empathy or aggression due to more inconsistent and limited evidence in the existing literature.

Our results revealed that pain observation was associated with increased LPP amplitudes over centroparietal regions, alongside a delayed decrease in theta and alpha power between 650–1300 ms post-stimulus onset – which replicates findings in previous similar research (Fabi & Leuthold, 2017; Fan & Han, 2008; Mu et al., 2008). Notably, and also consistent with earlier studies (Chen et al., 2023; Fabi & Leuthold, 2018; Perry et al., 2010; Yang et al., 2009), our results showed no significant correlations between self-reported empathy and EEG responses during pain observation. Similarly, there was no correlation between participants' subjective perceptions of pain and pain-related EEG activity. One possible explanation is that the neural processing of pain – as measured by EEG – may not fully capture the empathy facets assessed through self-report measures or state-level indices such as perceived pain intensity (Coll et al., 2017; Zaki & Ochsner, 2012).

In contrast, correlations with callous-unemotional traits and aggression revealed distinct patterns in pain-related EEG fluctuations. Consistent with our hypothesis, participants with more callous and uncaring traits exhibited reduced differentiation in LPP amplitudes between painful and neutral stimuli (Branchadell et al., 2024; Brislin et al., 2022; Decety et al., 2015). Notably, callous and uncaring traits reflect a lack of empathy, disregard for others' emotions, and overall emotional detachment. As such, their association with reduced specificity in pain-related ERPs may provide indirect evidence of a link between empathic predispositions and the neural processing of others' pain. Nevertheless, it is important to note that perceived pain intensity remained consistent across participants regardless of their reported levels of callousness and uncaring behaviours.

On the other hand, unemotional traits – which reflect poor affect and emotional arousal – were associated with greater theta modulation. This finding is particularly relevant given that late theta power desynchronisation – as observed in the present study – has been suggested to indicate enhanced cognitive appraisal of painful stimuli (see Mu et al., 2008). It is thus plausible that, similar to psychopathic cohorts (Lockwood, 2016), individuals with more unemotional traits process others' pain by relying more on cognitive appraisal than affective resonance. Conversely, participants more prone to proactive aggression exhibited reduced theta modulation during pain *vs* no-pain conditions, which may indicate a lower engagement in evaluating observed pain. However, the current evidence is insufficient to make these conclusions as neither callous-unemotional traits nor aggression correlated with overall pain ratings, and effect sizes in the correlations with EEG activity were relatively small.

To further explore the neural signatures of pain observation, we developed machine learning classifiers based on both time-domain (ERPs) and time-frequency (ERDs) features. Despite the growing promise of computational approaches in predicting neural responses during direct pain (Dinh et al., 2019; Mari et al., 2022, 2023a, 2023b), neither classifier in our study achieved above-chance accuracy. This result is consistent with previous research also failing to classify EEG signals related to second-hand pain observation (Mari et al., 2023c, 2025). Several factors likely contribute to these null results. For example, as recently highlighted by Mari *et al* (2025), it is possible that EEG activity to second-hand pain is not sufficiently discernible for machine learning classification, particularly in passive paradigms where explicit engagement is not required. Yet, research also shows that even with active paradigms, classification of second-hand pain does not exceed chance levels (Mari et al., 2025). An alternative explanation for the suboptimal performance of our model could lie in the high inter-individual variability of EEG responses to second-hand pain. This variability, combined with the relatively small sample size and the loss of trials during preprocessing, likely reduced the model's sensitivity and hindered its classification accuracy. It is well-established that larger datasets offer more reliable outcomes in machine learning, as classifier performance improves with increased data volume and variability (Gómez-Tapia et al., 2022; Rommel et al., 2022). Furthermore, classification accuracy may have been hindered by the inherent ambiguity of EEG signals in response to second-hand pain stimuli. Unlike first-hand pain, where EEG patterns tend to be more robust and distinct (Mari et al., 2022), second-hand pain often elicits subtler and less consistent neural responses, making it difficult for classifiers to reliably distinguish pain from no-pain conditions (Mari et al., 2025). In fact, the activation of pain-related neural networks has also been observed in individuals with congenital insensitivity to pain[9] (Salomons et al., 2016), suggesting that such activations may not exclusively reflect true pain experiences. Consistent with this, research has shown that non-painful stimuli may still elicit brain activation patterns similar to those triggered by nociceptive input (Legrain et al., 2011; Mouraux & Iannetti, 2009). Such discrepancies call into question the extent to which EEG signals recorded during second-hand pain observation truly reflect vicarious experiences of pain (see Grice-Jackson et al., 2017).

A recurrent narrative in pain empathy research has been that vicarious responses to observed pain in others is reflective of their capacity for empathy (Coll, 2018; Lamm et al., 2011). However, mixed evidence in the EEG literature indicates that this is a simplistic claim. In fact, as adults we become increasingly reliant on higher-order regions that support reflective and context-sensitive responses to others' emotions (Arain et al., 2013; Blakemore, 2008; Caballero et al., 2016), whereas vicarious pain responses are more reflective of sensorimotor resonance and emotion contagion, which are not sufficient for empathy in adult samples (Valentini, 2010; Zaki & Ochsner, 2012). This position highlights the importance of distinguishing what is necessary from what is sufficient for empathy to occur. The idea is that recognition is necessary but not sufficient for empathy without the accompanying affective and motivational components. Keeping track of these distinctions is essential, particularly when trying to understand empathy's role in behavioural pathologies such as psychopathy, where dysfunctions in empathy are a defining feature. It is important to note, however, that our sample predominantly consisted of individuals with normative empathy levels, which could have limited the sensitivity of the study to detect stronger associations with pain-related EEG fluctuations. In other words, the lack of variability in empathy levels within our sample might have constrained our study's ability to evaluate the extent to which individual differences in empathy affect the neural response to second-hand pain. To address this limitation, it would be valuable to investigate a broader spectrum of empathic variability, for instance by including populations with clinical levels of psychopathy. This approach would help understand how extreme differences in empathy affect neural responses to second-hand pain.

Another significant limitation of our study – and pain empathy research in general – lies in the use of static images to represent pain. While static images are commonly used in pain empathy research due to their experimental control and ease of presentation, they lack the richness of real-life pain experiences. That is, there is no compelling reason to assume that participants would experience empathy for static images of strangers' body parts undergoing simulated pain. The experience of empathy is influenced by situational factors, such as task demands, social context, and individual differences in motivation to engage empathically (see Zaki & Ochsner, 2012 for a discussion). This aligns with broader frameworks proposing that empathy is not driven by a singular neural mechanism but rather by a distributed network in which activation and interactions among brain regions can be context-dependent (Bird & Viding, 2014; Coll et al., 2017). The reliance on static images may therefore underestimate the complexity of pain empathy. To overcome this, future research could benefit from employing more realistic and immersive experimental paradigms, such as video clips or virtual reality. These approaches would better simulate real-world scenarios where individuals experience and react to others' pain, hence leading to more reliable conclusions regarding pain empathy.

Conclusion

This study highlights the challenges of using EEG to investigate the neural correlates of pain empathy. Although EEG responses to second-hand pain are relatively consistent, the failure of machine learning classifiers to reliably differentiate these signals suggests that the underlying neural responses may not be distinct or stable enough. Moreover, the absence of direct correlations with empathy traits indicates that these EEG patterns might not specifically reflect pain empathy, even though associations with callous-unemotional traits imply that predispositions toward empathy could indirectly influence how second-hand pain is processed. Nevertheless, given the small effect sizes and study limitations, these observations should be regarded as preliminary rather than conclusive evidence. Future research should aim to replicate these results with larger and more diverse samples for more reliable conclusions.

Chapter V.

The Promise of New Avenues: On the Possibility to Modulate Callous-Unemotional Traits via Brain Stimulation

Abstract

Psychopathy represents a challenging condition marked by high psychiatric comorbidity, with callous-unemotional traits contributing to symptom aggravation and treatment resistance. In recent years, brain stimulation has been proposed as a potential therapeutic approach in the treatment of psychopathy, yet its effectiveness in specifically targeting callous-unemotional traits remains underexplored. In this chapter, I evaluate the effects of different NIBS protocols on socio-affective processes related to callous-unemotional traits, focusing on empathy, prosociality and guilt as proxies. This investigation includes a systematic review and meta-analysis using multi-level random-effects models. Data from 66 studies (125 effects) were analysed based on stimulation modality (magnetic, electrical) and directionality (excitatory, inhibitory). The results indicate that excitatory protocols improved behavioural outcomes compared to sham/active controls, while inhibitory stimulation led to reductions in these behaviours. However, over 90% of the included studies were conducted in healthy adult samples, limiting direct generalisability to psychopathy. Moreover, sensitivity analyses indicated high between-study heterogeneity and effect estimates (particularly for excitatory stimulation) were influenced by single-study outliers. While these findings provide proof-of-concept for the potential of NIBS to modulate callous-unemotional traits, they also underscore the need for protocol standardisation and the inclusion of clinically-relevant samples to ascertain its therapeutic potential.

"…a prognosis of irreversibility reflects (…) our own therapeutic incompetency and inadequacy rather than an asseveration that change is not possible."

— Jacob Chwast (1961, p.223)

Introduction

Despite extensive research on psychopathy since Cleckley's (1941) and Hare's (Hare & Hart, 1993; Hare & Neumann, 2008) seminal works, effective treatment remains an ongoing challenge for both public health and criminal justice systems (Kiehl & Hoffman, 2011). Historically, attempts to treat psychopathy have resulted in misguided practices like electroconvulsive therapy[10], which not only raised ethical concerns but frequently worsened behavioural symptoms (Green et al., 1944; Hare, 1970). Today, psychological interventions offer more humane and effective alternatives, with therapies such as anger management, cognitive-behavioural therapy (CBT) and group therapies commonly employed to reduce violence and promote empathy in offenders meeting the criteria for psychopathy (David et al., 2018; Rice et al., 1992; Tew et al., 2012). Yet, these intervention efforts remain insufficient, as psychopathic offenders still exhibit high recidivism compared to their non-psychopathic counterparts, even after receiving treatment (Kiehl & Hoffman, 2011; Reidy et al., 2013; Rice et al., 1992; Serin, 1996). Such discouraging outcomes prompt many to label criminal psychopaths as *untreatable*, reinforcing a reliance on long-term incarceration or, in extreme cases, capital punishment (Harris & Rice, 2006). More recent perspectives, however, challenge the notion of psychopathy as an inherently untreatable condition, pointing to methodological limitations in existing research and clinical practices (Hecht et al., 2018; Kiehl & Hoffman, 2014; Reidy et al., 2013). For instance, critics argue that conventional therapies fundamentally fail to address the core characteristics contributing to treatment resistance in psychopathy (D'Silva et al., 2004; Salekin et al., 2010). This perspective stresses the need to put greater emphasis on treatment-resistant traits in order to improve therapeutic efficacy and long-term prognosis for psychopathic cohorts.

In this chapter, I discuss callousness as a key challenge for treating psychopathy (Frick et al., 2014; Muñoz & Frick, 2012). This discussion builds on current literature to give insights into how conventional therapies fall short in addressing callous-unemotional traits and how brain stimulation presents unique opportunities for overcoming these limitations.

The challenges of treating callous-unemotional traits

Empathy enhancement strategies have become a central component in many intervention programs targeting aggressive and antisocial behaviour. For example, treatments for children with conduct disorder often include modules designed to cultivate perspective-taking and compassion (Goldstein et al., 1998; Pecukonis, 1990). Similarly, empathy-based therapeutic approaches are commonly employed in correctional settings to reduce reoffending among violent offenders (Day et al., 2010; Trivedi-Bateman & Crook, 2022). These interventions are grounded in the belief that fostering empathy will promote prosocial behaviour and reduce aggression. However, despite their widespread use, empirical support for their effectiveness remains limited. Research indicates that while some individuals may show modest increases in empathy post-treatment, these changes are often inconsistent and do not reliably predict reductions in recidivism or aggressive behaviour (Day et al., 2010; Trivedi-Bateman & Crook, 2022). This is particularly evident among individuals with callous-unemotional traits, who show greater resistance to therapeutic interventions (Frick & White, 2008; Hawes et al., 2014; Hawes & Dadds, 2007). The challenge lies in the assumption that improving emotional understanding alone can foster genuine empathy and reduce antisocial behaviour. The success of conventional psychotherapies such as CBT depends on therapeutic alliance and patient compliance, both of which imply the establishment of rapport (Felthous, 2011). These therapies emphasise improving emotional understanding as a means to fostering empathy and prosocial behaviour (Chialant et al., 2016; Hecht et al., 2018). However, this approach fundamentally disregards that genuine empathy entails not only recognising and understanding others' emotions but also resonating with them on an affective level (Blair, 2005). In other words, merely improving emotional understanding does not equate to genuine empathy, nor does it guarantee reductions in aggressive or antisocial behaviour (Baron-Cohen & Wheelwright, 2004; Vachon et al., 2014a). This fundamental gap may explain the limited effectiveness of current interventions for the treatment of psychopathic cohorts.

As discussed in previous chapters, individuals who can effectively understand others' emotions but lack affective engagement may become more adept at manipulating others (Harris & Rice, 2006; Kiehl & Hoffman, 2014; Polaschek, 2014; Rice et al., 1992; Seto & Barbaree, 1999). In this regard, therapeutic approaches that prioritise emotional understanding over affective resonance might inadvertently enhance manipulation skills in patients with callous-unemotional traits, as these typically do not struggle with understanding others' emotions but rather with forming genuine emotional connections (Blair, 2013; Cheng et al., 2012). This issue complicates the evaluation of treatment success, which often relies on observable behavioural improvements (Harris & Rice, 2006). Therefore, individuals that can simulate empathy without genuinely experiencing it may give a misleading impression of treatment compliance (Chialant et al., 2016). In forensic settings, this poses a significant risk, as offenders may appear reformed despite retaining core emotional deficits, potentially leading to premature release and continued harmful behaviours. Despite these challenges, psychosocial interventions do show some positive outcomes when implemented early on and with sufficient intensity in children/adolescents (Salekin, 2019). This is likely due to the ongoing development of key socio-affective skills – such as impulse control, emotional regulation, and moral reasoning – during childhood and adolescence, which allows for greater flexibility in shaping behavioural patterns (Reidy et al., 2013; Salekin et al., 2012). In adulthood, by contrast, repeated reinforcement of maladaptive cognitive-affective schemas – such as lack of remorse, and insensitivity to punishment – contributes to the entrenchment of callousness and emotional detachment, which makes older psychopaths especially resistant to treatment (Felthous, 2011; Kiehl & Hoffman, 2011). Consequently, treating adult cohorts with pronounced callous-unemotional traits requires approaches that move beyond superficial behavioural modifications to address the deeply rooted social and emotional impairments that sustain these traits (da Silva et al., 2013; Felthous, 2015; Van Dongen, 2020).

Deficits in socio-affective behaviour correspond with a dysregulation of cortical excitability and inhibition in paralimbic and frontal brain regions (N. E. Anderson & Kiehl, 2012; Blair, 2003b; Blair et al., 2006; Blair, 2013; Kiehl & Hoffman, 2011). Structures in the paralimbic system, such as the amygdala, ACC, and insula, are essential for emotional processing and moral cognition (Blair, 2007). Studies using fMRI have shown that individuals with higher levels of callous-unemotional traits exhibit reduced activity in these paralimbic regions when processing distress cues like fear and sadness (Blair, 2013; Glenn & Raine, 2008; Jones et al., 2009; Kiehl et al., 2001; Marsh et al., 2008, 2013; White et al., 2012). This neural hypoactivity is coupled with diminished emotional reactivity, reflecting affective impairments in processing others' distress (Marsh et al., 2008; Viding & McCrory, 2019). As a result, individuals with pronounced callous-unemotional traits may not develop the normal aversive conditioning that typically deters harmful interpersonal behaviours, increasing their risk of engaging in recurrent and victim-based crimes (Gao et al., 2010; Kahn et al., 2013; Zych et al., 2019). Furthermore, callous-unemotional traits are associated with regional hypoactivity in orbitofrontal and ventromedial prefrontal brain regions (Birbaumer et al., 2005; Blair, 2013; Glenn & Raine, 2008; Kiehl & Hoffman, 2011; Marsh et al., 2011), which play critical roles in moral decision-making and emotional regulation (Kringelbach & Rolls, 2004). In healthy individuals, the activation of the OFC has been linked to the experience of guilt (Wagner et al., 2011), but individuals with callous-unemotional traits show dysfunctional activity in this area, often in tandem with reduced amygdala function (see Blair, 2007 for a discussion). Functional connectivity analyses additionally show that psychopaths exhibit lower connectivity between the amygdala and the OFC (Marsh et al., 2011), which results in a lack of top-down regulation of emotional responses, leading to poor impulse control and diminished sensitivity to social reinforcement (Carré et al., 2013). Similarly, dysfunctions in the VMPFC have been implicated in the expression of callousness, possibly as a secondary consequence of amygdala hypoactivity (Blair, 2013). This suggests that individuals with callous-unemotional traits may struggle to integrate emotional and social information into their decision-making processes, leading to impairments in behavioural flexibility and moral reasoning (Anderson et al., 2013; Koenigs, 2012). Without the ability to incorporate affective feedback, these individuals may continue engaging in antisocial behaviour despite punitive consequences, making rehabilitation more challenging (Salekin et al., 2010).

These findings highlight one of the main limitations of conventional behavioural therapies in the treatment of psychopathy, which is their inability to address the neural underpinnings of its underlying symptoms. Emerging perspectives thus advocate for the implementation of neuroscience-informed approaches to overcome these limitations (Canavero, 2014; Glenn & Raine, 2014; Van Dongen, 2020).

The potential of non-invasive brain stimulation

Neuromodulation interventions offer promising avenues for addressing the neural impairments associated with callous-unemotional traits. One conventional approach to neuromodulation involves pharmacological treatments that target neurotransmitter or endocrine imbalances (Demirtas-Tatlidede et al., 2013). However, these treatments have only been proven effective in alleviating psychiatric comorbidities in psychopathic cohorts, having minimal direct effects on the core symptoms of psychopathy itself (Chialant et al., 2016; Kiehl & Hoffman, 2011; Reidy et al., 2013). Moreover, the use of psychiatric medication often comes with adverse side effects that deter long-term adherence (Romero-Martínez et al., 2020). This has led researchers to explore alternative neuromodulatory strategies for psychiatric treatment, among which NIBS has gained increased popularity within the past decades. NIBS techniques like transcranial magnetic stimulation (TMS) and transcranial electrical stimulation (tES) present a more favourable cost-benefit trade-off than other methods for neuromodulation, as they overcome the adverse effects of medication (Canavero, 2014) and invasive stimulation methods like deep brain stimulation (DBS) (see Mackenzie, 2016 for a discussion). Essentially, NIBS offers a non-invasive means to modulate cortical neuronal excitability and underlying psychological processes by applying changing magnetic fields or electric currents to the skull's surface (Polanía et al., 2018; Stagg et al., 2018).

High-frequency repetitive TMS (HF-rTMS; 5–20 Hz), for example, has been shown to increase cortical excitability and enhance facilitatory neuroplasticity via mechanisms akin to long-term potentiation (Bliss & Cooke, 2011; Monte-Silva et al., 2013), while low-frequency rTMS (LF-rTMS; <5 Hz) reduces cortical excitability (Hallett, 2007). Additionally, rTMS can be applied in specific patterns through what is commonly known as theta-burst stimulation (TBS). In TBS, magnetic pulses are applied in bursts of three at 50 Hz with an inter-burst interval at 5 Hz, corresponding to theta oscillations. TBS trains can be applied intermittently (iTBS) to increase cortical excitability, or continuously (cTBS) to decrease it (Huang et al., 2005). Similarly, tES methods, including transcranial direct current stimulation (tDCS) and transcranial alternating current stimulation (tACS), offer versatile approaches. Electrical stimulation via tACS consists in adjusting cortical excitability by entraining neural oscillations corresponding to specific cognitive functions via oscillatory stimulation, while tDCS involves low-intensity application of direct currents for extended durations (Antal & Paulus, 2013). Depending on the polarity of the electrodes, tDCS can either enhance or reduce neural excitability through anodal (A-tDCS) or cathodal (C-tDCS) stimulation polarity (respectively), similar to TMS effects (Nitsche et al., 2003, 2008; Nitsche & Paulus, 2000). This polarity-dependent effect has been particularly demonstrated in motor cortical areas (Polanía et al., 2018; Stagg et al., 2018), although similar results have been observed in brain regions involved in social cognition. As such, NIBS parameters can be leveraged to investigate and potentially treat callous-unemotional symptoms in psychopathy (Canavero, 2014).

On one hand, inhibitory NIBS methods can model the socio-affective impairments observed in individuals with callous-unemotional traits, aiding in the identification of relevant neural substrates and biomarkers. For instance, C-tDCS applied to the OFC has been shown to reduce feelings of guilt (Karim et al., 2010), highlighting a possible intervention target. Conversely, excitatory NIBS techniques can counteract neural hypoactivity in key brain regions implicated in callous-unemotional traits. A-tDCS applied to the VMPFC, for example, has been found to enhance emotional responsiveness (Nejati et al., 2023), which suggests the potential of excitatory NIBS for improving empathy in individuals with callous-unemotional traits/psychopathy (see Sergiou et al., 2020 for a review). This potential is further supported by recent meta-analyses reporting significant improvements not only in empathic responses but also in prosocial behaviour following A-tDCS (Bahji et al., 2021; Darby & Pascual-Leone, 2017; Smits, Schutter, Van Honk, et al., 2020; B. Yuan et al., 2021), HF-rTMS (Christian & Soutschek, 2022; Darby & Pascual-Leone, 2017; Smits, Schutter, Van Honk, et al., 2020) and iTBS (C.-C. Yang et al., 2018). However, a critical limitation in the current literature is the lack of NIBS studies targeting behavioural outcomes specifically related to callous-unemotional traits. While previous reviews have investigated the impact of NIBS on empathy, most have addressed both the cognitive and affective components of empathy, rather than focusing exclusively on the emotional deficits characteristic of callous-unemotional traits. In fact, to our knowledge, no previous review has systematically investigated the effects of NIBS on guilt-related processes, which are central to the callous-unemotional phenotype. Furthermore, prior reviews typically focus on a single stimulation modality (either rTMS or tDCS), with only two meta-analyses comparing the effects of both methods on morality (Darby

& Pascual-Leone, 2017) or on emotional responses not specific to empathy (Smits, Schutter, van Honk, et al., 2020).

Project rationale

This study seeks to address the gaps in the current literature by evaluating the effects of both TMS and tES protocols on behavioural outcomes specifically related to affective empathy, prosociality, and guilt as proxies for callous-unemotional traits. This approach was theoretically informed by our current understanding of these behaviours as core characteristics defining callous-unemotional symptoms in psychopathy, further supported by a recent meta-analysis indicating their high correlation with the expression of callous-unemotional traits (see Waller et al., 2020). Given that callous-unemotional traits are associated with neural hypoactivation in multiple brain regions, we explored the effectiveness of different stimulation parameters (e.g., frequency, current polarity, intensity) through whole-brain analyses. We conducted separate analyses for high- and low-frequency rTMS as well as anodal and cathodal tDCS, hypothesising that excitatory protocols would lead to significant improvements in the targeted outcomes, whereas inhibitory protocols would attenuate the socio-affective response.

Study 6

Method

This research followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009; Page et al., 2021), with materials and protocol registered on OSF (https://doi.org/10.17605/OSF.IO/TH3S9).

Search strategy and study selection

We conducted an electronic search through November 2024 using Scopus, PubMed, and Web of Science. Our search strategy incorporated terms related to both magnetic and electrical forms of NIBS, including TMS, TBS, tDCS and tACS techniques. These were paired with terms associated with callous-unemotional traits, such as “callous”, “empathy”, “guilt”, and “prosocial”, as well as additional keywords like “emotional reactivity” or “cooperation” (see Supplement 5.1 for detailed search terms). Additional studies were identified through reference lists and forward citations (see PRISMA flowchart in Figure 5.1).

Study selection was limited to randomised controlled trials with human participants, excluding systematic reviews, case studies, and editorials. All articles were in English, with no limitations regarding publication status or year of publication. Out of 256 records (including grey literature), 66 studies met our inclusion criteria (Table 5.1). Data screening was carried out independently by multiple reviewers through Excel spreadsheets. Titles and abstracts were initially screened by three study authors, reaching an interrater agreement of 95%. Other two study authors subsequently reviewed the full texts of the selected articles, reaching an interrater agreement of 92.5%. Any disagreements were resolved through consensus among the reviewers.

Table 5.1. | *Eligibility criteria*

PICO Component	Criteria
P: Population	Target sample included adults aged 18 to 64 with no major neurological conditions. Only studies involving clinical samples that adequately controlled for the use of psychotropic medications were considered.
I: Intervention	Studies employed NIBS interventions, including TMS protocols (single pulse TMS, rTMS, cTBS, and iTBS) and/or tES protocols (tDCS, tACS, HD-tDCS, HD-tACS). Studies with multiple stimulation methods were only included if they applied each method separately.
C: Comparison	Studies included a control condition (whether within- or between-subjects), either through sham stimulation or an active stimulation site for comparison.
O: Outcome	Primary outcomes were assessed through behavioural tasks and/or self-report measures designed to measure callous-unemotional traits, affective empathy, prosociality and guilt. We only included studies in which behavioural data was collected during or after stimulation.

Note. The study's eligibility criteria followed the PICO (Population-Intervention-Comparison-Outcome) framework.

Data extraction

Extracted information included: a) study details (study ID, first author and year of publication, country where experiment was conducted); b) participant demographics (sample type, sample size, gender distribution, age); c) intervention details (experimental design, ROI, stimulation paradigm, intensity, duration, number of sessions, control method); d) behavioural targets and employed outcome measures (described in Supplement 5.2); and e) statistical data (number of participants per condition, and mean scores with corresponding standard deviations on behavioural outcomes in each condition).

For simplicity, we labelled all studies using electrical stimulation as 'tDCS', given that tACS involves constantly changing the direction of the current. If studies assessed identical outcomes using different stimulation protocols or various stimulation sites we treated each site trial as a separate unit of analysis. To mitigate the risk of bias and prevent double counting of outcomes, we averaged the effects for studies reporting multiple outcomes of the same measure. When studies incorporated various control conditions, we prioritised data extraction from sham control conditions, as these provide a reliable baseline for evaluating the true effects of stimulation. When numerical data were only presented graphically, we used Plot Digitizer software (plotdigitizer.sourceforge.net) to extract numerical values and corresponding standard deviations. In most studies, higher scores denoted greater levels of empathy, prosociality, or guilt. For scales with inverse trends, we adjusted mean scores by subtracting the group mean values from the maximum scale score, ensuring consistency across study observations (Smits, Schutter, Van Honk, et al., 2020). When final scores were not provided, we used change-from-baseline scores as a proxy, which are theoretically comparable to final scores in randomised controlled studies (Higgins & Green, 2011). Refer to Supplement 5.3 for additional details. Any discrepancies or missing information were addressed by contacting the study authors.

Statistical analysis

Studies were categorised by stimulation modality (magnetic *vs* electrical) and protocol (excitatory *vs* inhibitory) for separate multilevel random-effects meta-analyses, conducted using the metafor package in *R* (Viechtbauer, 2010). Intervention effects on behavioural targets were evaluated by calculating weighted standardised mean difference (Hedge's *g*) (Durlak, 2009). Effects were interpreted using $CI_{95\%}$ and *p* values for statistical significance (setting the significance threshold at $p < 0.05$). Significant effects were supplemented by sensitivity analyses. These included leave-one-out methods to determine the influence of individual studies (Willis & Riley, 2017), alongside influence diagnostics and Baujat plots to detect potential outliers (Wang, 2023). Additionally, we conducted subgroup analyses to test the moderating effects of stimulation parameters – including stimulation site, intensity, duration, paradigm (online/offline), and number of sessions. Main effects interpretation and subgroup analyses excluded identified outliers for more reliable conclusion. Furthermore, we assessed risks of publication bias and between study heterogeneity, with comparisons before and after outlier exclusion. Publication bias was assessed through funnel plots and Kendall's tau ($r\tau$) rank-order correlations (Begg & Mazumdar, 1994). Between-study heterogeneity was tested via prediction intervals, τ^2, and I^2 statistics (Borenstein, 2023; Higgins & Thompson, 2002).

Results

Study characteristics

This review covers research published between 2009 and 2024, including 21 studies using TMS and 45 using tDCS. Most studies were conducted on healthy participants with no reported criminal records, with only three TMS studies (de Wit et al., 2015; Jansen et al., 2019; Light et al., 2019) and two tDCS studies (Lisoni et al., 2024; Sergiou et al., 2022) – one of which included forensic inpatients (Sergiou et al., 2022). Regions of interest (ROIs) were primarily located within the PFC. Excitatory effects were induced via either HF-rTMS at 5–10 Hz (k = 9) or A-tDCS at 1–2 mA (k = 28), whereas inhibitory or excitatory-diminishing effects were achieved with LF-rTMS at 1 Hz (k = 9), cTBS (k = 6) or C-tDCS at 1–2 mA (k = 15). Some studies applied simultaneous anodal and cathodal stimulation over ROIs, using HD-tDCS (k = 6) or bilateral bipolar tDCS (k = 11). Because the precise direction of effects (excitatory versus inhibitory) is less clearly defined for these protocols, we used their results for exploratory analyses (Supplement 5.4). Additional study details are summarised in Supplement 5.5 and 5.6.

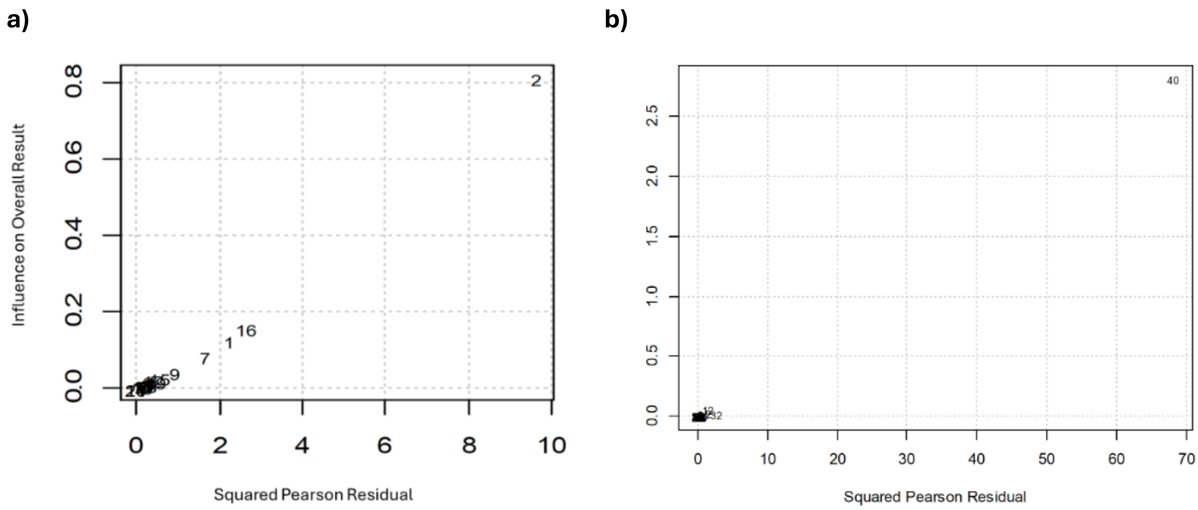
Effects of excitatory stimulation

Meta-analyses showed significant improvements in the socio-affective response following HF-rTMS (n = 21) and A-tDCS (n = 44). However, sensitivity analyses indicated that the pooled effect of HF-rTMS was primarily influenced by a single study (see Baujat plot in Figure 5.2a), with study removal reducing the effect estimate from g = 0.58, $CI_{95\%}$ [0.08, 1.08], p = 0.024 to g = 0.40, $CI_{95\%}$ [0.02, 0.79], p = .042. Similarly, for A-tDCS, the analyses suggested a potential overestimation due to the influence of one study (Figure 5.2b). Exclusion of this study reduced the effect estimate from $g = 0.56$, $CI_{95\%}$ [0.03, 1.10], $p = .040$ to $g = 0.32$, $CI_{95\%}$ [0.05, 0.60], $p = .022$. Forest plots in Supplement 5.7 illustrate the consistency of HF-rTMS and A-tDCS effects after these adjustments, with additional details provided in Supplement 5.8.

Effects of inhibitory stimulation

Inhibitory protocols showed more mixed results. On one hand, LF-rTMS (21 effects) and cTBS (15 effects) showed non-significant effects, with pooled effect estimates of $g = -0.89$, $CI_{95\%}$ [-1.96, 0.18], $p = .104$ and of $g = 0.17$, $CI_{95\%}$ [-0.13, 0.47], $p = .264$, respectively. In contrast, C-tDCS (24 effect sizes) was associated with a significant reduction in socio-affective responses ($g = -0.45$, $CI_{95\%}$ [-0.81, -0.09], $p = .014$). The forest plot in Supplement 5.7 illustrates the consistency of these effects. Sensitivity analyses further confirmed their robustness, showing minimal variation in effect estimates and significance upon the removal of individual studies (see Supplement 5.8 for specific values).

Figure 5.2. | *Baujat plots*



Note. Baujat plots display the influence of individual trials on the overall effect sizes in studies using high-frequency rTMS (Panel A) and anodal tDCS (Panel B). In the TMS analysis, Study 2 (Balconi & Canavesio, 2014) notably drives the overall effects. Likewise, for tDCS trials, the plot highlights that Study 40 (Yuan et al., 2017) significantly impacts the aggregated findings.

Subgroup analyses

After excluding identified outliers, subgroup analyses revealed larger effect estimates in HF-rTMS (g = 0.56, $CI_{95\%}$ [0.05, 1.07], p = .030), A- tDCS (g = 0.44, $CI_{95\%}$ [0.07, 0.80], p = .018) and C-tDCS (g = -0.50, $CI_{95\%}$ [-0.92, -0.09], p = .016) RCTs. Further analysis indicated greater excitatory effects following bilateral HF-rTMS applied to the PFC (g = 1.22, $CI_{95\%}$ [0.53, 1.90], p < 0.001), A-tDCS to the middle PFC (g = 1.06, $CI_{95\%}$ [0.37, 1.75], p = .002), and C-tDCS to the left PFC (g = -2.32, $CI_{95\%}$ [-4.23, -0.41], p = .017). Additionally, a moderation effect was observed with the number of stimulation sessions in LF-rTMS. Notably, this moderation increased the effect estimate to significance (g = -0.62, $CI_{95\%}$ [-1.19, -0.06], p = .031).

Between-study heterogeneity and publication bias

Heterogeneity assessments revealed large prediction intervals regardless of outlier exclusion, indicating high variability in effect sizes across studies (see Table 5.3). Additionally, funnel plot asymmetries (Supplement 5.9) suggest potential publication bias among studies involving HF-rTMS ($r\tau$ = 0.36; p = .022), LF-rTMS ($r\tau$ = -0.42; p = .007), A-tDCS ($r\tau$ = 0.26; p = .011), and C-tDCS ($r\tau$ = -0.38; p = .008).

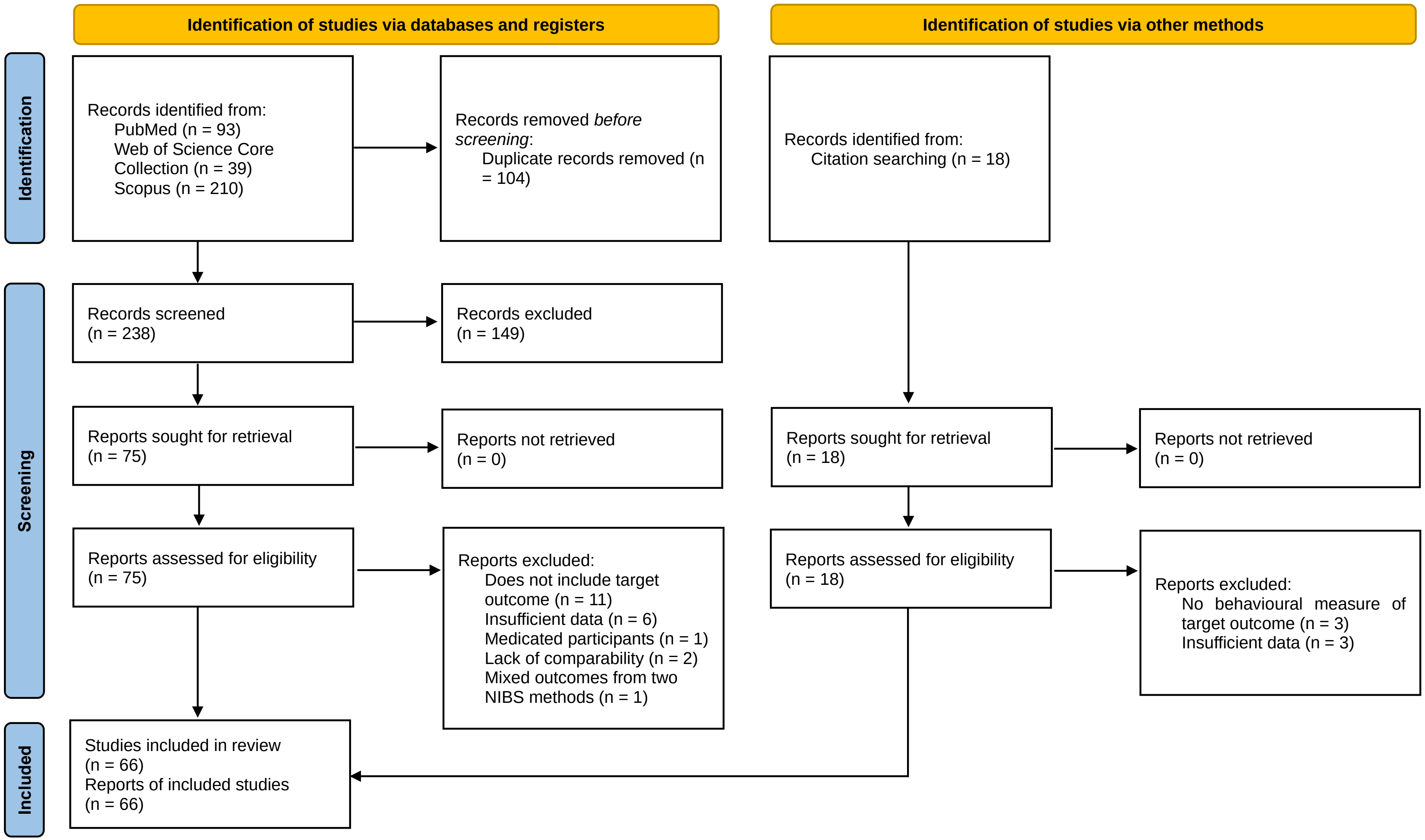
Table 5.3. | *Heterogeneity assessment*

	PI	τ^2	I^2
High-frequency rTMS	-1.68, 2.84	1.27	93.36%
High-frequency rTMS$_a$	-1.25, 2.95	0.67	88.51%
Low-frequency rTMS	-5.83, 4.06	6.07	97.65%
Continuous TBS	-0.85, 0.97	0.19	69.58%
Anodal tDCS	-2.97, 4.09	3.16	97.69%
Anodal tDCS$_a$	-1.42, 2.07	0.77	91.39%
Cathodal tDCS	-2.17, 1.26	0.73	90.68%

Note. PI : Prediction Interval; τ^2 : tau-squared statistic.

$_a$ Analysis after outlier exclusion.

Figure 5.1. | *PRISMA Flowchart*



Discussion

This study presents the first meta-analysis to systematically evaluate the effects of NIBS on behavioural outcomes specifically linked to callous-unemotional traits, with a focus on affective empathy, prosociality, and guilt. These socio-affective processes were selected based on their established relevance to the presentation of callous-unemotional traits, as documented in both diagnostic frameworks and empirical research (Hare, 2006; Hare & Neumann, 2008; Waller et al., 2020). Our hypotheses were grounded in evidence indicating neurofunctional impairments in individuals with high callous-unemotional traits (Blair, 2013; Kiehl & Hoffman, 2011), alongside evidence that excitatory and inhibitory NIBS protocols modulate cortical excitability in predictable patterns (Polanía et al., 2018; Stagg et al., 2018). We hypothesised that excitatory stimulation would lead to improvements in the targeted outcomes and that inhibitory stimulation would attenuate them. Our findings offer preliminary support for these hypotheses. Inhibitory stimulation, particularly via C-tDCS, resulted in small but statistically significant reductions in the targeted outcomes. Conversely, excitatory stimulation through HF-rTMS and A-tDCS was associated with modest improvements. Nevertheless, sensitivity analyses suggested the potential overestimation of effect sizes in both HF-rTMS and A-tDCS meta-analyses through outliers. Interestingly, the identified outlier studies either assessed participants' emotional responses to witnessing aggression (H. Yuan et al., 2017) or their predispositions to intervene in such scenarios (Balconi & Canavesio, 2014). This could suggest that excitatory effects on the socio-affective response (in healthy individuals) may be relative to the emotional saliency of the presented stimuli, which is exacerbated in explicit emotion tasks such as those depicting interpersonal conflict (J. Yuan et al., 2019).

Subgroup analyses removing outliers indicated that effect sizes for HF-rTMS, A-tDCS, and C-tDCS were larger in studies using a between-subjects design (RCTs). While this may suggest that NIBS effects are more pronounced when baseline neural activity remains unaffected by prior stimulation, it is equally plausible that such designs inflate effect sizes due to uncontrolled variability in individual differences (e.g., baseline cortical excitability), which crossover designs are better suited to control (Sedgwick, 2014). Additionally, we found larger effect sizes in studies targeting the PFC. However, it should be noted that callous-unemotional traits are not the result of dysfunction in a single brain region but rather emerge from disrupted connectivity within broader neural networks (Blair, 2013; Carré et al., 2013; Kiehl & Hoffman, 2011; Marsh et al., 2011). Therefore, approaches that move beyond focal stimulation may be more appropriate for addressing the neural disruptions associated with callous-unemotional traits. One promising direction could be using techniques that explicitly target functional connectivity between brain regions, such as cortico-cortical paired associative stimulation – commonly known as ccPAS –, which offers the possibility to strengthen or weaken network-level interactions by inducing temporally coordinated plasticity between interconnected cortical sites (J. J. Zhang, 2024). Furthermore, although the overall effects of LF-rTMS were not statistically significant, subgroup analyses revealed that treatment outcomes were moderated by the number of stimulation sessions. Specifically, an increased number of LF-rTMS sessions was associated with stronger inhibitory effects, suggesting a dose-response relationship. This finding reinforces the importance of treatment duration in optimising therapeutic efficacy, and aligns with prior research indicating that repeated NIBS sessions can enhance neural plasticity through cumulative effects on synaptic modulation (Stagg et al., 2018). However, given the non-significant main effects for LF-rTMS, caution is warranted in interpreting this moderation effect, as it may instead reflect variability in study methodologies or potential publication bias.

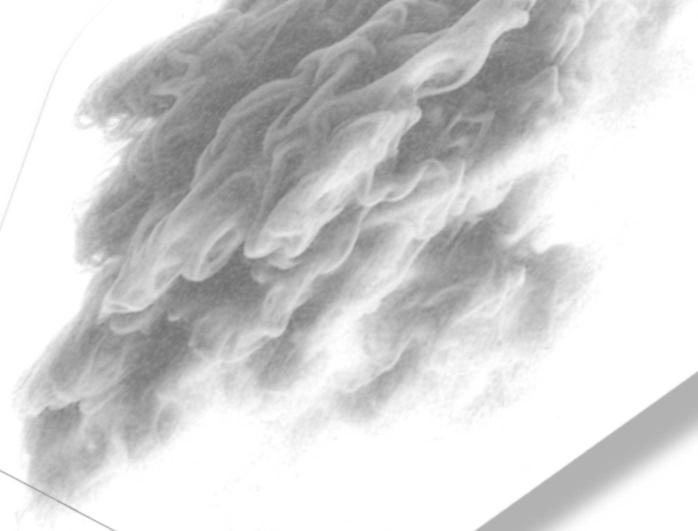
Despite the observed effects, several limitations constrain the interpretation and generalisability of this meta-analysis. One key challenge stems from the considerable variability in study designs, stimulation parameters, and targeted brain regions. For instance, while the PFC emerged as a key target, the studies included in our analysis stimulated different PFC subregions – each of which plays a distinct role in the expression of callous-unemotional traits (Blair, 2013; Glenn & Raine, 2008; Kiehl & Hoffman, 2011). This issue is further compounded by the differing spatial precision of NIBS techniques. For example, tDCS lacks spatial resolution but offers the possibility to target brain regions within a broader neural networks such as those involved in empathy (Bernhardt & Singer, 2012; Engen & Singer, 2013; Fan et al., 2011). In contrast, rTMS provides greater focality but may yield inconsistent effects due to variations in coil placement – particularly when applied online (Walsh & Cowey, 2000). These methodological differences challenge the identification of optimal stimulation parameters.

Another significant limitation is the limited generalisability of our findings to individuals with clinically elevated callous-unemotional traits. Given that individuals with high callous-unemotional traits exhibit distinct neurobiological features (Blair, 2013), it is plausible that their reactions to NIBS differ from those of non-clinical populations. For example, previous evidence has shown that excitatory NIBS renders greater improvements in empathic responses among individuals with lower levels of baseline empathy as compared to high-empathic individuals (Peled-Avron et al., 2019). This raises the possibility that callous-unemotional individuals – who typically have reduced affective empathy – may be more responsive to excitatory NIBS protocols, regardless of the emotional saliency of the task. However, this hypothesis remains speculative as most studies included in this meta-analysis were conducted in healthy samples. In fact, the only available study we identified in which excitatory NIBS (specifically A-tDCS) was applied to a psychopathic cohort found no significant effects (Sergiou et al., 2022). Furthermore, the generalisability of our results to callous-unemotional traits is also constrained by the lack of specificity in outcome measures. That is, although the targeted outcomes (affective empathy, prosociality, guilt) are theoretically and empirically linked to callous-unemotional traits (Waller et al., 2020), the measures used to assess these behaviours are not direct measures of callous-unemotional traits. As such, their use as outcome measures inevitably adds variability to our data.

Despite these limitations, our findings provide the first meta-analytic evidence that NIBS can influence socio-affective behaviours conceptually relevant to callous-unemotional traits – at least in healthy adult populations. Although these results do not establish clinical efficacy, they provide preliminary proof-of-concept useful for informing future research, especially in the design of studies aimed at outcomes relevant in the expression of callous-unemotional traits. This is particularly relevant given the lack of effective treatments for individuals high in callous-unemotional traits (e.g., criminal psychopaths), for whom traditional psychotherapeutic approaches often fail. If similar effects can be demonstrated in these cohorts, it would provide a clearer path into the feasibility of leveraging NIBS as a therapeutic tool for psychopathy. Our study specifically highlights that the exploration of this possibility is significantly constrained by the scarcity of studies directly addressing this issue. Moreover, our review underscores the need for greater protocol standardisation in NIBS research addressing socio-affective processes. Despite between-study heterogeneity, we identified consistent trends that provide a starting point for refining stimulation parameters in future research. Additionally, while we focused on behavioural outcomes in this meta-analysis, we argue that future studies would benefit from the inclusion of neurophysiological measures to assess changes in brain activity before and after stimulation. Such data would help clarify the mechanisms through which NIBS influences socio-affective behaviour and determine whether observed behavioural changes correspond to alterations in cortical excitability or connectivity. Incorporating neuroimaging methods such as fMRI, for instance, could help map the regional and network-level effects of NIBS, hence facilitating the identification of relevant ROIs (Polanía et al., 2018).

Conclusion

The treatability of callous-unemotional symptoms in psychopathy remains a pressing and unresolved challenge. In this meta-analysis, we provide preliminary evidence that NIBS can modulate behaviours relevant to the expression of callous-unemotional traits. Nonetheless, current evidence is predominantly derived from healthy samples, and methodological approaches vary substantially across studies, hence limiting the scope of our conclusions. Despite these limitations, our study represents the first attempt at examining meta-analytic evidence on the possibility to alter behaviours underlying callous-unemotional traits via neuromodulation, offering a critical foundation for future research. As such, this study contributes to a growing discourse on innovative interventions for psychopathy, emphasising the exploration of NIBS methods as relevant targets for further research.



Chapter VI.

Broad Discussion and Concluding Remarks

"Lost within a man who murdered, there was a soul like any other soul."

– William Trevor in 'Felicia's Journey' (1994, p.212)

Broad summary and discussion

The capacity for cruelty is not exclusive to individuals seen as malicious or evil. In fact, there are numerous factors that can facilitate cruel behaviour even among ordinary people – such as their social context, upbringing, or psychological predispositions, just to name a few (Baron-Cohen, 2011). Psychopathy has been a focal point in studying such predispositions, as those with this condition are notorious for their capacity for cruelty (Hare & Neumann, 2008). Importantly, research has demonstrated that psychopathic traits – particularly those involving emotional shallowness and interpersonal coldness, collectively referred to as callous-unemotional – are not restricted to clinical or criminal populations. Instead, these traits are found to varying degrees within the general population, suggesting a dimensional rather than categorical distribution (Edens et al., 2006; Patrick et al., 2009; Skeem et al., 2003). Building on this perspective, the present thesis investigated how callous-unemotional traits manifest within normative samples and how they relate to behavioural patterns typically associated with psychopathy, with a particular emphasis on empathy and aggression. Furthermore, the thesis explores the cognitive and neural mechanisms that may underlie these behavioural tendencies, seeking to identify potential targets and avenues for intervention. In this chapter, I synthesise and discuss the main findings from this body of work, critically evaluate its limitations, and provide recommendations for future research.

Main findings

This thesis comprises four empirical investigations that together explore the relationship between callous-unemotional traits, empathy, and aggression within normative populations. Collectively, these studies provide insight into the behavioural, neurophysiological, and cognitive mechanisms associated with subclinical expressions of psychopathic traits.

Chapter II evaluated the role of different facets of empathy (including affective and cognitive subcomponents) on the link between aggression and the expression of callous traits in normative adult samples. Results indicated that, at higher levels of interpersonal callousness, participants exhibited a tendency towards proactive aggression. Notably, this association was mediated by affective dissonance, suggesting that emotional discord (rather than a mere absence of empathy) might be key in understanding psychopathy-related aggression.

Chapter III examines the dual role of empathy in punitive attitudes. Participants who reported higher concern for victims tended to advocate for harsher punishments, especially when the aggression was intentional. However, this effect diminished among individuals with more callous-unemotional traits, who not only expressed less concern for victims but were also more inclined to endorse harsh punishments regardless of perceived intent. This suggests that callous-unemotional traits may promote a decoupling between emotional response and moral decision-making.

Chapter IV investigated the EEG correlates of empathy in the context of perceiving physical pain in others. Replicating previous research, the study found that second-hand pain perception elicited distinct neural patterns but failed to find significant correlations with participants' empathy levels. Instead, these neural responses seem to vary relative to the expression of callous-unemotional traits, suggesting that these traits can better predict individual differences in pain processing than empathy alone at subclinical levels.

Finally, ***Chapter V*** presented a meta-analytic review of NIBS studies targeting behaviours associated with callous-unemotional traits – particularly affective empathy, prosociality, and guilt. The analysis revealed that excitatory stimulation of brain regions involved in socio-cognitive processing shows potential for improving these behaviours, whereas inhibitory stimulation to the same brain regions had the opposite effects. The significance of these effects, however, was constrained by the high degree of methodological variability among the studies analysed and thus are considered preliminary.

Synthesis and theoretical implications

The interest in psychopathy research primarily lies in the understanding of psychopaths as a notorious minority known for posing significant threats to both individuals and society, largely due to their capacity to recurrently engage in severe aggressive behaviours like predatory violence (e.g., Camp et al., 2013; Meloy et al., 2018). Nevertheless, current perspectives of psychopathy as a spectrum have motivated research seeking to examine these patterns in non-criminal and non-clinical cohorts. The findings presented in this thesis add to this research. In this section, I offer a more in-depth discussion on the study results and explore their respective theoretical implications.

The risks underlying psychopathy manifest subclinically

The risks and challenges in psychopathy can be better understood through the struggle of psychopathic individuals to resonate with and care about others' suffering, which is believed to facilitate interpersonal harm by weakening the natural aversive response to causing distress (De Ridder et al., 2016; Lovett & Sheffield, 2007; Lui et al., 2016; White et al., 2015). Our research reveal that these risks are similarly manifested in subclinical populations. For example, our investigation in ***Chapter II*** emphasises the role of interpersonal callousness in predicting proactive aggression in community (non-criminal) samples. Notably, this form of aggression has been recognised as a potential marker of overall aggression severity due to its intentional and goal-oriented nature, as well as its frequent co-occurrence with reactive aggression (Brugman et al., 2017; Euler et al., 2017). The predictive relationship between interpersonal callousness and proactive aggression thus underscores the utility of assessing callous-unemotional traits in identifying individuals at risk for engaging in more severe forms of aggression (Camara et al., 2025).

Findings in ***Chapter IV*** further reinforce the idea that these risks might be indeed associated with reduced responsiveness to others' painful experiences. For example, the results showed that callous-unemotional traits were associated with delayed desynchronisation in theta-power oscillations during the observation of others in pain. Previous work (Mu et al., 2008) has interpreted such delayed responses as indicative of increased reliance on cognitive appraisal rather than automatic emotional resonance. This pattern aligns with the emotional detachment seen in individuals with high callous-unemotional traits, who may intellectually register distressing stimuli without genuinely feeling concern for the other. Furthermore, in ***Chapter III***, punishment severity among participants with more callous-unemotional traits was less influenced by their empathy for victims, indicating a punitive stance detached from concern for others' suffering. This finding is also consistent with the understanding that diminished moral sensitivity at higher levels of callousness may predispose individuals exhibiting these traits to engage in more deliberately aggressive behaviours (e.g., Gini et al., 2014; White et al., 2015). This is further supported by our results in ***Chapter II***, where moral disengagement positively correlated with proactive aggression among participants with higher interpersonal callousness.

The role of affective processing in behaviour is nuanced

While the problems underlying psychopathy and callous-unemotional traits are most often linked to deficits in affective processing, research suggests that this relationship is far more nuanced than originally anticipated. A compelling example is our finding that affective dissonance is a more powerful predictor of callous aggression than other aspects of empathy in community samples. This finding suggests that the emotional detachment typical of callous-unemotional traits does not simply reflect an absence of empathy, but rather a more complex disruption in how emotional information is processed and regulated. In particular, affective dissonance may reflect a form of emotional interference or conflict that facilitate aggressive behaviour. This is consistent with the idea that affective dissonance might explain key motivational aspects underlying aggressive behaviour in psychopathic cohorts (Dryburgh & Vachon, 2019; Levitan & Vachon, 2021; Vachon & Lynam, 2016), and further points to the need for a more differentiated understanding of empathy-related constructs in psychopathy research.

On the other hand, it is also important to emphasise that functional empathy does not equate the lack of aggression. Rather, empathy is often selectively applied, and its influence on behaviour is shaped by contextual and interpersonal factors. This selective application is particularly evident in the findings from ***Chapter III***, which illustrate how empathy can sometimes be directed toward certain individuals (i.e., victims of aggression) while being withheld from others (i.e., aggressors), leading to biases in judgments related to punishment. Specifically, our study revealed that participants advocated more for harsher punishments for perpetrators that were perceived more negatively, as indicated by their ratings of the perpetrator's meanness. These insights are crucial as the public's perceptions of the rehabilitation potential of offenders can influence the development of intervention efforts for such. In fact, criminal psychopaths are a prime example of this issue, as prevailing pessimistic views on their treatability often discourage therapeutic efforts for these cohorts (Harris & Rice, 2006; Rice et al., 1992). Advances in neuroscience, however, challenge the narrative that psychopathy cannot be treated. Studies in this area suggest that callousness in psychopathy is likely driven by abnormalities in the brain networks responsible for socio-affective processing and further highlight the potential of modifying these patterns via NIBS (see Canavero, 2014 for a discussion). While still in early stages and far from conclusive, our findings in ***Chapter***

V partially support the potential of NIBS techniques in addressing the affective deficits underlying callousness in psychopathy and encourages further inquiry into this possibility.

Taken together, these findings underscore the need for a multidimensional approach to psychopathy. Given its complex interplay of emotional, cognitive, and behavioural components, psychopathy should not be studied – or treated – through a singular lens. Integrative models that account for neural, psychological, and contextual factors are essential not only for advancing theoretical understanding but also for informing more effective, compassionate, and potentially transformative interventions.

Research limitations and future directions

Despite the theoretical significance of the research presented in this thesis, the described studies carry important limitations that need to be acknowledged when interpreting our results. Specifically, this research faces two primary limitations related to sample variability, and the relationships among our targeted constructs. In this section, I delve into how these factors may limit the scope of our study results, offering potential avenues for future investigation.

Limitations of comparability across the psychopathy spectrum

This thesis was motivated by the question of whether approaching psychopathy as a spectrum can provide useful insights into the role of empathy in aggressive behaviour. However, the studies here presented primarily focus on one end of the spectrum – i.e., university or community samples – and as such conclusions on symptom dimensionality are limited. The inclusion of a clinical sample would enable comparisons with healthy populations, thereby enriching our understanding of psychopathy as a spectrum. Comparing individuals diagnosed with psychopathy to healthy controls allows for a more accurate examination of the extent and nature of differences in socio-affective processing between these groups. By identifying specific deficits in empathy, emotional responsiveness, and other relevant domains in psychopathic individuals compared to their non-psychopathic counterparts, we can better delineate where psychopathy falls on the spectrum of socio-affective functioning. Moreover, comparing psychopathic individuals to healthy controls can also help identify potential protective factors or compensatory mechanisms that may exist within this spectrum. Understanding such nuances can inform interventions and treatment strategies tailored to the specific needs of individuals at different points along the psychopathy spectrum. This is particularly pertinent for the studies discussed in ***Chapters IV*** and ***V***, as individual differences in neural activity associated with psychopathy are more likely to manifest among individuals exhibiting a wide range of psychopathic traits. For instance, follow-up research for Study 5 could investigate the correlations between EEG responses to pain and callous-unemotional traits by including a psychiatric sample that meets clinical criteria for psychopathy.

Revisiting the empathy-aggression link

The research combined insights from social psychology and cognitive neuroscience to explore the complex relationship between empathy and aggression. However, these studies are limited in their ability to draw definitive conclusions about empathy's role in driving aggression. For example, the premise for the research presented in ***Chapter V*** was that it might be feasible to improve psychopathy prognosis by modulating its underlying symptoms via NIBS, focusing on affective empathy, prosociality and guilt-related emotions as key targets. However, this is based on the assumption that these socio-affective functions are directly responsible for the behavioural problems manifested in psychopathic cohorts, when in fact research shows that variations in their expression may also serve adaptive purposes (Benning et al., 2018). Thus, a more nuanced understanding of how these socio-affective traits operate within different contexts is necessary. In fact, one important consideration in the empathy-aggression debate is that context matters. ***Chapter III***, for instance, shows the critical role of moral contexts in understanding both the expression of empathy and its potential impact on aggressive behaviour.

Therefore, integrating contextual information into the examination of empathy and aggression in relation to psychopathic traits could yield more valuable insights, not only into the connections between empathy and aggression but also into the understanding of how this relationship is expressed across the psychopathy spectrum. In fact, our EEG study (***Chapter IV***) lacked such contextual information, thereby limiting insights into how processing others' pain may truly evoke empathic responses. Similarly, while ***Chapter II*** revealed important associations between callousness, empathy, and aggression, these findings were derived from studies that also lacked sufficient attention to contextual nuance, thereby limiting their explanatory depth. To overcome this limitation, follow-up research could integrate more ecologically valid experimental paradigms. For example, incorporating virtual reality or interactive social games would better capture the complexities of real-world social interactions. These methods would not only allow for more accurate identification of empathy-related brain activity but also offer deeper insight into the specific situations in which empathy may aggression.

Another promising direction for future research is to closely examine the neural patterns associated with affective dissonance in relation to aggressive behaviour in psychopathic cohorts. Affective dissonance – characterised by the experience of conflicting emotional responses, such as deriving pleasure from another's pain – is a central component of sadism, a trait that has been identified as a significant predictor of aggression in psychopathy (Kirsch & Becker, 2007; Meloy, 1997). This fundamental resemblance suggests that the neural mechanisms underlying affective dissonance may overlap with those driving sadistic tendencies, thus providing a potential neurobiological link to aggressive behaviour. Future investigations could employ advanced neuroimaging and electrophysiological techniques to elucidate whether the same brain networks that regulate affective dissonance are also implicated in the manifestation of sadistic aggression, thereby clarifying their respective roles and interactions. Such research would not only further our understanding of the neurobiological underpinnings of psychopathy but also inform the development of targeted interventions aimed at mitigating aggressive behaviours. Although exploring these relationships was beyond the scope of the current thesis, it represents a promising avenue for future studies that could significantly enhance our understanding of the complex interplay between empathy and aggression along the spectrum of psychopathy.

Conclusion

Understanding the humanity within individuals who commit extreme acts of violence, such as those labelled as *criminal psychopaths*, prompts the exploration of the underlying factors and challenges inherent in such behaviours. When discussing psychopathy, it is crucial to move beyond mere labels to understand the experiences and psychological makeup of individuals with this condition. Rather than viewing psychopathy solely through the lens of pathology or criminality, we can explore the complex interplay of biological, psychological, and environmental factors that shape behaviour at different levels of psychopathy. Nevertheless, advancements in intervention are not just hindered by methodological shortcomings but also by societal perceptions that often portray psychopaths as irredeemable. These perceptions discourage exploration of innovative treatment modalities such as brain stimulation in clinical settings, thereby hindering overall advancements in the field. Therefore, rather than resigning to the notion that criminal psychopaths are beyond help, there is a pressing need to improve both research methodologies and advocacy efforts to better understand the their potential for rehabilitation (Felthous, 2011). Without such advancements, our society risks perpetuating a simplistic and punitive view of psychopathy that neglects the possibility of rehabilitation.

This thesis not only contributes to the understanding of how subclinical psychopathic traits affect empathy and aggression through novel combinations of EEG analysis, behavioural assessment, and meta-analytical review of NIBS interventions; it also prompts us to consider not only the intricate nature of psychopathy but also the capacity of our society to embrace a more empathetic and nuanced perspective on such individuals. Only then can we truly assess whether those we label as irredeemable are in fact capable of change – or if, in our failure to seek this understanding, we have simply resigned ourselves to viewing them as *evil*.

Notes

[1] "Cruelty (from the Latin *crudelem*, "morally rough") is the deliberate infliction of physical or psychological pain on a living creature, (often with the) delight of the perpetrators" (Nell, 2006, p.211).

[2] The term "sadism" originates from the German concept of *schadenfreude* – or malicious joy –, first used by the psychologist Richard von Krafft-Ebing in reference to Marquis de Sade, an 18th-century French nobleman and writer notorious for his depravity. Today, sadism is defined as the tendency to find pleasure or satisfaction in causing or observing pain, suffering, or humiliation in others (Foulkes, 2019).

[3] The term "conduct disorder" defines a pattern of antisocial, aggressive, and defiant behaviours in children and adolescents that violate societal norms and the rights of others (American Psychiatric Association, 2013).

[4] Mental state attribution – often called *mentalising* or *Theory of Mind (ToM)* – refers to the capacity to infer other's mental states (thoughts, beliefs, feelings) to predict behavioural responses (Baron-Cohen et al., 1985; Byom & Mutlu, 2013).

[5] Pain empathy refers to the capacity to understand and share another individual's experience of pain – whether physical or emotional. It involves not only a cognitive recognition of another's suffering but also an affective response, where the observer may experience a vicarious form of the distress (Decety & Jackson, 2004; Singer et al., 2004).

[6] Personal distress refers to a self-focused, aversive emotional reaction to witnessing another's suffering that is characterised by feelings of discomfort and anxiety, primarily aimed at alleviating one's own distress rather than addressing the needs of the other (Batson et al., 1983; Eisenberg & Miller, 1987). Unlike affective empathy – which involves sharing and resonating with the others' emotions and experiences while still differentiating the *self* from others –

personal distress is more self-focused and not necessarily congruent with the other's emotional state but rather centred on the individual's own emotional turmoil. As such, personal distress is more likely to lead to withdrawal or self-protective behaviours, with research suggesting that this response is not reflective of genuine empathy (Vachon & Lynam, 2016).

[7] "Morality encompasses notions of justice, fairness, and rights, as well as maxims regarding interpersonal relations. Another theoretical view contends that morality includes the full array of psychological mechanisms that are active in the moral lives of people across cultures. Rather than stating the content of moral issues (e.g., justice and welfare), this definition specifies the function of moral systems as an interlocking sets of values, virtues, norms, practices, and identities that work together to suppress or regulate selfishness and make cooperative social life possible. What seems clear is that, regardless of the definition, a central focus of morality is the judgment of the rightness or wrongness of acts or behaviours that knowingly cause harm to people." (Decety & Cowell, 2015, p.3)

[8] Psychopathic meanness refers to a cluster of personality traits marked by callousness, a tendency towards cruelty, and a profound lack of empathy for others (Patrick et al., 2009).

[9] Congenital insensitivity to pain (CIP), also known as congenital analgesia, is a rare genetic condition first described by Thrush in 1973. Individuals with CIP are born without the ability to perceive physical pain, despite having otherwise intact sensory modalities and normal cognitive development (see Schon et al., 2020 for a review)

[10] Electroconvulsive therapy (ECT) – also known as electroshock therapy – is a medical treatment that uses controlled electrical currents to induce a brief seizure under general anaesthesia, primarily to treat severe mental health conditions (Abrams, 2002). ECT has also been explored as a way to reduce aggression in individuals with psychiatric disorders (Ujkaj et al., 2012).

Bibliography

Abbink, K., & Harris, D. (2019). In-group favouritism and out-group discrimination in naturally occurring groups. *PLOS ONE*, *14*(9), e0221616. https://doi.org/10.1371/journal.pone.0221616

Abrams, R. (2002). *Electroconvulsive therapy*. Oxford University Press.

Almeida, R., Prata, C., Pereira, M. R., Barbosa, F., & Ferreira-Santos, F. (2024). Neuronal Correlates of Empathy: A Systematic Review of Event-Related Potentials Studies in Perceptual Tasks. *Brain Sciences*, *14*(5), 504.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Ames, D. L., Jenkins, A. C., Banaji, M. R., & Mitchell, J. P. (2008). Taking Another Person's Perspective Increases Self-Referential Neural Processing. *Psychological Science*, *19*(7), 642–644. https://doi.org/10.1111/j.1467-9280.2008.02135.x

Amrhein, V., Greenland, S., & McShane, B. (2019). Scientists rise up against statistical significance. *Nature*, *567*(7748), 305–307.

Anderson, N. E., & Kiehl, K. A. (2012). The psychopath magnetized: Insights from brain imaging. *Trends in Cognitive Sciences*, *16*(1), 52–60. https://doi.org/10.1016/j.tics.2011.11.008

Anderson, N. E., & Kiehl, K. A. (2013). Psychopathy and Aggression: When Paralimbic Dysfunction Leads to Violence. In K. A. Miczek & A. Meyer-Lindenberg (Eds.), *Neuroscience of Aggression* (Vol. 17, pp. 369–393). Springer Berlin Heidelberg. https://doi.org/10.1007/7854_2013_257

Anderson, S. W., Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (2013). Impairment of social and moral behavior related to early damage in human prefrontal cortek. In *Social neuroscience* (pp. 29–39). Psychology Press.

Antal, A., & Paulus, W. (2013). Transcranial alternating current stimulation (tACS). *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00317

Arain, M., Haque, M., Johal, L., Mathur, P., Nel, W., Rais, A., Sandhu, R., & Sharma, S. (2013). Maturation of the adolescent brain. *Neuropsychiatric Disease and Treatment*, *9*(null), 449–461. https://doi.org/10.2147/NDT.S39776

Babiak, P., Neumann, C. S., & Hare, R. D. (2010). Corporate psychopathy: Talking the walk. *Behavioral Sciences & the Law*, *28*(2), 174–193.

Babiak, P., & O'Toole, M. E. (2012). The corporate psychopath. *FBI L. Enforcement Bull.*, *81*, 7.

Babiloni, C., Brancucci, A., Percio, C. D., Capotosto, P., Arendt-Nielsen, L., Chen, A. C. N., & Rossini, P. M. (2006). Anticipatory Electroencephalography Alpha Rhythm Predicts Subjective Perception of Pain Intensity. *The Journal of Pain*, *7*(10), 709–717. https://doi.org/10.1016/j.jpain.2006.03.005

Baez, S., Herrera, E., García, A. M., Manes, F., Young, L., & Ibáñez, A. (2017). Outcome-oriented moral evaluation in terrorists. *Nature Human Behaviour*, *1*(6), 0118.

Baez, S., Manes, F., Huepe, D., Torralva, T., Fiorentino, N., Richter, F., Huepe-Artigas, D., Ferrari, J., Montañes, P., & Reyes, P. (2014). Primary empathy deficits in frontotemporal dementia. *Frontiers in Aging Neuroscience*, *6*, 262.

Bahji, A., Forth, E., Yang, C.-C., & Khalifa, N. (2021). Transcranial direct current stimulation for empathy: A systematic review and meta-analysis. *Social Neuroscience*, *16*(3), 232–255. https://doi.org/10.1080/17470919.2021.1889657

Baker, L. A., Raine, A., Liu, J., & Jacobson, K. C. (2008). Differential Genetic and Environmental Influences on Reactive and Proactive Aggression in Children. *Journal of Abnormal Child Psychology*, *36*(8), 1265–1278. https://doi.org/10.1007/s10802-008-9249-1

Balconi, M., & Canavesio, Y. (2014). High-frequency rTMS on DLPFC increases prosocial attitude in case of decision to support people. *Social Neuroscience*, *9*(1), 82–93. https://doi.org/10.1080/17470919.2013.861361

Bandura, A. (n.d.). *Moral Disengagement in the Perpetration of Inhumanities*.

Bandura, A. (1990). Selective activation and disengagement of moral control. *Journal of Social Issues*, *46*(1), 27–46. https://doi.org/10.1111/j.1540-4560.1990.tb00270.x

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (n.d.). *Mechanisms of Moral Disengagement in the Exercise of Moral Agency*.

Banissy, M. J., & Ward, J. (2007). Mirror-touch synesthesia is linked with empathy. *Nature Neuroscience*, *10*(7), 815–816. https://doi.org/10.1038/nn1926

Barchia, K., & Bussey, K. (2011). Individual and collective social cognitive influences on peer aggression: Exploring the contribution of aggression efficacy, moral disengagement, and collective efficacy. *Aggressive Behavior*, *37*(2), 107–120. https://doi.org/10.1002/ab.20375

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology*, *13*(4), 379–398.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*(1), 37–46.

Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient: An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163–175. https://doi.org/10.1023/B:JADD.0000022607.19833.00

Barr, K. N., & Quinsey, V. L. (2004). Is psychopathy a pathology or a life strategy? Implications for social policy. *Evolutionary Psychology, Public Policy, and Personal Decisions*, 293–317.

Bastian, B., Denson, T. F., & Haslam, N. (2013). The Roles of Dehumanization and Moral Outrage in Retributive Justice. *PLOS ONE*, *8*(4), e61842. https://doi.org/10.1371/journal.pone.0061842

Bastian, B., Laham, S. M., Wilson, S., Haslam, N., & Koval, P. (2011). Blaming, praising, and protecting our humanity: The implications of everyday dehumanization for judgments of moral status. *British Journal of Social Psychology*, *50*(3), 469–483. https://doi.org/10.1348/014466610X521383

Batson, C. D. (1997). *Self–other merging and the empathy–altruism hypothesis: Reply to Neuberg et al.(1997).*

Batson, C. D., O'Quin, K., Fultz, J., Vanderplas, M., & Isen, A. M. (1983). Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. *Journal of Personality and Social Psychology*, *45*(3), 706–718. https://doi.org/10.1037/0022-3514.45.3.706

Begg, C. B., & Mazumdar, M. (1994). Operating Characteristics of a Rank Correlation Test for Publication Bias. *Biometrics*, *50*(4), 1088. https://doi.org/10.2307/2533446

Benning, S. D., Venables, N. C., & Hall, J. R. (2018). Successful psychopathy. *Handbook of Psychopathy*, *2*, 585–608.

Berkout, O. V., Young, J. N., & Gross, A. M. (2011). Mean girls and bad boys: Recent research on gender differences in conduct disorder. *Aggression and Violent Behavior*, *16*(6), 503–511. https://doi.org/10.1016/j.avb.2011.06.001

Bernhardt, B. C., & Singer, T. (2012). The Neural Basis of Empathy. In *Annual Review of Neuroscience* (Vol. 35, Issue Volume 35, 2012, pp. 1–23). Annual Reviews. https://doi.org/10.1146/annurev-neuro-062111-150536

Birbaumer, N., Veit, R., Lotze, M., Erb, M., Hermann, C., Grodd, W., & Flor, H. (2005). Deficient Fear Conditioning in Psychopathy: A Functional Magnetic Resonance Imaging Study. *Archives of General Psychiatry*, *62*(7), 799–805. https://doi.org/10.1001/archpsyc.62.7.799

Bird, G., & Viding, E. (2014a). The self to other model of empathy: Providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neuroscience & Biobehavioral Reviews*, *47*, 520–532. https://doi.org/10.1016/j.neubiorev.2014.09.021

Bird, G., & Viding, E. (2014b). The self to other model of empathy: Providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neuroscience & Biobehavioral Reviews*, *47*, 520–532.

Blair, R. J. R. (2003a). Neurobiological basis of psychopathy. *British Journal of Psychiatry*, *182*(1), 5–7. Cambridge Core. https://doi.org/10.1192/bjp.182.1.5

Blair, R. J. R. (2003b). Neurobiological basis of psychopathy. *The British Journal of Psychiatry*, *182*(1), 5–7.

Blair, R. J. R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition*, *14*(4), 698–718. https://doi.org/10.1016/j.concog.2005.06.004

Blair, R. J. R. (2007). The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends in Cognitive Sciences*, *11*(9), 387–392. https://doi.org/10.1016/j.tics.2007.07.003

Blair, R. J. R. (2013). The neurobiology of psychopathic traits in youths. *Nature Reviews Neuroscience*, *14*(11), 786–799. https://doi.org/10.1038/nrn3577

Blair, R. J. R., Peschardt, K. S., Budhani, S., Mitchell, D. G. V., & Pine, D. S. (2006). The development of psychopathy. *Journal of Child Psychology and Psychiatry*, *47*(3–4), 262–276. https://doi.org/10.1111/j.1469-7610.2006.01596.x

Blakemore, S.-J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, *9*(4), 267–277. https://doi.org/10.1038/nrn2353

Blankertz, B., Lemm, S., Treder, M., Haufe, S., & Müller, K.-R. (2011). Single-trial analysis and classification of ERP components—A tutorial. *NeuroImage*, *56*(2), 814–825.

Bliss, T. V. P., & Cooke, S. F. (2011). Long-term potentiation and long-term depression: A clinical perspective. *Clinics*, *66*, 3–17. https://doi.org/10.1590/S1807-59322011001300002

Booth, N., Williams, K., & Vea Papadopoulou. (2016). *Prisoners' Childhood and Family Backgrounds*. https://doi.org/10.13140/RG.2.1.1244.6488

Borenstein, M. (2023). How to understand and report heterogeneity in a meta-analysis: The difference between I-squared and prediction intervals. *Integrative Medicine Research*, *12*(4), 101014. https://doi.org/10.1016/j.imr.2023.101014

Botvinick, M., Jha, A. P., Bylsma, L. M., Fabian, S. A., Solomon, P. E., & Prkachin, K. M. (2005). Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. *Neuroimage*, *25*(1), 312–319.

Branchadell, V., Poy, R., Ribes Guardiola, P., Segarra, P., & Moltó, J. (2024). Psychopathic callousness and perspective taking in pain processing: An ERP study. *Social Cognitive and Affective Neuroscience, 19*. https://doi.org/10.1093/scan/nsae022

Brislin, S. J., Perkins, E. R., Ribes-Guardiola, P., Patrick, C. J., & Foell, J. (2022). Pain processing and antisocial behavior: A multimodal investigation of the roles of boldness and meanness. *Personality Disorders: Theory, Research, and Treatment*, *13*(6), 685–696. https://doi.org/10.1037/per0000556

Brower, M. C., & Price, B. H. (2001). Neuropsychiatry of frontal lobe dysfunction in violent and criminal behaviour: A critical review. *Journal of Neurology, Neurosurgery & Psychiatry*, *71*(6), 720. https://doi.org/10.1136/jnnp.71.6.720

Brown, M. (2012). Empathy and punishment. *Punishment & Society*, *14*(4), 383–401. https://doi.org/10.1177/1462474512452507

Brugman, S., Cornet, L. J. M., Smeijers, D., Smeets, K., Oostermeijer, S., Buitelaar, J. K., Verkes, R., Lobbestael, J., De Kogel, C. H., & Jansen, L. M. C. (2017). Examining the reactive proactive questionnaire in adults in forensic and non-forensic settings: A variable- and person-based approach. *Aggressive Behavior*, *43*(2), 155–162. https://doi.org/10.1002/ab.21671

Bruneau, E. G., Jacoby, N., & Saxe, R. (2015). Empathic control through coordinated interaction of amygdala, theory of mind and extended pain matrix brain regions. *NeuroImage*, *114*, 105–119. https://doi.org/10.1016/j.neuroimage.2015.04.034

Brunoni, A. R., Vanderhasselt, M.-A., Boggio, P. S., Fregni, F., Dantas, E. M., Mill, J. G., Lotufo, P. A., & Benseñor, I. M. (2013). Polarity- and valence-dependent effects of prefrontal transcranial direct current stimulation on heart rate variability and salivary cortisol. *Psychoneuroendocrinology*, *38*(1), 58–66. https://doi.org/10.1016/j.psyneuen.2012.04.020

Buchanan, A., & Mathieu, D. (1986). Philosophy and Justice. In R. L. Cohen (Ed.), *Justice: Views from the Social Sciences* (pp. 11–45). Springer US. https://doi.org/10.1007/978-1-4899-3511-3_2

Bushman, B. J., & Anderson, C. A. (2001). Is it time to pull the plug on hostile versus instrumental aggression dichotomy? *Psychological Review*, *108*(1), 273–279. https://doi.org/10.1037/0033-295X.108.1.273

Buss, A. H., & Perry, M. (1992). *The aggression questionnaire. 63*(3), 452.

Byom, L. J., & Mutlu, B. (2013). Theory of mind: Mechanisms, methods, and new directions. *Frontiers in Human Neuroscience*, *7*, 413.

Byrd, A. L., Kahn, R. E., & Pardini, D. A. (2013). A Validation of the Inventory of Callous-Unemotional Traits in a Community Sample of Young Adult Males. *Journal of Psychopathology and Behavioral Assessment*, *35*(1), 20–34. https://doi.org/10.1007/s10862-012-9315-4

Caballero, A., Granberg, R., & Tseng, K. Y. (2016). Mechanisms contributing to prefrontal cortex maturation during adolescence. *The Adolescent Brain*, *70*, 4–12. https://doi.org/10.1016/j.neubiorev.2016.05.013

Camara, C. F., Sel, A., de Klerk, C. C., & Hanel, P. H. (2025). Direct and indirect effects of interpersonal callousness on aggression through empathy and moral disengagement. *Personality and Individual Differences*, *232*, 112836.

Camp, J. P., Skeem, J. L., Barchard, K., Lilienfeld, S. O., & Poythress, N. G. (2013). Psychopathic predators? Getting specific about the relation between psychopathy and violence. *Journal of Consulting and Clinical Psychology*, *81*(3), 467–480. https://doi.org/10.1037/a0031349

Campos, C., Pasion, R., Azeredo, A., Ramião, E., Mazer, P., Macedo, I., & Barbosa, F. (2022). Refining the link between psychopathy, antisocial behavior, and empathy: A meta-analytical approach across different conceptual frameworks. *Clinical Psychology Review*, *94*, 102145. https://doi.org/10.1016/j.cpr.2022.102145

Canavero, S. (2014). Criminal Minds: Neuromodulation of the Psychopathic Brain. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00124

Caprara, G. V., Tisak, M. S., Alessandri, G., Fontaine, R. G., Fida, R., & Paciello, M. (2014). The contribution of moral disengagement in mediating individual tendencies toward aggression and violence. *Developmental Psychology*, *50*(1), 71–85. https://doi.org/10.1037/a0034488

Caravita, S. C. S., Gini, G., & Pozzoli, T. (2012). Main and Moderated Effects of Moral Cognition and Status on Bullying and Defending. *Aggressive Behavior*, *38*(6), 456–468. https://doi.org/10.1002/ab.21447

Carlsmith, K. M., & Darley, J. M. (2008). Psychological Aspects of Retributive Justice. In *Advances in Experimental Social Psychology* (Vol. 40, pp. 193–236). Academic Press. https://doi.org/10.1016/S0065-2601(07)00004-4

Carlsmith, K. M., Darley, J. M., & Robinson, P. H. (2002). Why do we punish? Deterrence and just deserts as motives for punishment. *Journal of Personality and Social Psychology*, *83*(2), 284–299. https://doi.org/10.1037/0022-3514.83.2.284

Carré, J. M., Hyde, L. W., Neumann, C. S., Viding, E., & Hariri, A. R. (2013). The neural signatures of distinct psychopathic traits. *Social Neuroscience*, *8*(2), 122–135.

Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., Taylor, A., & Poulton, R. (2002). Role of Genotype in the Cycle of Violence in Maltreated Children. *Science*, *297*(5582), 851–854. https://doi.org/10.1126/science.1072290

Chang, S.-A. A., Tillem, S., Benson-Williams, C., & Baskin-Sommers, A. (2021). Cognitive Empathy in Subtypes of Antisocial Individuals. *Frontiers in Psychiatry*, *12*, 677975. https://doi.org/10.3389/fpsyt.2021.677975

Chen, S., Shi, J., Yang, X., Ye, H., & Luo, J. (2019). Modulating Activity in the Dorsolateral Prefrontal Cortex Changes Punishment in the 3-Player Prisoner's Dilemma: A Transcranial Direct Current Stimulation Study. *Frontiers in Neuroscience*, *13*. https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2019.01160

Chen, T., Li, Q., Peng, M., & Li, X. (2023). Moral transgression modulates empathy for pain: Evidence from ERP and EEG data. *Biological Psychology*, *176*, 108467. https://doi.org/10.1016/j.biopsycho.2022.108467

Cheng, Y., Hung, A.-Y., & Decety, J. (2012). Dissociation between affective sharing and emotion understanding in juvenile psychopaths. *Development and Psychopathology*, *24*(2), 623–636.

Cheng, Y., Yang, C.-Y., Lin, C.-P., Lee, P.-L., & Decety, J. (2008). The perception of pain in others suppresses somatosensory oscillations: A magnetoencephalography study. *NeuroImage*, *40*(4), 1833–1840. https://doi.org/10.1016/j.neuroimage.2008.01.064

Chialant, D., Edersheim, J., & Price, B. H. (2016). The Dialectic Between Empathy and Violence: An Opportunity for Intervention? *The Journal of Neuropsychiatry and Clinical Neurosciences*, *28*(4), 273–285. https://doi.org/10.1176/appi.neuropsych.15080207

Chowdhury, R. M., & Fernando, M. (2014). The relationships of empathy, moral identity and cynicism with consumers' ethical beliefs: The mediating role of moral disengagement. *Journal of Business Ethics*, *124*, 677–694.

Christian, P., & Soutschek, A. (2022). Causal role of right dorsolateral prefrontal cortex for norm-guided social decision making: A meta-analysis of TMS studies. *Neuropsychologia*, *176*, 108393. https://doi.org/10.1016/j.neuropsychologia.2022.108393

Cikara, M., & Fiske, S. T. (2013). Their pain, our pleasure: Stereotype content and schadenfreude. *Annals of the New York Academy of Sciences*, *1299*(1), 52–59.

Cima, M., Tonnaer, F., & Hauser, M. D. (2010). Psychopaths know right from wrong but don't care. *Social Cognitive and Affective Neuroscience*, *5*(1), 59–67. https://doi.org/10.1093/scan/nsp051

Ciucci, E., & Baroncelli, A. (2014). The emotional core of bullying: Further evidences of the role of callous–unemotional traits and empathy. *Personality and Individual Differences*, *67*, 69–74. https://doi.org/10.1016/j.paid.2013.09.033

Cleckley, H. (1941). *The mask of sanity; an attempt to reinterpret the so-called psychopathic personality.* (p. 298). Mosby.

Cogoni, C., Monachesi, B., Mazza, V., Grecucci, A., & Vaes, J. (2023). Neural dynamics of vicarious physical pain processing reflect impaired empathy toward sexually objectified versus non-sexually objectified women. *Psychophysiology*, *60*(12), e14400. https://doi.org/10.1111/psyp.14400

Colins, O. F., Fanti, K. A., Salekin, R. T., & Andershed, H. (2017). Psychopathic Personality in the General Population: Differences and Similarities Across Gender. *Journal of Personality Disorders*, *31*(1), 49–74. https://doi.org/10.1521/pedi_2016_30_237

Coll, M.-P. (2018). Meta-analysis of ERP investigations of pain empathy underlines methodological issues in ERP research. *Social Cognitive and Affective Neuroscience*, *13*(10), 1003–1017. https://doi.org/10.1093/scan/nsy072

Coll, M.-P., Viding, E., Rütgen, M., Silani, G., Lamm, C., Catmur, C., & Bird, G. (2017). Are we really measuring empathy? Proposal for a new measurement framework.

Neuroscience & Biobehavioral Reviews, *83*, 132–139. https://doi.org/10.1016/j.neubiorev.2017.10.009

Combrisson, E., & Jerbi, K. (2015). Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *Journal of Neuroscience Methods*, *250*, 126–136.

Corbera, S., Ikezawa, S., Bell, M. D., & Wexler, B. E. (2014). Physiological evidence of a deficit to enhance the empathic response in schizophrenia. *European Psychiatry*, *29*(8), 463–472. Cambridge Core. https://doi.org/10.1016/j.eurpsy.2014.01.005

Corrado, R. R., Vincent, G. M., Hart, S. D., & Cohen, I. M. (2004). Predictive validity of the Psychopathy Checklist: Youth Version for general and violent recidivism. *Behavioral Sciences & the Law*, *22*(1), 5–22. https://doi.org/10.1002/bsl.574

da Silva, D. R., Rijo, D., & Salekin, R. T. (2013). Child and adolescent psychopathy: Assessment issues and treatment needs. *Aggression and Violent Behavior*, *18*(1), 71–78.

Dallaire, D. H. (2007). Incarcerated Mothers and Fathers: A Comparison of Risks for Children and Families. *Family Relations*, *56*(5), 440–453. https://doi.org/10.1111/j.1741-3729.2007.00472.x

Darby, R. R., & Pascual-Leone, A. (2017). Moral Enhancement Using Non-invasive Brain Stimulation. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00077

Darley, J. M., & Pittman, T. S. (2003). The Psychology of Compensatory and Retributive Justice. *Personality and Social Psychology Review*, *7*(4), 324–336. https://doi.org/10.1207/S15327957PSPR0704_05

David, D., Lynn, S., & Montgomery, G. (2018). *Evidence-Based Psychotherapy: The State of the Science and Practice* (1st ed.). Wiley. https://doi.org/10.1002/9781119462996

Davidov, M., Paz, Y., Roth-Hanania, R., Uzefovsky, F., Orlitsky, T., Mankuta, D., & Zahn-Waxler, C. (2021). Caring babies: Concern for others in distress during infancy. *Developmental Science*, *24*(2), e13016. https://doi.org/10.1111/desc.13016

Day, A., Casey, S., & Gerace, A. (2010). Interventions to improve empathy awareness in sexual and violent offenders: Conceptual, empirical, and clinical issues. *Aggression and Violent Behavior*, *15*(3), 201–208.

De Ridder, J., Pihet, S., Suter, M., & Caldara, R. (2016). Empathy in Institutionalized Adolescents With Callous-Unemotional Traits: An Ecological Momentary Assessment Study of Emotion Recognition. *Criminal Justice and Behavior*, *43*(5), 653–669. https://doi.org/10.1177/0093854815618431

de Tommaso, M., Ricci, K., Conca, G., Vecchio, E., Delussi, M., & Invitto, S. (2019). Empathy for pain in fibromyalgia patients: An EEG study. *International Journal of Psychophysiology*, *146*, 43–53. https://doi.org/10.1016/j.ijpsycho.2019.09.007

de Vignemont, F., & Singer, T. (2006). The empathic brain: How, when and why? *Trends in Cognitive Sciences*, *10*(10), 435–441. https://doi.org/10.1016/j.tics.2006.08.008

de Wit, S. J., van der Werf, Y. D., Mataix-Cols, D., Trujillo, J. P., van Oppen, P., Veltman, D. J., & van den Heuvel, O. A. (2015). Emotion regulation before and after transcranial magnetic stimulation in obsessive compulsive disorder. *Psychological Medicine*, *45*(14), 3059–3073. Cambridge Core. https://doi.org/10.1017/S0033291715001026

Decety, J. (2010). The Neurodevelopment of Empathy in Humans. *Developmental Neuroscience*, *32*(4), 257–267. https://doi.org/10.1159/000317771

Decety, J. (2015). The neural pathways, development and functions of empathy. *Social Behavior, 3*, 1–6. https://doi.org/10.1016/j.cobeha.2014.12.001

Decety, J., Bartal, I. B.-A., Uzefovsky, F., & Knafo-Noam, A. (2016). Empathy as a driver of prosocial behaviour: Highly conserved neurobehavioural mechanisms across species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150077. https://doi.org/10.1098/rstb.2015.0077

Decety, J., & Cowell, J. M. (2015). Empathy, Justice, and Moral Behavior. *AJOB Neuroscience*, *6*(3), 3–14. https://doi.org/10.1080/21507740.2015.1047055

Decety, J., & Jackson, P. L. (2004). The Functional Architecture of Human Empathy. *Behavioral and Cognitive Neuroscience Reviews*, *3*(2), 71–100. https://doi.org/10.1177/1534582304267187

Decety, J., Lewis, K. L., & Cowell, J. M. (2015). Specific electrophysiological components disentangle affective sharing and empathic concern in psychopathy. *Journal of Neurophysiology*, *114*(1), 493–504. https://doi.org/10.1152/jn.00253.2015

Decety, J., & Meyer, M. (2008). From emotion resonance to empathic understanding: A social developmental neuroscience account. *Development and Psychopathology*, *20*(4), 1053–1080. Cambridge Core. https://doi.org/10.1017/S0954579408000503

Decety, J., & Michalska, K. J. (2010). Neurodevelopmental changes in the circuits underlying empathy and sympathy from childhood to adulthood. *Developmental Science*, *13*(6), 886–899.

Decety, J., Michalska, K. J., & Kinzler, K. D. (2012). The contribution of emotion and cognition to moral sensitivity: A neurodevelopmental study. *Cerebral Cortex*, *22*(1), 209–220. https://doi.org/10.1093/cercor/bhr111

Decety, J., Norman, G. J., Berntson, G. G., & Cacioppo, J. T. (2012). A neurobehavioral evolutionary perspective on the mechanisms underlying empathy. *Progress in Neurobiology*, *98*(1), 38–48. https://doi.org/10.1016/j.pneurobio.2012.05.001

Decety, J., Skelly, L. R., & Kiehl, K. A. (2013). Brain Response to Empathy-Eliciting Scenarios Involving Pain in Incarcerated Individuals With Psychopathy. *JAMA Psychiatry*, *70*(6), 638–645. https://doi.org/10.1001/jamapsychiatry.2013.27

Decety, J., & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience, 2*(1), 1–24. https://doi.org/10.1016/j.dcn.2011.05.003

Decety, J., Yang, C.-Y., & Cheng, Y. (2010). Physicians down-regulate their pain empathy response: An event-related brain potential study. *Neuroimage*, *50*(4), 1676–1682.

Decety, J., & Yoder, K. J. (2016). Empathy and motivation for justice: Cognitive empathy and concern, but not emotional empathy, predict sensitivity to injustice for others. *Social Neuroscience*, *11*(1), 1–14. https://doi.org/10.1080/17470919.2015.1029593

Del Gaizo, A. L., & Falkenbach, D. M. (2008). Primary and secondary psychopathic-traits and their relationship to perception and experience of emotion. *Personality and Individual Differences*, *45*(3), 206–212. https://doi.org/10.1016/j.paid.2008.03.019

Demirtas-Tatlidede, A., Vahabzadeh-Hagh, A. M., & Pascual-Leone, A. (2013). Can noninvasive brain stimulation enhance cognition in neuropsychiatric disorders? *Neuropharmacology*, *64*, 566–578. https://doi.org/10.1016/j.neuropharm.2012.06.020

DeWall, C. N., Anderson, C. A., & Bushman, B. J. (2013). Aggression. *Handbook of Psychology: Personality and Social Psychology, Vol. 5, 2nd Ed.*, 449–466.

Dinh, S. T., Nickel, M. M., Tiemann, L., May, E. S., Heitmann, H., Hohn, V. D., Edenharter, G., Utpadel-Fischler, D., Tölle, T. R., & Sauseng, P. (2019). Brain dysfunction in chronic pain patients assessed by resting-state electroencephalography. *Pain*, *160*(12), 2751–2765.

Dolan, M. C., & Fullam, R. S. (2009). Psychopathy and Functional Magnetic Resonance Imaging Blood Oxygenation Level-Dependent Responses to Emotional Faces in Violent Patients with Schizophrenia. *Cortical Development and Glutamatergic Dysregulation in Schizophrenia*, *66*(6), 570–577. https://doi.org/10.1016/j.biopsych.2009.03.019

Dolan, M., & Doyle, M. (2000). Violence risk prediction: Clinical and actuarial measures and the role of the Psychopathy Checklist. *British Journal of Psychiatry*, *177*(4), 303–311. Cambridge Core. https://doi.org/10.1192/bjp.177.4.303

Dotterer, H. L., Waller, R., Neumann, C. S., Shaw, D. S., Forbes, E. E., Hariri, A. R., & Hyde, L. W. (2017). Examining the Factor Structure of the Self-Report of Psychopathy Short-Form Across Four Young Adult Samples. *Assessment*, *24*(8), 1062–1079. https://doi.org/10.1177/1073191116640355

Douglas, K. S., Herbozo, S., Poythress, N. G., Belfrage, H., & Edens, J. F. (2006). Psychopathy and suicide: A multisample investigation. *Psychological Services*, *3*(2), 97–116. https://doi.org/10.1037/1541-1559.3.2.97

Douglas, K. S., Lilienfeld, S. O., Skeem, J. L., Poythress, N. G., Edens, J. F., & Patrick, C. J. (2008). Relation of Antisocial and Psychopathic Traits to Suicide-related Behavior Among Offenders. *Law and Human Behavior*, *32*(6), 511–525. https://doi.org/10.1007/s10979-007-9122-8

Dryburgh, N. S. J., & Vachon, D. D. (2019). Relating sex differences in aggression to three forms of empathy. *Personality and Individual Differences*, *151*, 109526. https://doi.org/10.1016/j.paid.2019.109526

D'Silva, K., Duggan, C., & McCarthy, L. (2004). Does treatment really make psychopaths worse? A review of the evidence. *Journal of Personality Disorders, 18*(2), 163–177.

Duerden, E. G., Arsalidou, M., Lee, M., & Taylor, M. J. (2013). Lateralization of affective processing in the insula. *NeuroImage*, *78*, 159–175. https://doi.org/10.1016/j.neuroimage.2013.04.014

Durlak, J. A. (2009). How to Select, Calculate, and Interpret Effect Sizes. *Journal of Pediatric Psychology*, *34*(9), 917–928. https://doi.org/10.1093/jpepsy/jsp004

Dutton, K. (2012). *The wisdom of psychopaths*. Random House.

Easycap Equidistant Layouts, F. (2018). *Easycap Equidistant Electrode Layouts*.

Edens, J. F., Marcus, D. K., Lilienfeld, S. O., & Poythress Jr., N. G. (2006). Psychopathic, not psychopath: Taxometric evidence for the dimensional structure of psychopathy. *Journal of Abnormal Psychology*, *115*(1), 131–144. https://doi.org/10.1037/0021-843X.115.1.131

Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin*, *101*(1), 91.

Eman, S., Nicolson, R., Blades, M., & Jha, R. (2022). Gender-based personality traitsin physically aggressive and non-aggressiveantisocial behaviours. *Current Issues in Personality Psychology*. https://doi.org/10.5114/cipp.2021.110942

Engen, H. G., & Singer, T. (2013). Empathy circuits. *Macrocircuits*, *23*(2), 275–282. https://doi.org/10.1016/j.conb.2012.11.003

Ermer, E., Cope, L. M., Nyalakanti, P. K., Calhoun, V. D., & Kiehl, K. A. (2013). Aberrant paralimbic gray matter in incarcerated male adolescents with psychopathic traits. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*(1), 94-103.e3. https://doi.org/10.1016/j.jaac.2012.10.013

Erzi, S. (2020). Dark Triad and schadenfreude: Mediating role of moral disengagement and relational aggression. *Personality and Individual Differences*, *157*, 109827. https://doi.org/10.1016/j.paid.2020.109827

Espinoza, F. A., Vergara, V. M., Reyes, D., Anderson, N. E., Harenski, C. L., Decety, J., Rachakonda, S., Damaraju, E., Rashid, B., Miller, R. L., Koenigs, M., Kosson, D. S., Harenski, K., Kiehl, K. A., & Calhoun, V. D. (2018). Aberrant functional network connectivity in psychopathy from a large (N = 985) forensic sample. *Human Brain Mapping*, *39*(6), 2624–2634. https://doi.org/10.1002/hbm.24028

Essau, C. A., Sasagawa, S., & Frick, P. J. (2006). Callous-unemotional traits in a community sample of adolescents. *Assessment*, *13*(4), 454–469.

Euler, F., Steinlin, C., & Stadler, C. (2017). Distinct profiles of reactive and proactive aggression in adolescents: Associations with cognitive and affective empathy. *Child and Adolescent Psychiatry and Mental Health*, *11*(1), 1. https://doi.org/10.1186/s13034-016-0141-4

Fabi, S., & Leuthold, H. (2017). Empathy for pain influences perceptual and motor processing: Evidence from response force, ERPs, and EEG oscillations. *Social Neuroscience*, *12*(6), 701–716. https://doi.org/10.1080/17470919.2016.1238009

Fabi, S., & Leuthold, H. (2018). Racial bias in empathy: Do we process dark- and fair-colored hands in pain differently? An EEG study. *Neuropsychologia*, *114*, 143–157. https://doi.org/10.1016/j.neuropsychologia.2018.04.024

Falla, D., Romera, E. M., & Ortega-Ruiz, R. (2021). Aggression, Moral Disengagement and Empathy. A Longitudinal Study Within the Interpersonal Dynamics of Bullying. *Frontiers in Psychology*, *12*, 703468. https://doi.org/10.3389/fpsyg.2021.703468

Fallon, N., Roberts, C., & Stancak, A. (2020). Shared and distinct functional networks for empathy and pain processing: A systematic review and meta-analysis of fMRI studies. *Social Cognitive and Affective Neuroscience*, *15*(7), 709–723. https://doi.org/10.1093/scan/nsaa090

Fan, Y., Duncan, N. W., de Greck, M., & Northoff, G. (2011). Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neuroscience & Biobehavioral Reviews*, *35*(3), 903–911. https://doi.org/10.1016/j.neubiorev.2010.10.009

Fan, Y., & Han, S. (2008). Temporal dynamic of neural mechanisms involved in empathy for pain: An event-related brain potential study. *Neuropsychologia*, *46*(1), 160–173.

Farrington, D. P. (2000). Psychosocial predictors of adult antisocial personality and adult convictions. *Behavioral Sciences & the Law*, *18*(5), 605–622. https://doi.org/10.1002/1099-0798(200010)18:5<605::AID-BSL406>3.0.CO;2-0

Feather, N. T. (1994). Human values and their relation to justice. *Journal of Social Issues*, *50*(4), 129–151. https://doi.org/10.1111/j.1540-4560.1994.tb01201.x

Feather, N. T., & Dawson, S. (1998). Judging deservingness and affect in relation to another's employment or unemployment: A test of a justice model. *European Journal of Social Psychology*, *28*(3), 361–381. https://doi.org/10.1002/(SICI)1099-0992(199805/06)28:3<361::AID-EJSP867>3.0.CO;2-E

Fecteau, S., Boggio, P., Fregni, F., & Pascual-Leone, A. (2013). Modulation of untruthful responses with non-invasive brain stimulation. *Frontiers in Psychiatry*, *3*. psyh. https://doi.org/10.3389/fpsyt.2012.00097

Felthous, A. R. (2011). The "untreatability" of psychopathy and hospital commitment in the USA. *International Journal of Law and Psychiatry*, *34*(6), 400–405. https://doi.org/10.1016/j.ijlp.2011.10.004

Felthous, A. R. (2015). The appropriateness of treating psychopathic disorders. *CNS Spectrums*, *20*(3), 182–189.

Fontaine, R. G., Fida, R., Paciello, M., Tisak, M. S., & Caprara, G. V. (2014). The mediating role of moral disengagement in the developmental course from peer rejection in adolescence to crime in early adulthood. *Psychology, Crime & Law*, *20*(1), 1–19.

Foulkes, L. (2019). Sadism: Review of an elusive construct. *Personality and Individual Differences*, *151*, 109500. https://doi.org/10.1016/j.paid.2019.07.010

Fragkaki, I., Cima, M., & Meesters, C. (2016). The Association Between Callous–Unemotional Traits, Externalizing Problems, and Gender in Predicting Cognitive and Affective Morality Judgments in Adolescence. *Journal of Youth and Adolescence*, *45*(9), 1917–1930. https://doi.org/10.1007/s10964-016-0527-x

Frick, P. J., Cornell, A. H., Barry, C. T., Bodin, S. D., & Dane, H. E. (2003). Callous-Unemotional Traits and Conduct Problems in the Prediction of Conduct Problem Severity, Aggression, and Self-Report of Delinquency. *Journal of Abnormal Child Psychology*.

Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014). Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. *Psychological Bulletin*, *140*(1), 1–57. https://doi.org/10.1037/a0033076

Frick, P. J., & White, S. F. (2008). Research Review: The importance of callous-unemotional traits for developmental models of aggressive and antisocial behavior. *Journal of Child Psychology and Psychiatry*, *49*(4), 359–375. https://doi.org/10.1111/j.1469-7610.2007.01862.x

Gantiva, C., Cendales, R., Díaz, M., & González, Y. (2021). Is There Really a Relationship Between Empathy and Aggression? Evidence From Physiological and Self-Report Measures. *Journal of Interpersonal Violence*, *36*(7–8), 3438–3458. https://doi.org/10.1177/0886260518776999

Gao, Y., Raine, A., Venables, P. H., Dawson, M. E., & Mednick, S. A. (2010). Association of poor childhood fear conditioning and adult crime. *American Journal of Psychiatry*, *167*(1), 56–60.

Gini, G., Pozzoli, T., & Bussey, K. (2015). The Role of Individual and Collective Moral Disengagement in Peer Aggression and Bystanding: A Multilevel Analysis. *Journal of Abnormal Child Psychology*, *43*(3), 441–452. https://doi.org/10.1007/s10802-014-9920-7

Gini, G., Pozzoli, T., & Hymel, S. (2014). Moral disengagement among children and youth: A meta-analytic review of links to aggressive behavior. *Aggressive Behavior*, *40*(1), 56–68. https://doi.org/10.1002/ab.21502

Glenn, A. L., & Raine, A. (2008). The neurobiology of psychopathy. *Psychiatric Clinics of North America*, *31*(3), 463–475.

Glenn, A. L., & Raine, A. (2009). Psychopathy and instrumental aggression: Evolutionary, neurobiological, and legal perspectives. *International Journal of Law and Psychiatry*, *32*(4), 253–258. https://doi.org/10.1016/j.ijlp.2009.04.002

Glenn, A. L., & Raine, A. (2014). Neurocriminology: Implications for the punishment, prediction and prevention of criminal behaviour. *Nature Reviews Neuroscience*, *15*(1), 54–63.

Goldfine, A. M., Victor, J. D., Conte, M. M., Bardin, J. C., & Schiff, N. D. (2011). Determination of awareness in patients with severe brain injury using EEG power spectral analysis. *Clinical Neurophysiology*, *122*(11), 2157–2168. https://doi.org/10.1016/j.clinph.2011.03.022

Goldstein, A. P., Glick, B., & Gibbs, J. C. (1998). *Aggression replacement training: A comprehensive intervention for aggressive youth, Rev*. Research Press.

Gómez-Tapia, C., Bozic, B., & Longo, L. (2022). On the Minimal Amount of EEG Data Required for Learning Distinctive Human Features for Task-Dependent Biometric Applications. *Frontiers in Neuroinformatics*, *16*. https://www.frontiersin.org/journals/neuroinformatics/articles/10.3389/fninf.2022.844667

Gonzalez-Liencres, C., Brown, E. C., Tas, C., Breidenstein, A., & Brüne, M. (2016). Alterations in event-related potential responses to empathy for pain in schizophrenia. *Psychiatry Research*, *241*, 14–21. https://doi.org/10.1016/j.psychres.2016.04.091

Goubert, L., Craig, K. D., Vervoort, T., Morley, S., Sullivan, M. J., de CAC, W., Cano, A., & Crombez, G. (2005). Facing others in pain: The effects of empathy. *Pain*, *118*(3), 285–288.

Green, E., Silverman, M., & Geil, G. (1944). Petit Mal Electro-shock therapy of criminal psychopaths. *Criminal Psychopathology*, *5*(4), 667–673.

Grice-Jackson, T., Critchley, H. D., Banissy, M. J., & Ward, J. (2017). Consciously feeling the pain of others reflects atypical functional connectivity between the pain matrix and frontal-parietal regions. *Frontiers in Human Neuroscience*, *11*, 507.

Hall, J. R., & Benning, S. D. (2006). The 'Successful' Psychopath: Adaptive and Subclinical Manifestations of Psychopathy in the General Population. *Handbook of Psychopathy.*, 459–478.

Hallett, M. (2007). Transcranial Magnetic Stimulation: A Primer. *Neuron*, *55*(2), 187–199. https://doi.org/10.1016/j.neuron.2007.06.026

Han, S., Fan, Y., & Mao, L. (2008). Gender difference in empathy for pain: An electrophysiological investigation. *Brain Research*, *1196*, 85–93. https://doi.org/10.1016/j.brainres.2007.12.062

Hanel, P. H. P., & Vione, K. C. (2016). Do Student Samples Provide an Accurate Estimate of the General Public? *PLOS ONE*, *11*(12), e0168354. https://doi.org/10.1371/journal.pone.0168354

Hare, R. D. (1970). *Psychopathy: Theory and research.*

Hare, R. D. (1980). A research scale for the assessment of psychopathy in criminal populations. *Personality and Individual Differences*, *1*(2), 111–119. https://doi.org/10.1016/0191-8869(80)90028-8

Hare, R. D. (2003). *The Hare Psychopathy Checklist Revised (2nd ed.)*. Multi-Health Systems.

Hare, R. D. (2006). Psychopathy: A clinical and forensic overview. *Psychiatric Clinics*, *29*(3), 709–724.

Hare, R. D., & Hart, S. D. (1993). Psychopathy, mental disorder, and crime. In *Mental disorder and crime.* (pp. 104–115). Sage Publications, Inc. https://doi.org/10.1001/archpsyc.55.1.86

Hare, R. D., Hart, S. D., & Harpur, T. J. (1991). Psychopathy and the DSM-IV criteria for antisocial personality disorder. *Journal of Abnormal Psychology*, *100*(3), 391.

Hare, R. D., & Neumann, C. S. (2008). Psychopathy as a Clinical and Empirical Construct. *Annual Review of Clinical Psychology*, *4*(1), 217–246. https://doi.org/10.1146/annurev.clinpsy.3.022806.091452

Harmon-Jones, E., & Winkielman, P. (2007). Asymmetrical frontal cortical activity, affective valence, and motivational direction. *Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior*, 137–156.

Harris, G. T., & Rice, M. E. (2006). Treatment of Psychopathy: A Review of Empirical Findings. In *Handbook of psychopathy.* (pp. 555–572). The Guilford Press.

Haslam, S. A., O'Brien, A., Jetten, J., Vormedal, K., & Penna, S. (2005). Taking the strain: Social identity, social support, and the experience of stress. *British Journal of Social Psychology*, *44*(3), 355–370. https://doi.org/10.1348/014466605X37468

Hawes, D. J., & Dadds, M. R. (2007). Stability and malleability of callous-unemotional traits during treatment for childhood conduct problems. *Journal of Clinical Child and Adolescent Psychology*, *36*(3), 347–355.

Hawes, D. J., Price, M. J., & Dadds, M. R. (2014). Callous-Unemotional Traits and the Treatment of Conduct Problems in Childhood and Adolescence: A Comprehensive Review. *Clinical Child and Family Psychology Review*, *17*(3), 248–267. https://doi.org/10.1007/s10567-014-0167-1

Hayes, A. F. (2012). *PROCESS [macro]*.

Hecht, L. K., Berg, J. M., Lilienfeld, S. O., & Latzman, R. D. (2016). Parsing the heterogeneity of psychopathy and aggression: Differential associations across dimensions and gender. *Personality Disorders: Theory, Research, and Treatment*, *7*(1), 2–14. https://doi.org/10.1037/per0000128

Heym, N., Firth, J., Kibowski, F., Sumich, A., Egan, V., & Bloxsom, C. A. J. (2019). Empathy at the Heart of Darkness: Empathy Deficits That Bind the Dark Triad and Those That Mediate Indirect Relational Aggression. *Frontiers in Psychiatry*, *10*. https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2019.00095

Higgins, J., & Green, S. (2011). Internet. *Cochrane Handbook for Systematic Reviews of Interventions Version*, *5*(0).

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Hoenen, M., Lübke, K. T., & Pause, B. M. (2015). Somatosensory mu activity reflects imagined pain intensity of others. *Psychophysiology*, *52*(12), 1551–1558. https://doi.org/10.1111/psyp.12522

Hoffman, M. L. (1975). Developmental synthesis of affect and cognition and its implications for altruistic motivation. *Developmental Psychology*, *11*(5), 607.

Hofman, D., & Schutter, D. J. L. G. (2012). Asymmetrical frontal resting-state beta oscillations predict trait aggressive tendencies and behavioral inhibition. *Social Cognitive and Affective Neuroscience*, *7*(7), 850–857. https://doi.org/10.1093/scan/nsr060

Holt, S. E., Meloy, J. R., & Strack, S. (1999). Sadism and Psychopathy in Violent and Sexually Violent Offenders. *J Am Acad Psychiatry Law*, *27*(1).

Hu, J., Li, Y., Yin, Y., Blue, P. R., Yu, H., & Zhou, X. (2017). How do self-interest and other-need interact in the brain to determine altruistic behavior? *NeuroImage*, *157*, 598–611. psyh. https://doi.org/10.1016/j.neuroimage.2017.06.040

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424.

Hu, L., Peng, W., Valentini, E., Zhang, Z., & Hu, Y. (2013). Functional features of nociceptive-induced suppression of alpha band electroencephalographic oscillations. *The Journal of Pain*, *14*(1), 89–99.

Hu, Y., Strang, S., & Weber, B. (2015). Helping or punishing strangers: Neural correlates of altruistic decisions as third-party and of its relation to empathic concern. *Frontiers in Behavioral Neuroscience*, *9*. https://www.frontiersin.org/journals/behavioral-neuroscience/articles/10.3389/fnbeh.2015.00024

Huang, Y.-Z., Edwards, M. J., Rounis, E., Bhatia, K. P., & Rothwell, J. C. (2005). Theta Burst Stimulation of the Human Motor Cortex. *Neuron*, *45*(2), 201–206. https://doi.org/10.1016/j.neuron.2004.12.033

Hyde, L. W., Shaw, D. S., & Moilanen, K. L. (2010). Developmental Precursors of Moral Disengagement and the Role of Moral Disengagement in the Development of Antisocial Behavior. *Journal of Abnormal Child Psychology*, *38*(2), 197–209. https://doi.org/10.1007/s10802-009-9358-5

Iannetti, G. D., & Mouraux, A. (2010). From the neuromatrix to the pain matrix (and back). *Experimental Brain Research*, *205*(1), 1–12.

Ickes, W. (1993). Empathic accuracy. *Journal of Personality*, *61*(4), 587–610.

Ickes, W. (2001). Measuring empathic accuracy. In *Interpersonal sensitivity* (pp. 237–260). Psychology Press.

Jackson, P. L., Rainville, P., & Decety, J. (2006). To what extent do we share the pain of others? Insight from the neural bases of pain empathy. *Pain*, *125*(1), 5–9. https://doi.org/10.1016/j.pain.2006.09.013

Jakobwitz, S., & Egan, V. (2006). The dark triad and normal personality traits. *Personality and Individual Differences*, *40*(2), 331–339. https://doi.org/10.1016/j.paid.2005.07.006

Jansen, J. M., van den Heuvel, O. A., van der Werf, Y. D., de Wit, S. J., Veltman, D. J., van den Brink, W., & Goudriaan, A. E. (2019). The Effect of High-Frequency Repetitive Transcranial Magnetic Stimulation on Emotion Processing, Reappraisal, and Craving in Alcohol Use Disorder Patients and Healthy Controls: A Functional Magnetic Resonance Imaging Study. *Frontiers in Psychiatry*, *10*. https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2019.00272

Jauniaux, J., Khatibi, A., Rainville, P., & Jackson, P. L. (2019). A meta-analysis of neuroimaging studies on pain empathy: Investigating the role of visual information and observers' perspective. *Social Cognitive and Affective Neuroscience*, *14*(8), 789–813. https://doi.org/10.1093/scan/nsz055

Jin, W. Y., & Peng, M. (2021). The effects of social perception on moral judgment. *Frontiers in Psychology*, *11*, 557216.

Jolliffe, D., & Farrington, D. P. (2004). Empathy and offending: A systematic review and meta-analysis. *Aggression and Violent Behavior*, *9*(5), 441–476. https://doi.org/10.1016/j.avb.2003.03.001

Jolliffe, D., & Farrington, D. P. (2011). Is low empathy related to bullying after controlling for individual and social background variables? *Journal of Adolescence*, *34*(1), 59–71. https://doi.org/10.1016/j.adolescence.2010.02.001

Jones, A. P., Laurens, K. R., Herba, C. M., Barker, G. J., & Viding, E. (2009). Amygdala hypoactivity to fearful faces in boys with conduct problems and callous-unemotional traits. *American Journal of Psychiatry*, *166*(1), 95–102.

Kahn, R. E., Byrd, A. L., & Pardini, D. A. (2013). Callous-unemotional traits robustly predict future criminal offending in young men. *Law and Human Behavior*, *37*(2), 87–97. https://doi.org/10.1037/b0000003

Karim, A. A., Schneider, M., Lotze, M., Veit, R., Sauseng, P., Braun, C., & Birbaumer, N. (2010). The Truth about Lying: Inhibition of the Anterior Prefrontal Cortex Improves Deceptive Behavior. *Cerebral Cortex*, *20*(1), 205–213. https://doi.org/10.1093/cercor/bhp090

Keysers, C., Kaas, J. H., & Gazzola, V. (2010). Somatosensation in social perception. *Nature Reviews Neuroscience*, *11*(6), 417–428. https://doi.org/10.1038/nrn2833

Kiehl, K. A., & Buckholtz, J. W. (2010). Inside the Mind of a Psychopath. *Scientific American Mind*, *21*(4), 22–29. JSTOR.

Kiehl, K. A., & Hoffman, M. B. (2011). The criminal psychopathy: History, neuroscience, treatment, and ecconomics. *Jurimetrics*, *51*, 355–397.

Kiehl, K. A., Smith, A. M., Hare, R. D., Mendrek, A., Forster, B. B., Brink, J., & Liddle, P. F. (2001). Limbic abnormalities in affective processing by criminal psychopaths as revealed by functional magnetic resonance imaging. *Biological Psychiatry*, *50*(9), 677–684. https://doi.org/10.1016/S0006-3223(01)01222-7

Kimonis, E. R., Frick, P. J., Skeem, J. L., Marsee, M. A., Cruise, K., Munoz, L. C., Aucoin, K. J., & Morris, A. S. (2008). Assessing callous–unemotional traits in adolescent offenders: Validation of the Inventory of Callous–Unemotional Traits. *Psychopathic Traits and Risk Assessment in Children and Adolescents*, *31*(3), 241–252. https://doi.org/10.1016/j.ijlp.2008.04.002

Kirsch, L. G., & Becker, J. V. (2007). Emotional deficits in psychopathy and sexual sadism: Implications for violent and sadistic behavior. *Clinical Psychology Review*, *27*(8), 904–922. https://doi.org/10.1016/j.cpr.2007.01.011

Knight, G. P., Guthrie, I. K., Page, M. C., & Fabes, R. A. (2002). Emotional arousal and gender differences in aggression: A meta-analysis. *Aggressive Behavior*, *28*(5), 366–393. https://doi.org/10.1002/ab.80011

Koch, J. L. A. (1892). *Die psychopathischen Minderwertigkeiten...* (Vol. 2). O. Maier.

Koenigs, M. (2012). The role of prefrontal cortex in psychopathy. *Reviews in the Neurosciences*, *23*(3). https://doi.org/10.1515/revneuro-2012-0036

Koenigs, M., Kruepke, M., Zeier, J., & Newman, J. P. (2012). Utilitarian moral judgment in psychopathy. *Social Cognitive and Affective Neuroscience*, *7*(6), 708–714. https://doi.org/10.1093/scan/nsr048

Kohlberg, L. (1981). *The philosophy of moral development: Moral stages and the idea of justice*.

Kolivoski, K. M., & Shook, J. J. (2016). Incarcerating Juveniles in Adult Prisons: Examining the Relationship Between Age and Prison Behavior in Transferred Juveniles. *Criminal Justice and Behavior*, *43*(9), 1242–1259. https://doi.org/10.1177/0093854816631793

Kringelbach, M. L., & Rolls, E. T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: Evidence from neuroimaging and neuropsychology. *Progress in Neurobiology*, *72*(5), 341–372.

Krstic, S., Neumann, C. S., Roy, S., Robertson, C. A., Knight, R. A., & Hare, R. D. (2018). Using latent variable- and person-centered approaches to examine the role of psychopathic traits in sex offenders. *Personality Disorders: Theory, Research, and Treatment*, *9*(3), 207–216. https://doi.org/10.1037/per0000249

Kuhlman, W. N. (1978). Functional topography of the human mu rhythm. *Electroencephalography and Clinical Neurophysiology*, *44*(1), 83–93.

Lalonde, R. N., & Silverman, R. A. (1994). Behavioral preferences in response to social injustice: The effects of group permeability and social identity salience. *Journal of Personality and Social Psychology*, *66*(1), 78.

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage*, *54*(3), 2492–2502. https://doi.org/10.1016/j.neuroimage.2010.10.014

Lamm, C., & Majdandžić, J. (2015). The role of shared neural activations, mirror neurons, and morality in empathy – A critical comment. *Social Neuroscience*, *90*, 15–24. https://doi.org/10.1016/j.neures.2014.10.008

Legrain, V., Iannetti, G. D., Plaghki, L., & Mouraux, A. (2011). The pain matrix reloaded: A salience detection system for the body. *Progress in Neurobiology*, *93*(1), 111–124.

Leistico, A.-M. R., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A Large-Scale Meta-Analysis Relating the Hare Measures of Psychopathy to Antisocial Conduct. *Law and Human Behavior*, *32*(1), 28–45. https://doi.org/10.1007/s10979-007-9096-6

Lemm, S., Blankertz, B., Curio, G., & Muller, K.-R. (2005). Spatio-spectral filters for improving the classification of single trial EEG. *IEEE Transactions on Biomedical Engineering*, *52*(9), 1541–1548.

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology*, *68*(1), 151–158. https://doi.org/10.1037/0022-3514.68.1.151

Levitan, J. A., & Vachon, D. D. (2021). The nuanced association between deficient empathy and sexual aggression. *Personality and Individual Differences*, *177*, 110812. https://doi.org/10.1016/j.paid.2021.110812

Levy, D. A. (2022). Optimizing the social utility of judicial punishment: An evolutionary biology and neuroscience perspective. *Frontiers in Human Neuroscience*, *16*. https://www.frontiersin.org/journals/human-neuroscience/articles/10.3389/fnhum.2022.967090

Levy, J., Goldstein, A., Pratt, M., & Feldman, R. (2018). Maturation of Pain Empathy from Child to Adult Shifts from Single to Multiple Neural Rhythms to Support Interoceptive Representations. *Scientific Reports*, *8*(1), 1810. https://doi.org/10.1038/s41598-018-19810-3

Li, F., Ball, S., Zhang, X., & Smith, A. (2020). Focal stimulation of the temporoparietal junction improves rationality in prosocial decision-making. *Scientific Reports*, *10*(1), 20275. https://doi.org/10.1038/s41598-020-76956-9

Li, X., Liu, Y., Ye, Q., Lu, X., & Peng, W. (2020). The linkage between first-hand pain sensitivity and empathy for others' pain: Attention matters. *Human Brain Mapping*, *41*(17), 4815–4828. https://doi.org/10.1002/hbm.25160

Li, X., Meng, X., Li, H., Yang, J., & Yuan, J. (2017). The impact of mood on empathy for pain: Evidence from an EEG study. *Psychophysiology*, *54*(9), 1311–1322. https://doi.org/10.1111/psyp.12882

Light, S. N., Bieliauskas, L. A., & Taylor, S. F. (2019). Measuring change in anhedonia using the "Happy Faces" task pre- to post-repetitive transcranial magnetic stimulation (rTMS) treatment to left dorsolateral prefrontal cortex in Major Depressive Disorder (MDD): Relation to empathic happiness. *Translational Psychiatry*, *9*(1), 217. https://doi.org/10.1038/s41398-019-0549-8

Lilienfeld, S. O., Widows, M. R., & Staff, P. (2005). Psychopathic personality inventory TM-revised. *Social Influence (SOI)*, *61*(65), 97.

Lin, Z., Cui, F., Wu, Y., & Wei, Q. (2024). The effect of wrongdoer's status on observer punishment recommendations: The mediating role of envy and the moderating role of belief in a just world. *Frontiers in Psychology*, *15*. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2024.1227961

Lisoni, J., Nibbio, G., Baldacci, G., Zucchetti, A., Cicale, A., Zardini, D., Miotto, P., Deste, G., Barlati, S., & Vita, A. (2024). Improving depressive symptoms in patients with schizophrenia using bilateral bipolar-nonbalanced prefrontal tDCS: Results from a double-blind sham-controlled trial. *Journal of Affective Disorders*, *349*, 165–175. https://doi.org/10.1016/j.jad.2024.01.050

Lo Cricchio, M. G., Musso, P., Lo Coco, A., Cassibba, R., & Liga, F. (2022). The relation between empathy and aggression: The role of attachment style. *Europe's Journal of Psychology*, *18*(3), 319–336. https://doi.org/10.5964/ejop.4509

Lockwood, P. L. (2016). The anatomy of empathy: Vicarious experience and disorders of social cognition. *Behavioural Brain Research*, *311*, 255–266. https://doi.org/10.1016/j.bbr.2016.05.048

Lockwood, P. L., Sebastian, C. L., McCrory, E. J., Hyde, Z. H., Gu, X., De Brito, S. A., & Viding, E. (2013). Association of callous traits with reduced neural response to others' pain in children with conduct problems. *Current Biology*, *23*(10), 901–905.

Long, J. A., & Long, M. (2021). Package 'interactions'. *Comprehensive, User-Friendly Toolkit for Probing Interactions. Available Online at: Https://Cran. r-Project. Org/Web/Packages/Interactions/Interactions. Pdf (Assessed July 29, 2021)*.

Long, Y., Zhong, M., Aili, R., Zhang, H., Fang, X., & Lu, C. (2023). Transcranial direct current stimulation of the right anterior temporal lobe changes interpersonal neural synchronization and shared mental processes. *Brain Stimulation*, *16*(1), 28–39. psyh. https://doi.org/10.1016/j.brs.2022.12.009

Lovett, B., & Sheffield, R. (2007). Affective empathy deficits in aggressive children and adolescents: A critical review. *Clinical Psychology Review*, *27*(1), 1–13. https://doi.org/10.1016/j.cpr.2006.03.003

Lu, T., & McKeown, S. (2018). The effects of empathy, perceived injustice and group identity on altruistic preferences: Towards compensation or punishment. *Journal of Applied Social Psychology*, *48*(12), 683–691.

Lübke, K. T., Sachse, C., Hoenen, M., & Pause, B. M. (2020). Mu-Suppression as an Indicator of Empathic Processes in Lesbian, Gay, and Heterosexual Adults. *Archives of Sexual Behavior*, *49*(2), 635–644. https://doi.org/10.1007/s10508-019-01491-2

Lui, J. H. L., Barry, C. T., & Sacco, D. F. (2016). Callous-unemotional traits and empathy deficits: Mediating effects of affective perspective-taking and facial emotion recognition. *Cognition and Emotion*, *30*(6), 1049–1062. https://doi.org/10.1080/02699931.2015.1047327

Luo, A., & Bussey, K. (2023). Moral disengagement in youth: A meta-analytic review. *Developmental Review*, *70*, 101101.

Mackenzie, R. (2016). Deep Brain Stimulation for Psychopaths—A No Brainer. *AJOB Neuroscience*, *7*(3), 137–139. https://doi.org/10.1080/21507740.2016.1225849

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation Analysis. *Annual Review of Psychology*, *58*(1), 593–614. https://doi.org/10.1146/annurev.psych.58.110405.085542

Mahmut, M. K., Homewood, J., & Stevenson, R. J. (2008). The characteristics of non-criminals with high psychopathy traits: Are they similar to criminal psychopaths? *Journal of Research in Personality*, *42*(3), 679–692.

Mari, T., Ali, S. H., Pacinotti, L., Powsey, S., & Fallon, N. (2025). Machine learning classification of active viewing of pain and non-pain images using EEG does not exceed chance in external validation samples. *Cognitive, Affective, & Behavioral Neuroscience*. https://doi.org/10.3758/s13415-025-01268-2

Mari, T., Asgard, O., Henderson, J., Hewitt, D., Brown, C., Stancak, A., & Fallon, N. (2023). External validation of binary machine learning models for pain intensity perception classification from EEG in healthy individuals. *Scientific Reports*, *13*(1), 242.

Mari, T., Henderson, J., Ali, S. H., Hewitt, D., Brown, C., Stancak, A., & Fallon, N. (2023b). External Validation of Machine Learning and EEG for Continuous Pain Intensity Prediction in Healthy Individuals. *bioRxiv*.

Mari, T., Henderson, J., Ali, S. H., Hewitt, D., Brown, C., Stancak, A., & Fallon, N. (2023c). Machine learning and EEG can classify passive viewing of discrete categories of visual stimuli but not the observation of pain. *BMC Neuroscience*, *24*(1), 50.

Mari, T., Henderson, J., Maden, M., Nevitt, S., Duarte, R., & Fallon, N. (2022). Systematic Review of the Effectiveness of Machine Learning Algorithms for Classifying Pain Intensity, Phenotype or Treatment Outcomes Using Electroencephalogram Data. *The Journal of Pain*, *23*(3), 349–369. https://doi.org/10.1016/j.jpain.2021.07.011

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Maroney, T. A. (2024). Law and emotion. In *Research Handbook on Law and Psychology* (pp. 64–82). Edward Elgar Publishing.

Marsh, A. A. (2016). Neural, cognitive, and evolutionary foundations of human altruism. *WIREs Cognitive Science*, *7*(1), 59–71. https://doi.org/10.1002/wcs.1377

Marsh, A. A. (2018). The neuroscience of empathy. *Emotion-Cognition Interactions*, *19*, 110–115. https://doi.org/10.1016/j.cobeha.2017.12.016

Marsh, A. A., Finger, E. C., Fowler, K. A., Adalio, C. J., Jurkowitz, I. T. N., Schechter, J. C., Pine, D. S., Decety, J., & Blair, R. J. R. (2013). Empathic responsiveness in amygdala and anterior cingulate cortex in youths with psychopathic traits. *Journal of Child Psychology and Psychiatry*, *54*(8), 900–910. https://doi.org/10.1111/jcpp.12063

Marsh, A. A., Finger, E. C., Fowler, K. A., Jurkowitz, I. T., Schechter, J. C., Henry, H. Y., Pine, D. S., & Blair, R. J. R. (2011). Reduced amygdala–orbitofrontal connectivity during moral judgments in youths with disruptive behavior disorders and psychopathic traits. *Psychiatry Research: Neuroimaging*, *194*(3), 279–286.

Marsh, A. A., Finger, E. C., Mitchell, D. G. V., Reid, M. E., Sims, C., Kosson, D. S., Towbin, K. E., Leibenluft, E., Pine, D. S., & Blair, R. J. R. (2008). Reduced Amygdala Response to Fearful Expressions in Children and Adolescents With Callous-Unemotional Traits and Disruptive Behavior Disorders. *American Journal of Psychiatry*, *165*(6), 712–720. https://doi.org/10.1176/appi.ajp.2007.07071145

Marshall, L. E., & Marshall, W. L. (2011). Empathy and antisocial behaviour. *Journal of Forensic Psychiatry & Psychology*, *22*(5), 742–759. https://doi.org/10.1080/14789949.2011.617544

McMEEKAN, E. R. L., & LISHMAN, W. A. (1975). RETEST RELIABILITIES AND INTERRELATIONSHIP OF THE ANNETT HAND PREFERENCE QUESTIONNAIRE AND THE EDINBURGH HANDEDNESS INVENTORY. *British Journal of Psychology*, *66*(1), 53–59. https://doi.org/10.1111/j.2044-8295.1975.tb01439.x

Meloy, J. R. (1997). The Psychology of Wickedness: Psychopathy and Sadism. *Psychiatric Annals*, *27*(9), 630–633. https://doi.org/10.3928/0048-5713-19970901-10

Meloy, J. R. (2006). Empirical basis and forensic application of affective and predatory violence. *Australian & New Zealand Journal of Psychiatry*, *40*(6–7), 539–547.

Meloy, J. R., Book, A., Hosker-Field, A., Methot-Jones, T., & Roters, J. (2018). Social, Sexual, and Violent Predation: Are Psychopathic Traits Evolutionarily Adaptive? *Violence and Gender*, *5*(3), 153–165. https://doi.org/10.1089/vio.2018.0012

Meng, J., Hu, L., Shen, L., Yang, Z., Chen, H., Huang, X., & Jackson, T. (2012). Emotional primes modulate the responses to others’ pain: An ERP study. *Experimental Brain Research, 220*(3), 277–286. https://doi.org/10.1007/s00221-012-3136-2

Michail, G., Dresel, C., Witkovský, V., Stankewitz, A., & Schulz, E. (2016). Neuronal oscillations in various frequency bands differ between pain and touch. *Frontiers in Human Neuroscience*, *10*, 182.

Mikula, G. (2003). Testing an attribution-of-blame model of judgments of injustice. *European Journal of Social Psychology*, *33*(6), 793–811. https://doi.org/10.1002/ejsp.184

Miller, P. A., & Eisenberg, N. (1988a). The relation of empathy to aggressive and externalizing/antisocial behavior. *Psychological Bulletin*, *103*(3), 324–344. https://doi.org/10.1037/0033-2909.103.3.324

Miller, P. A., & Eisenberg, N. (1988b). The relation of empathy to aggressive and externalizing/antisocial behavior. *Psychological Bulletin*, *103*(3), 324–344. https://doi.org/10.1037/0033-2909.103.3.324

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Research methodes and reporting. *Bmj*, *8*, 332–336.

Mohr, P., Howells, K., Gerace, A., Day, A., & Wharton, M. (2007). The role of perspective taking in anger arousal. *Personality and Individual Differences*, *43*(3), 507–517. https://doi.org/10.1016/j.paid.2006.12.019

Molero-Chamizo, A., Riquel, R. M., Moriana, J. A., Nitsche, M. A., & Rivera-Urbina, G. N. (2019). Bilateral Prefrontal Cortex Anodal tDCS Effects on Self-reported Aggressiveness in Imprisoned Violent Offenders. *Neuroscience*, *397*, 31–40. https://doi.org/10.1016/j.neuroscience.2018.11.018

Monte-Silva, K., Kuo, M.-F., Hessenthaler, S., Fresnoza, S., Liebetanz, D., Paulus, W., & Nitsche, M. A. (2013). Induction of Late LTP-Like Plasticity in the Human Motor Cortex by Repeated Non-Invasive Brain Stimulation. *Brain Stimulation*, *6*(3), 424–432. https://doi.org/10.1016/j.brs.2012.04.011

Moran, P., Rowe, R., Flach, C., Briskman, J., Ford, T., Maughan, B., Scott, S., & Goodman, R. (2009). Predictive Value of Callous-Unemotional Traits in a Large Community Sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(11), 1079–1084. https://doi.org/10.1097/CHI.0b013e3181b766ab

Morrison, I., Lloyd, D., Di Pellegrino, G., & Roberts, N. (2004). Vicarious responses to pain in anterior cingulate cortex: Is empathy a multisensory issue? *Cognitive, Affective, & Behavioral Neuroscience*, *4*(2), 270–278. https://doi.org/10.3758/CABN.4.2.270

Morrow, E. P. (2020). Cognitive, Affective, and General Empathy in Individuals Convicted of a Sexual Offense: A Meta-Analysis. *Sexual Abuse*, *32*(8), 883–906. https://doi.org/10.1177/1079063219858062

Motoyama, Y., Ogata, K., Hoka, S., & Tobimatsu, S. (2017). Frequency-dependent changes in sensorimotor and pain affective systems induced by empathy for pain. *Journal of Pain Research*, *Volume 10*, 1317–1326. https://doi.org/10.2147/JPR.S129791

Mouraux, A., & Iannetti, G. D. (2009). Nociceptive laser-evoked brain potentials do not reflect nociceptive-specific neural activity. *Journal of Neurophysiology*, *101*(6), 3258–3269.

Mu, Y., Fan, Y., Mao, L., & Han, S. (2008). Event-related theta and alpha oscillations mediate empathy for pain. *Brain Research*, *1234*, 128–136. https://doi.org/10.1016/j.brainres.2008.07.113

Müller, K.-R., Tangermann, M., Dornhege, G., Krauledat, M., Curio, G., & Blankertz, B. (2008). Machine learning for real-time single-trial EEG-analysis: From brain–computer interfacing to mental state monitoring. *Journal of Neuroscience Methods*, *167*(1), 82–90.

Mullins-Nelson, J. L., Salekin, R. T., & Leistico, A.-M. R. (2006). Psychopathy, Empathy, and Perspective -Taking Ability in a Community Sample: Implications for the Successful Psychopathy Concept. *International Journal of Forensic Mental Health*, *5*(2), 133–149. https://doi.org/10.1080/14999013.2006.10471238

Muñoz, L. C., & Frick, P. J. (2012). Callous-Unemotional Traits and Their Implication for Understanding and Treating Aggressive and Violent Youths. *Criminal Justice and Behavior*, *39*(6), 794–813. https://doi.org/10.1177/0093854812437019

Murphy, J. G. (2019). Marxism and retribution. In *Retribution* (pp. 107–133). Routledge.

Nejati, V., Mardanpour, A., Zabihzaheh, A., Estaji, R., Vaziri, Z. S., & Shahidi, S. (2023). The role of prefrontal cortex and temporoparietal junction in interpersonal comfort and emotional approach. *Scientific Reports*, *13*(1), 21636.

Nell, V. (2006). Cruelty's rewards: The gratifications of perpetrators and spectators. *Behavioral and Brain Sciences*, *29*(3), 211–224. Cambridge Core. https://doi.org/10.1017/S0140525X06009058

Neuberg, S. L., Cialdini, R. B., Brown, S. L., Luce, C., Sagarin, B. J., & Lewis, B. P. (1997). *Does empathy lead to anything more than superficial helping? Comment on Batson et al.(1997).*

Neumann, C. S., & Hare, R. D. (2008). Psychopathic traits in a large community sample: Links to violence, alcohol use, and intelligence. *Journal of Consulting and Clinical Psychology*, *76*(5), 893–899. https://doi.org/10.1037/0022-006X.76.5.893

Niel, K. A., Hunnicutt-Ferguson, K., Reidy, D. E., Martinez, M. A., & Zeichner, A. (2007). Relationship of Pain Tolerance with Human Aggression. *Psychological Reports*, *101*(1), 141–144. https://doi.org/10.2466/pr0.101.1.141-144

Nitsche, M. A., Cohen, L. G., Wassermann, E. M., Priori, A., Lang, N., Antal, A., Paulus, W., Hummel, F., Boggio, P. S., Fregni, F., & Pascual-Leone, A. (2008). Transcranial direct current stimulation: State of the art 2008. *Brain Stimulation*, *1*(3), 206–223. https://doi.org/10.1016/j.brs.2008.06.004

Nitsche, M. A., Liebetanz, D., Antal, A., Lang, N., Tergau, F., & Paulus, W. (2003). Chapter 27 Modulation of cortical excitability by weak direct current stimulation – technical, safety and functional aspects. In *Supplements to Clinical Neurophysiology* (Vol. 56, pp. 255–276). Elsevier. https://doi.org/10.1016/S1567-424X(09)70230-2

Nitsche, M. A., & Paulus, W. (2000). Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. *The Journal of Physiology*, *527*(3), 633–639. https://doi.org/10.1111/j.1469-7793.2000.t01-1-00633.x

Niv, S., Ashrafulla, S., Joshi, A., Raine, A., Leahy, R., Baker, L. A., & Tuvblad, C. (2018). Relationships of Alpha, Beta, and Theta EEG Spectra Properties with Aggressive and Nonaggressive Antisocial Behavior in Children and Adolescents. *The American Journal of Psychology*, *131*(4), 429–437. https://doi.org/10.5406/amerjpsyc.131.4.0429

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, *2011*, 1–9. https://doi.org/10.1155/2011/156869

Osofsky, M. J., Bandura, A., & Zimbardo, P. G. (2005). The Role of Moral Disengagement in the Execution Process. *Law and Human Behavior*, *29*(4), 371–393. https://doi.org/10.1007/s10979-005-4930-1

Paciello, M., Fida, R., Cerniglia, L., Tramontano, C., & Cole, E. (2013). High cost helping scenario: The role of empathy, prosocial reasoning and moral disengagement on helping behavior. *Personality and Individual Differences*, *55*(1), 3–7.

Paciello, M., Fida, R., Tramontano, C., Lupinetti, C., & Caprara, G. V. (2008). Stability and change of moral disengagement and its impact on aggression and violence in late adolescence. *Child Development*, *79*(5), 1288–1309.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

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*(3), 913–938. Cambridge Core. https://doi.org/10.1017/S0954579409000492

Paulhus, D. L. (2014). Toward a Taxonomy of Dark Personalities. *Current Directions in Psychological Science*, *23*(6), 421–426. https://doi.org/10.1177/0963721414547737

Paulhus, D. L. (2016). *Self-Report Psychopathy Scale-Fourth Edition™*.

Pearce, S. (2011). Zero Degrees of Empathy: A New Theory of Human Cruelty By Simon Baron-Cohen. Allen Lane Publishing. 2011. £20.00 (hb). 208pp. ISBN: 9780713997910. *British Journal of Psychiatry*, *199*(6), 520–520. https://doi.org/10.1192/bjp.bp.111.098434

Pecukonis, E. V. (1990). A cognitive/affective empathy training program as a function of ego development in aggressive adolescent females. *Adolescence*, *25*(97), 59.

Peled-Avron, L., Glasner, L., Gvirts, H. Z., & Shamay-Tsoory, S. G. (2019). The role of the inferior frontal gyrus in vicarious social touch: A transcranial direct current stimulation (tDCS) study. *Developmental Cognitive Neuroscience*, *35*, 115–121. psyh. https://doi.org/10.1016/j.dcn.2018.04.010

Perry, A., Bentin, S., Bartal, I. B.-A., Lamm, C., & Decety, J. (2010). “Feeling” the pain of those who are different from us: Modulation of EEG in the mu/alpha range. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(4), 493–504. https://doi.org/10.3758/CABN.10.4.493

Piaget, J. (1932). The moral judgment of the. *Chi1d. New York: McGraw-Hil1 Book Co*.

Piguet, O., Hornberger, M., Mioshi, E., & Hodges, J. R. (2011). Behavioural-variant frontotemporal dementia: Diagnosis, clinical staging, and management. *The Lancet Neurology*, *10*(2), 162–172.

Pinel, P. (1801). *Traité sur l’Alienation mentale*. Chez Richard, Caille Et Ravier, Libraires.

Ploner, M., Gross, J., Timmermann, L., Pollok, B., & Schnitzler, A. (2006). Oscillatory activity reflects the excitability of the human somatosensory system. *Neuroimage*, *32*(3), 1231–1236.

Polanía, R., Nitsche, M. A., & Ruff, C. C. (2018). Studying and modifying brain function with non-invasive brain stimulation. *Nature Neuroscience*, *21*(2), 174–187. https://doi.org/10.1038/s41593-017-0054-4

Polaschek, D. L. L. (2014). Adult criminals with psychopathy: Common beliefs about treatability and change have little empirical support. *Current Directions in Psychological Science*, *23*(4), 296–301. https://doi.org/10.1177/0963721414535211

Porter, S., Woodworth, M., Earle, J., Drugge, J., & Boer, D. (2003). Characteristics of sexual homicides committed by psychopathic and nonpsychopathic offenders. *Law and Human Behavior*, *27*(5), 459–470. https://doi.org/10.1023/A:1025461421791

Preston, O. C., & Anestis, J. C. (2020). The Indirect Relationships between Psychopathic Traits and Proactive and Reactive Aggression through Empathy and Emotion Dysregulation. *Journal of Psychopathology and Behavioral Assessment*, *42*(3), 409–423. https://doi.org/10.1007/s10862-019-09760-z

Preston, S. D., & De Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(1), 1–20. https://doi.org/10.1017/S0140525X02000018

Prichard. (1835). *Scientific intelligence: Proceedings of the medical section op the british association at Dublin, in August, 1835*.

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*(2), 159–171. https://doi.org/10.1002/ab.20115

Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., Van Swieten, J. C., Seelaar, H., Dopper, E. G., & Onyike, C. U. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, *134*(9), 2456–2477.

Rêgo, G. G., Lapenta, O. M., Marques, L. M., Costa, T. L., Leite, J., Carvalho, S., Gonçalves, Ó. F., Brunoni, A. R., Fregni, F., & Boggio, P. S. (2015). Hemispheric dorsolateral prefrontal cortex lateralization in the regulation of empathy for pain. *Neuroscience Letters*, *594*, 12–16. https://doi.org/10.1016/j.neulet.2015.03.042

Reidy, D. E., Dimmick, K., MacDonald, K., & Zeichner, A. (2009). The relationship between pain tolerance and trait aggression: Effects of sex and gender role. *Aggressive Behavior*, *35*(5), 422–429. https://doi.org/10.1002/ab.20315

Reidy, D. E., Kearns, M. C., & DeGue, S. (2013). Reducing psychopathic violence: A review of the treatment literature. *Aggression and Violent Behavior*, *18*(5), 527–538. https://doi.org/10.1016/j.avb.2013.07.008

Reidy, D. E., Shelley-Tremblay, J. F., & Lilienfeld, S. O. (2011). Psychopathy, reactive aggression, and precarious proclamations: A review of behavioral, cognitive, and biological research. *Aggression and Violent Behavior*, *16*(6), 512–524. https://doi.org/10.1016/j.avb.2011.06.002

Ricarte, J. J., Caravaca-Sánchez, F., Barry, T. J., & Aizpurua, E. (2022). Suicide behaviours in incarcerated males: Links to psychopathic traits, forms of aggression to others, personal characteristics, and current penitentiary variables. *Journal of Forensic and Legal Medicine*, *89*, 102357. https://doi.org/10.1016/j.jflm.2022.102357

Rice, M. E., Harris, G. T., & Cormier, C. A. (1992). An evaluation of a maximum security therapeutic community for psychopaths and other mentally disordered offenders. *Law and Human Behavior*, *16*(4), 399–412. https://doi.org/10.1007/BF02352266

Richardson, H., Lisandrelli, G., Riobueno-Naylor, A., & Saxe, R. (2018). Development of the social brain from age three to twelve years. *Nature Communications*, *9*(1), 1027. https://doi.org/10.1038/s41467-018-03399-2

Richter, S. M., & Bixler, R. P. (2022). Complexifying urban expansion: An exploratory, gradient-based approach. *Buildings & Cities*, *3*(1).

Riečanský, I., Paul, N., Kölble, S., Stieger, S., & Lamm, C. (2015). Beta oscillations reveal ethnicity ingroup bias in sensorimotor resonance to pain of others. *Social Cognitive and Affective Neuroscience, 10*(7), 893–901.

Risser, S., & Eckert, K. (2016). Investigating the relationships between antisocial behaviors, psychopathic traits, and moral disengagement. *International Journal of Law and Psychiatry*, *45*, 70–74. https://doi.org/10.1016/j.ijlp.2016.02.012

Ritchie, M. B., Neufeld, R. W. J., Yoon, M., Li, A., & Mitchell, D. G. V. (2022). Predicting youth aggression with empathy and callous unemotional traits: A Meta-analytic review. *Clinical Psychology Review*, *98*, 102186. https://doi.org/10.1016/j.cpr.2022.102186

Rivet, B., Souloumiac, A., Attina, V., & Gibert, G. (2009). xDAWN algorithm to enhance evoked potentials: Application to brain–computer interface. *IEEE Transactions on Biomedical Engineering*, *56*(8), 2035–2043.

Rodrigues, J., Leipold, N., Hewig, J., & Hein, G. (2024). The relationship of the source of punishment and personality traits with investment and punishment in a public goods game. *Scientific Reports*, *14*(1), 21046. https://doi.org/10.1038/s41598-024-71106-x

Romero-Martínez, Á., Bressanutti, S., & Moya-Albiol, L. (2020). A Systematic Review of the Effectiveness of Non-Invasive Brain Stimulation Techniques to Reduce Violence Proneness by Interfering in Anger and Irritability. *Journal of Clinical Medicine*, *9*(3), 882. https://doi.org/10.3390/jcm9030882

Rommel, C., Paillard, J., Moreau, T., & Gramfort, A. (2022). Data augmentation for learning predictive models on EEG: a systematic comparison. *Journal of Neural Engineering*, *19*(6), 066020.

Rosseel, Y. (2012). **lavaan**: An *R* Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2). https://doi.org/10.18637/jss.v048.i02

Ruby, P., & Decety, J. (2004). How Would You Feel versus How Do You Think She Would Feel? A Neuroimaging Study of Perspective-Taking with Social Emotions. *Journal of Cognitive Neuroscience*, *16*(6), 988–999. https://doi.org/10.1162/0898929041502661

Rybak, M., Crayton, J. W., Young, I. J., Herba, E., & Konopka, L. M. (2006). Frontal Alpha Power Asymmetry in Aggressive Children and Adolescents with Mood and Disruptive Behavior Disorders. *Clinical EEG and Neuroscience*, *37*(1), 16–24. https://doi.org/10.1177/155005940603700105

Sagi, A., & Hoffman, M. L. (1976). Empathic distress in the newborn. *Developmental Psychology*, *12*(2), 175.

Saladino, V., Lin, H., Zamparelli, E., & Verrastro, V. (2021). Neuroscience, Empathy, and Violent Crime in an Incarcerated Population: A Narrative Review. *Frontiers in Psychology*, *12*, 694212. https://doi.org/10.3389/fpsyg.2021.694212

Salekin, R. T. (2019). Psychopathy and therapeutic pessimism Clinical lore or clinical reality? *Clinical Forensic Psychology and Law*, 257–290.

Salekin, R. T., Tippey, J. G., & Allen, A. D. (2012). Treatment of conduct problem youth with interpersonal callous traits using mental models: Measurement of risk and change. *Behavioral Sciences & the Law*, *30*(4), 470–486.

Salekin, R. T., Worley, C., & Grimes, R. D. (2010). Treatment of psychopathy: A review and brief introduction to the mental model mpproach for psychopathy. *Behavioral Sciences & the Law*, *28*(2), 235–266.

Salomons, T. V., Iannetti, G. D., Liang, M., & Wood, J. N. (2016). The “pain matrix” in pain-free individuals. *JAMA Neurology*, *73*(6), 755–756.

Santamaría-García, H., Baez, S., Reyes, P., Santamaría-García, J. A., Santacruz-Escudero, J. M., Matallana, D., Arévalo, A., Sigman, M., García, A. M., & Ibáñez, A. (2017). A lesion model of envy and Schadenfreude: Legal, deservingness and moral dimensions as revealed by neurodegeneration. *Brain*, *140*(12), 3357–3377.

Schiller, B., Baumgartner, T., & Knoch, D. (2014). Intergroup bias in third-party punishment stems from both ingroup favoritism and outgroup discrimination. *Evolution and Human Behavior*, *35*(3), 169–175. https://doi.org/10.1016/j.evolhumbehav.2013.12.006

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Psychology Software Tools, Incorporated.

Schon, K. R., Parker, A. P. J., & Woods, C. G. (2020). *Congenital insensitivity to pain overview*.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, *99*(6), 323–338.

Schutter, D. J., de Weijer, A. D., Meuwese, J. D., Morgan, B., & van Honk, J. (2008). Interrelations between motivational stance, cortical excitability, and the frontal electroencephalogram asymmetry of emotion: A transcranial magnetic stimulation study. *Human Brain Mapping*, *29*(5), 574–580.

Seara-Cardoso, A., Viding, E., Lickley, R. A., & Sebastian, C. L. (2015). Neural responses to others’ pain vary with psychopathic traits in healthy adult males. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(3), 578–588. https://doi.org/10.3758/s13415-015-0346-7

Sedgwick, P. (2014). Randomised controlled trials: “Within subject” versus “between subject” designs. *BMJ : British Medical Journal*, *349*, g6435. https://doi.org/10.1136/bmj.g6435

Seibert, L. A., Miller, J. D., Few, L. R., Zeichner, A., & Lynam, D. R. (2011). An examination of the structure of self-report psychopathy measures and their relations with general traits and externalizing behaviors. *Personality Disorders: Theory, Research, and Treatment*, *2*(3), 193–208. https://doi.org/10.1037/a0019232

Sergiou, C. S., Santarnecchi, E., Franken, I. H. A., & van Dongen, J. D. M. (2020). The effectiveness of Transcranial Direct Current Stimulation as an intervention to improve empathic abilities and reduce violent behavior: A literature review. *Aggression and Violent Behavior*, *55*, 101463. https://doi.org/10.1016/j.avb.2020.101463

Sergiou, C. S., Santarnecchi, E., Romanella, S. M., Wieser, M. J., Franken, I. H. A., Rassin, E. G. C., & van Dongen, J. D. M. (2022). Transcranial Direct Current Stimulation Targeting the Ventromedial Prefrontal Cortex Reduces Reactive Aggression and Modulates Electrophysiological Responses in a Forensic Population. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *7*(1), 95–107. https://doi.org/10.1016/j.bpsc.2021.05.007

Serin, R. C. (1996). Violent recidivism in criminal psychopaths. *Law and Human Behavior*, *20*, 207–217.

SETO, M. C., & BARBAREE, H. E. (1999). Psychopathy, Treatment Behavior, and Sex Offender Recidivism. *Journal of Interpersonal Violence*, *14*(12), 1235–1248. https://doi.org/10.1177/088626099014012001

Shamay-Tsoory, S. G. (2011). The Neural Bases for Empathy. *The Neuroscientist*, *17*(1), 18–24. https://doi.org/10.1177/1073858410379268

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, *132*(3), 617–627. https://doi.org/10.1093/brain/awn279

Shaver, K. G., & Shaver, K. G. (1985). A Theory of Blame. *The Attribution of Blame: Causality, Responsibility, and Blameworthiness*, 155–176.

Shichman, R., & Weiss, A. (2022). Sending a message or seeking retributive justice? An experimental investigation of why people punish. *Journal of Behavioral and Experimental Economics*, *100*, 101918. https://doi.org/10.1016/j.socec.2022.101918

Shulman, E. P., Cauffman, E., Piquero, A. R., & Fagan, J. (2011). Moral disengagement among serious juvenile offenders: A longitudinal study of the relations between morally disengaged attitudes and offending. *Developmental Psychology*, *47*(6), 1619–1632. https://doi.org/10.1037/a0025404

Simas, E. N., Clifford, S., & Kirkland, J. H. (2020). How empathic concern fuels political polarization. *American Political Science Review*, *114*(1), 258–269.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for Pain Involves the Affective but not Sensory Components of Pain. *Science*, *303*(5661), 1157–1162. https://doi.org/10.1126/science.1093535

Skeem, J. L., & Cooke, D. J. (2010). Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate. *Psychological Assessment*, *22*(2), 433–445. https://doi.org/10.1037/a0008512

Skeem, J. L., Poythress, N., Edens, J. F., Lilienfeld, S. O., & Cale, E. M. (2003). Psychopathic personality or personalities? Exploring potential variants of psychopathy and their implications for risk assessment. *Aggression and Violent Behavior*, *8*(5), 513–546. https://doi.org/10.1016/S1359-1789(02)00098-8

Smits, F. M., Schutter, D. J. L. G., Van Honk, J., & Geuze, E. (2020). Does non-invasive brain stimulation modulate emotional stress reactivity? *Social Cognitive and Affective Neuroscience*, *15*(1), 23–51. https://doi.org/10.1093/scan/nsaa011

Smits, F. M., Schutter, D. J., van Honk, J., & Geuze, E. (2020). Does non-invasive brain stimulation modulate emotional stress reactivity? *Social Cognitive and Affective Neuroscience*, *15*(1), 23–51.

Snowdon, M. E., & Cathcart, S. (2018). tDCS potentiation provides no evidence for a link between right dorsal–lateral prefrontal cortical activity and empathic responding. *Social Neuroscience*, *13*(2), 190–201. psyh. https://doi.org/10.1080/17470919.2016.1273133

Stagg, C. J., Antal, A., & Nitsche, M. A. (2018). Physiology of Transcranial Direct Current Stimulation. *The Journal of ECT*, *34*(3). https://journals.lww.com/ectjournal/fulltext/2018/09000/physiology_of_transcranial_direct_current.3.aspx

Sticca, F., & Perren, S. (2015). The chicken and the egg: Longitudinal associations between moral deficiencies and bullying: A parallel process latent growth model. *Merrill-Palmer Quarterly*, *61*(1), 85–100.

Sutton, J., Smith, P. K., & Swettenham, J. (1999). Social cognition and bullying: Social inadequacy or skilled manipulation? *British Journal of Developmental Psychology*, *17*(3), 435–450. https://doi.org/10.1348/026151099165384

Tager-Flusberg, H., Baron-Cohen, S., & Dohen, D. J. (2000). *Understanding other minds: Perspectives from developmental cognitive neuroscience*. Oxford University Press.

Tarrant, M., Branscombe, N. R., Warner, R. H., & Weston, D. (2012). Social identity and perceptions of torture: It's moral when we do it. *Journal of Experimental Social Psychology*, *48*(2), 513–518.

Tarrant, M., Dazeley, S., & Cottom, T. (2009). Social categorization and empathy for outgroup members. *British Journal of Social Psychology*, *48*(3), 427–446.

ter Harmsel, J. F., van Dongen, J. D. M., Zijlmans, J., van der Pol, T. M., Marhe, R., & Popma, A. (2022). Neurobiological Responses towards Stimuli Depicting Aggressive Interactions in Delinquent Young Adults and Controls: No Relation to Reactive and Proactive Aggression. *Brain Sciences*, *12*(2). https://doi.org/10.3390/brainsci12020124

Tew, J., Dixon, L., Harkins, L., & Bennett, A. (2012). Investigating changes in anger and aggression in offenders with high levels of psychopathic traits attending the Chromis violence reduction programme. *Criminal Behaviour and Mental Health*, *22*(3), 191–201.

Thrush, D. (1973). Congenital insensitivity to pain. *Brain*, *96*(2), 369–386.

Tillem, S., van Dongen, J., Brazil, I. A., & Baskin-Sommers, A. (2018). Psychopathic traits are differentially associated with efficiency of neural communication. *Psychophysiology*, *55*(9), e13194. https://doi.org/10.1111/psyp.13194

Timmers, I., Park, A. L., Fischer, M. D., Kronman, C. A., Heathcote, L. C., Hernandez, J. M., & Simons, L. E. (2018). Is Empathy for Pain Unique in Its Neural Correlates? A Meta-Analysis of Neuroimaging Studies of Empathy. *Frontiers in Behavioral Neuroscience*, *12*. https://www.frontiersin.org/articles/10.3389/fnbeh.2018.00289

Trivedi-Bateman, N., & Crook, E. L. (2022). The optimal application of empathy interventions to reduce antisocial behaviour and crime: A review of the literature. *Psychology, Crime & Law*, *28*(8), 796–819. https://doi.org/10.1080/1068316X.2021.1962870

Tusche, A., Böckler, A., Kanske, P., Trautwein, F.-M., & Singer, T. (2016). Decoding the Charitable Brain: Empathy, Perspective Taking, and Attention Shifts Differentially Predict Altruistic Giving. *The Journal of Neuroscience*, *36*(17), 4719. https://doi.org/10.1523/JNEUROSCI.3392-15.2016

Tyler, T. R., & Blader, S. L. (2003). The group engagement model: Procedural justice, social identity, and cooperative behavior. *Personality and Social Psychology Review*, *7*(4), 349–361.

Ujkaj, M., Davidoff, D. A., Seiner, S. J., Ellison, J. M., Harper, D. G., & Forester, B. P. (2012). Safety and efficacy of electroconvulsive therapy for the treatment of agitation and aggression in patients with dementia. *The American Journal of Geriatric Psychiatry*, *20*(1), 61–72.

Ullman, J. B., & Bentler, P. M. (2012). Structural Equation Modeling. In *Handbook of Psychology, Second Edition*. https://doi.org/10.1002/9781118133880.hop202023

Underwood, B., & Moore, B. (1982). Perspective-taking and altruism. *Psychological Bulletin*, *91*(1), 143–173. https://doi.org/10.1037/0033-2909.91.1.143

Vachon, D. D., & Lynam, D. R. (2016). Fixing the Problem With Empathy: Development and Validation of the Affective and Cognitive Measure of Empathy. *Assessment*, *23*(2), 135–149. https://doi.org/10.1177/1073191114567941

Vachon, D. D., Lynam, D. R., & Johnson, J. A. (2014a). The (non)relation between empathy and aggression: Surprising results from a meta-analysis. *Psychological Bulletin*, *140*(3), 751–773. https://doi.org/10.1037/a0035236

Vachon, D. D., Lynam, D. R., & Johnson, J. A. (2014b). The (non)relation between empathy and aggression: Surprising results from a meta-analysis. *Psychological Bulletin*, *140*(3), 751–773. https://doi.org/10.1037/a0035236

Vachon-Presseau, E., Roy, M., Martel, M.-O., Albouy, G., Chen, J., Budell, L., Sullivan, M. J., Jackson, P. L., & Rainville, P. (2012). Neural processing of sensory and emotional-communicative information associated with the perception of vicarious pain. *NeuroImage*, *63*(1), 54–62.

Vaes, J., Meconi, F., Sessa, P., & Olechowski, M. (2016). Minimal humanity cues induce neural empathic reactions towards non-human entities. *Neuropsychologia*, *89*, 132–140. https://doi.org/10.1016/j.neuropsychologia.2016.06.004

Valentini, E. (2010). The role of anterior insula and anterior cingulate in empathy for pain. *Journal of Neurophysiology*, *104*(2), 584–586.

Valentini, E., Liang, M., Aglioti, S. M., & Iannetti, G. D. (2012). Seeing touch and pain in a stranger modulates the cortical responses elicited by somatosensory but not auditory stimulation. *Human Brain Mapping*, *33*(12), 2873–2884.

Van Dongen, J. D. M. (2020). The Empathic Brain of Psychopaths: From Social Science to Neuroscience in Empathy. *Frontiers in Psychology*, *11*, 695. https://doi.org/10.3389/fpsyg.2020.00695

Van Dongen, J. D. M., Brazil, I. A., van der Veen, F. M., & Franken, I. H. A. (2018). Electrophysiological correlates of empathic processing and its relation to psychopathic meanness. *Neuropsychology*, *32*(8), 996–1006. https://doi.org/10.1037/neu0000477

Van Hazebroek, B. C. M., Olthof, T., & Goossens, F. A. (2017). Predicting aggression in adolescence: The interrelation between (a lack of) empathy and social goals. *Aggressive Behavior*, *43*(2), 204–214. https://doi.org/10.1002/ab.21675

Vanderhasselt, M.-A., De Raedt, R., Namur, V., Valiengo, L. C., Lotufo, P. A., Bensenor, I. M., Baeken, C., Boggio, P. S., & Brunoni, A. R. (2016). Emotional reactivity to valence-loaded stimuli are related to treatment response of neurocognitive therapy. *Journal of Affective Disorders*, *190*, 443–449.

Vanman, E. J. (2016). The role of empathy in intergroup relations. *Current Opinion in Psychology*, *11*, 59–63.

Vasconcelos, M., Viding, E., Sebastian, C. L., Faria, S., Almeida, P. R., Gonçalves, Ó. F., Gonçalves, R. A., Sampaio, A., & Seara-Cardoso, A. (2021). Callous-Unemotional Traits Moderate Anticipated Guilt and Wrongness Judgments to Everyday Moral Transgressions in Adolescents. *Frontiers in Psychiatry*, *12*. https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2021.625328

Vaughn, M. G., Edens, J. F., Howard, M. O., & Smith, S. T. (2009). An Investigation of Primary and Secondary Psychopathy in a Statewide Sample of Incarcerated Youth. *Youth Violence and Juvenile Justice*, *7*(3), 172–188. https://doi.org/10.1177/1541204009333792

Verona, E., Patrick, C. J., & Joiner, T. E. (2001). Psychopathy, antisocial personality, and suicide risk. *Journal of Abnormal Psychology*, *110*(3), 462–470. https://doi.org/10.1037/0021-843X.110.3.462

Viechtbauer, W. (2010). Conducting Meta-Analyses in *R* with the metafor Package. *Journal of Statistical Software*, *36*(3). https://doi.org/10.18637/jss.v036.i03

Vieira, J. B., Ferreira-Santos, F., Almeida, P. R., Barbosa, F., Marques-Teixeira, J., & Marsh, A. A. (2015). Psychopathic traits are associated with cortical and subcortical volume alterations in healthy individuals. *Social Cognitive and Affective Neuroscience*, *10*(12), 1693–1704. https://doi.org/10.1093/scan/nsv062

Viki, G. T., Fullerton, I., Raggett, H., Tait, F., & Wiltshire, S. (2012). The Role of Dehumanization in Attitudes Toward the Social Exclusion and Rehabilitation of Sex Offenders. *Journal of Applied Social Psychology*, *42*(10), 2349–2367. https://doi.org/10.1111/j.1559-1816.2012.00944.x

Vitaglione, G. D., & Barnett, M. A. (2003). Assessing a new dimension of empathy: Empathic anger as a predictor of helping and punishing desires. *Motivation and Emotion*, *27*, 301–325.

von Grundherr, M., Jauernig, J., & Uhl, M. (2021). To Condemn Is Not to Punish: An Experiment on Hypocrisy. *Games*, *12*(2). https://doi.org/10.3390/g12020038

W. Wu, Z. Chen, X. Gao, Y. Li, E. N. Brown, & S. Gao. (2015). Probabilistic Common Spatial Patterns for Multichannel EEG Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *37*(3), 639–653. https://doi.org/10.1109/TPAMI.2014.2330598

Waller, R., Wagner, N. J., Barstead, M. G., Subar, A., Petersen, J. L., Hyde, J. S., & Hyde, L. W. (2020). A meta-analysis of the associations between callous-unemotional traits and empathy, prosociality, and guilt. *Clinical Psychology Review*, *75*, 101809. https://doi.org/10.1016/j.cpr.2019.101809

Walsh, V., & Cowey, A. (2000). Transcranial magnetic stimulation and cognitive neuroscience. *Nature Reviews Neuroscience*, *1*(1), 73–80.

Wang, G., Li, J., Yin, X., Li, S., & Wei, M. (2016). Modulating activity in the orbitofrontal cortex changes trustees' cooperation: A transcranial direct current stimulation study. *Behavioural Brain Research*, *303*, 71–75. psyh. https://doi.org/10.1016/j.bbr.2016.01.047

Wang, H., Zhang, Y., Ao, L., Huang, R., Meng, Y., Jia, S., Zhang, X., & Liu, Y. (2025). Can guilt enhance sensitivity to other's suffering? An EEG investigation into moral emotions and pain empathy. *Cerebral Cortex*, *35*(2), bhae501. https://doi.org/10.1093/cercor/bhae501

Wang, J., Wang, Y., Hu, Z., & Li, X. (2014). Transcranial direct current stimulation of the dorsolateral prefrontal cortex increased pain empathy. *Neuroscience*, *281*, 202–207. https://doi.org/10.1016/j.neuroscience.2014.09.044

Wang, N. (2023). Conducting meta-analyses of proportions in R. *Journal of Behavioral Data Science*, *3*(2), 64–126.

Wang, X., Lei, L., Yang, J., Gao, L., & Zhao, F. (2017). Moral Disengagement as Mediator and Moderator of the Relation Between Empathy and Aggression Among Chinese Male Juvenile Delinquents. *Child Psychiatry & Human Development*, *48*(2), 316–326. https://doi.org/10.1007/s10578-016-0643-6

Wang, X., Yang, J., & Yang, L. (2014). A meta-analysis of the relationship between moral disengagement and aggressive behavior. *Advances in Psychological Science*, *22*(7), 1092.

Weaver, S. S., Dargis, M., Kiehl, K. A., & Koenigs, M. (2022). Criminal Histories and Rates of Recidivism Among Two Subtypes of Psychopathic Individuals. *Criminal Justice and Behavior*, *49*(4), 471–491. https://doi.org/10.1177/00938548211033329

Weiner, B. (1995). Inferences Of Responsibility And Social Motivation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 27, pp. 1–47). Academic Press. https://doi.org/10.1016/S0065-2601(08)60402-5

White, B. A., Gordon, H., & Guerra, R. C. (2015). Callous–unemotional traits and empathy in proactive and reactive relational aggression in young women. *Personality and Individual Differences*, *75*, 185–189. https://doi.org/10.1016/j.paid.2014.11.031

White, S. F., Marsh, A. A., Fowler, K. A., Schechter, J. C., Adalio, C., Pope, K., Sinclair, S., Pine, D. S., & Blair, R. J. R. (2012). Reduced amygdala response in youths with disruptive behavior disorders and psychopathic traits: Decreased emotional response versus increased top-down attention to nonemotional features. *American Journal of Psychiatry*, *169*(7), 750–758.

Whitmarsh, S., Nieuwenhuis, I., Barendregt, H., & Jensen, O. (2011). Sensorimotor Alpha Activity is Modulated in Response to the Observation of Pain in Others. *Frontiers in Human Neuroscience*, *5*. https://www.frontiersin.org/articles/10.3389/fnhum.2011.00091

Williams, K. M., & Paulhus, D. L. (2004). Factor structure of the Self-Report Psychopathy scale (SRP-II) in non-forensic samples. *Personality and Individual Differences*, *37*(4), 765–778.

Willis, B. H., & Riley, R. D. (2017). Measuring the statistical validity of summary meta-analysis and meta-regression results for use in clinical practice. *Statistics in Medicine*, *36*(21), 3283–3301. https://doi.org/10.1002/sim.7372

Woodworth, M., & Porter, S. (2002). In cold blood: Characteristics of criminal homicides as a function of psychopathy. *Journal of Abnormal Psychology*, *111*(3), 436–445. https://doi.org/10.1037/0021-843X.111.3.436

Wu, X., Xu, F., Chen, X., Wang, L., Huang, W., Wan, K., Ji, G.-J., Xiao, G., Xu, S., Yu, F., Zhu, C., Xi, C., & Wang, K. (2018). The effect of high-definition transcranial direct current stimulation of the right inferior frontal gyrus on empathy in healthy individuals. *Frontiers in Human Neuroscience*, *12*. psyh. https://doi.org/10.3389/fnhum.2018.00446

Yang, C.-C., Khalifa, N., & Völlm, B. (2018). The effects of repetitive transcranial magnetic stimulation on empathy: A systematic review and meta-analysis. *Psychological Medicine*, *48*(5), 737–750. Cambridge Core. https://doi.org/10.1017/S003329171700232X

Yang, C.-Y., Decety, J., Lee, S., Chen, C., & Cheng, Y. (2009). Gender differences in the mu rhythm during empathy for pain: An electroencephalographic study. *Brain Research*, *1251*, 176–184. https://doi.org/10.1016/j.brainres.2008.11.062

Yao, S., Becker, B., Geng, Y., Zhao, Z., Xu, X., Zhao, W., Ren, P., & Kendrick, K. M. (2016). Voluntary control of anterior insula and its functional connections is feedback-independent and increases pain empathy. *NeuroImage*, *130*, 230–240. https://doi.org/10.1016/j.neuroimage.2016.02.035

Yildirim, B. O., & Derksen, J. J. L. (2015). Clarifying the heterogeneity in psychopathic samples: Towards a new continuum of primary and secondary psychopathy. *Aggression and Violent Behavior, 24,* 9–41. https://doi.org/10.1016/j.avb.2015.05.001

Young, L., & Saxe, R. (2008). The neural basis of belief encoding and integration in moral judgment. *Neuroimage*, *40*(4), 1912–1920.

Yuan, B., Tolomeo, S., Yang, C., Wang, Y., & Yu, R. (2021). tDCS effect on prosocial behavior: A meta-analytic review. *Social Cognitive and Affective Neuroscience*, *17*(1), 26–42. https://doi.org/10.1093/scan/nsab067

Yuan, H., Tabarak, S., Su, W., Liu, Y., Yu, J., & Lei, X. (2017). Transcranial Direct Current Stimulation of the Medial Prefrontal Cortex Affects Judgments of Moral Violations. *Frontiers in Psychology, 8,* 1812. https://doi.org/10.3389/fpsyg.2017.01812

Yuan, J., Tian, Y., Huang, X., Fan, H., & Wei, X. (2019). Emotional bias varies with stimulus type, arousal and task setting: Meta-analytic evidences. *Neuroscience & Biobehavioral Reviews*, *107*, 461–472.

Zahn-Waxler, C., Radke-Yarrow, M., & King, R. A. (1979). Child rearing and children's prosocial initiations toward victims of distress. *Child Development*, 319–330.

Zaki, J., Bolger, N., & Ochsner, K. (2008). It takes two: The interpersonal nature of empathic accuracy. *Psychological Science*, *19*(4), 399–404.

Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: Progress, pitfalls and promise. *Nature Neuroscience*, *15*(5), 675–680.

Zaki, J., Ochsner, K. N., Hanelin, J., Wager, T. D., & Mackey, S. C. (2007). Different circuits for different pain: Patterns of functional connectivity reveal distinct networks for processing pain in self and others. *Social Neuroscience*, *2*(3–4), 276–291. https://doi.org/10.1080/17470910701401973

Zaki, J., Wager, T. D., Singer, T., Keysers, C., & Gazzola, V. (2016). The Anatomy of Suffering: Understanding the Relationship between Nociceptive and Empathic Pain. *Trends in Cognitive Sciences*, *20*(4), 249–259. https://doi.org/10.1016/j.tics.2016.02.003

Zebarjadi, N., Adler, E., Kluge, A., Jääskeläinen, I. P., Sams, M., & Levy, J. (2021). Rhythmic neural patterns during empathy to vicarious pain: Beyond the affective-cognitive empathy dichotomy. *Frontiers in Human Neuroscience*, *15*, 708107.

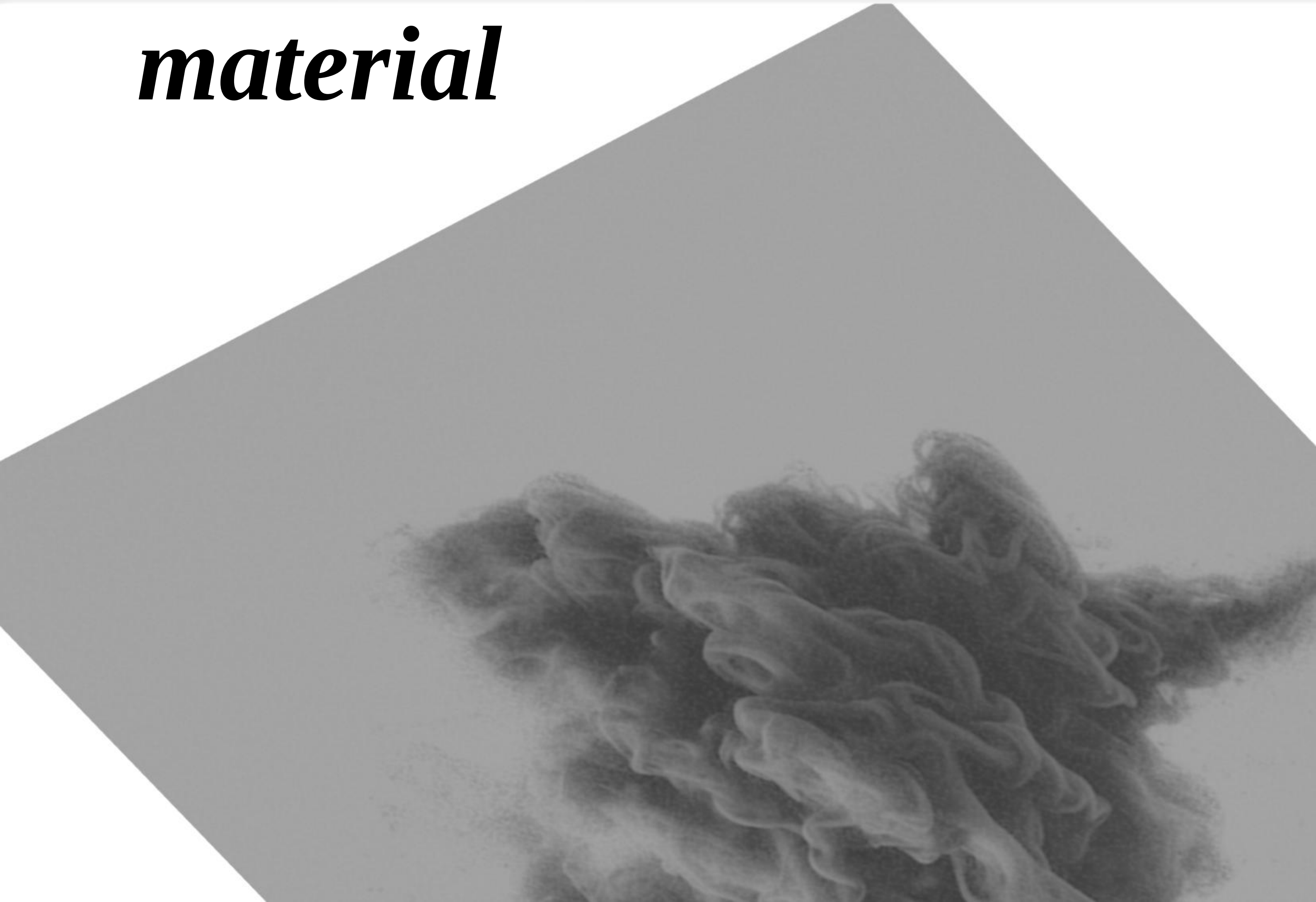
Zhang, J. J. (2024). Cortico-cortical paired associative stimulation: A novel neurostimulation solution for modulating brain connectivity and networks. *Frontiers in Neuroscience*, *17*, 1336134.

Zhang, Q., Chen, T., Liu, S., Liu, X., Zhang, Y., Yu, F., Ji, G.-J., Li, X., & Zhu, C. (2023). Effects of high-definition transcranial direct current stimulation on implicit emotion regulation of social pain in healthy individuals. *Journal of Affective Disorders*, *338*, 74–82. https://doi.org/10.1016/j.jad.2023.05.075

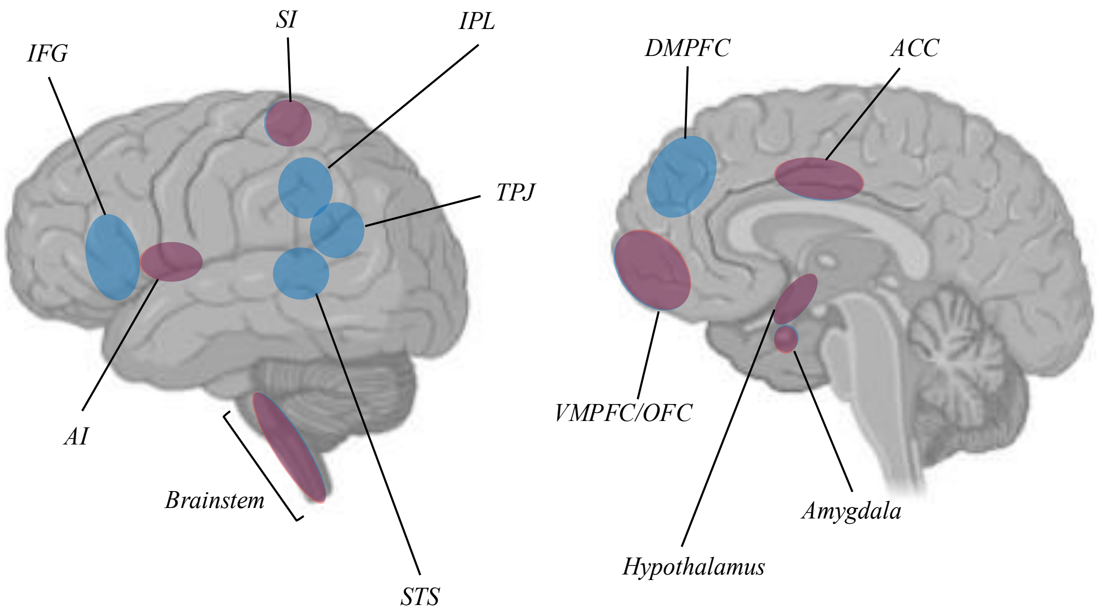
Zheng, W., Li, Y., Ye, H., & Luo, J. (2021). Effect of modulating DLPFC activity on antisocial and prosocial behavior: Evidence from a tDCS study. *Frontiers in Psychology*, *11*. psyh. https://doi.org/10.3389/fpsyg.2020.579792

Zych, I., Ttofi, M. M., & Farrington, D. P. (2019). Empathy and callous–unemotional traits in different bullying roles: A systematic review and meta-Analysis. *Trauma, Violence, & Abuse*, *20*(1), 3–21. https://doi.org/10.1177/1524838016683456

Supplementary material



Supplement 1.1. | *Empathy and the pain matrix*



Note. This figure illustrates the brain regions most consistently associated with empathy The left side of the figure presents an external lateral view of the left hemisphere, highlighting the Inferior Frontal Gyrus (IFG), Anterior Insula (AI), Primary Somatosensory Cortex (S1), Inferior Parietal Lobule (IPL), Temporal Parietal Junction (TPJ), Superior Temporal Sulcus (STS), and Brainstem. On the right, a mid-sagittal view of the left hemisphere is depicted, showcasing the Ventromedial Prefrontal Cortex (VMPFC), Orbitofrontal Cortex (OFC), Dorsomedial Prefrontal Cortex (DMPFC), Anterior Cingulate Cortex (ACC), Hypothalamus, and Amygdala. Regions highlighted in purple represent the areas where the empathy network overlaps with the pain matrix.

The *pain matrix* refers to a network of brain regions responsible for the experience of pain (Botvinick et al., 2005; Singer et al., 2004). This network is divided into two primary components: 1) sensory-discriminatory: with regions involved in the detection and processing the physical attributes of pain (e.g., location, intensity, and quality), including the primary and secondary somatosensory cortices, the thalamus, and the posterior insula; 2) affective-cognitive-evaluative: recruiting the insula, cingulate and prefrontal cortices, which are crucial for the emotional and cognitive evaluation of pain.

Neuroimaging studies have shown that many of the same regions activated during personal pain are also engaged when observing pain in others, overlapping with regions observed within the empathy network (see Figure above). Key areas involved in this response include the AI, dorsal ACC, anterior MCC, supplementary motor area (SMA), amygdala, periaqueductal grey (PAG), and the VMPFC (Lamm et al., 2011). Early interpretations took this as evidence for shared affective pain processing ("pain empathy"). However, several reviews and meta-analyses caution that these activations need not be pain- or empathy-specific. For example, Legrain *et al* (2011) note that many non-painful stimuli (and even innocuous cues in a threatening context) elicit neural activity in the AI and ACC similar to that seen in response to pain, which suggests that pain-matrix responses are not pain-exclusive but rather context-dependent. As such, they propose that this network may function as a multimodal system responsible for detecting, directing attention to, and responding to salient sensory information. Likewise, Decety and Svetlova (2012, p.9) emphasise that the activation of the pain matrix during second-hand pain observation is "not necessarily specific to the emotional experience of pain but may be related to other processes such as negative stimulus evaluation, attention to noxious stimuli, somatic monitoring, and the selection of appropriate skeletomuscular defensive movements". These interpretations suggest that shared neural activity during first-hand and second-hand pain experiences could be more related to the salience and relevance of pain-related cues rather than to empathic processing.

Supplement 2.1. | *Reliability and normality distribution of study measures*

Questionnaires/Scales	***α***	***Shapiro-Wilk***	***Skewness***	***Kurtosis***
Interpersonal Callousness (SRP-SF)	.83	.983	.311	-.014
Social Deviance (SRP-SF)	.74	$.971^{*}$	.364	-.601
Reactive Aggression (RPQ)	.81	$.974^{*}$	.391	.044
Proactive Aggression (RPQ)	.69	$.673^{***}$	2.199	5.184
Physical Aggression (BPAQ)	.81	$.976^{*}$	.169	-.746
Perspective-Taking (IRI)	.81	$.971^{*}$	.540	.211
Empathic Concern (IRI)	.77	.977	.504	-.249
Emotion Understanding (ACME)	.90	$.968^{*}$	.467	.621
Affective Resonance (ACME)	.83	$.956^{**}$	.657	-.063
Affective Dissonance (ACME)	.91	$.827^{***}$	1.905	5.335

Note. SRP-SF = Self-Report Psychopathy-Short Form; RPQ = Reactive-Proactive Aggression; IRI = Interpersonal Reactivity Index; ACME = Affective and Cognitive Measure of Empathy.

$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$. Statistically significant values in the Shapiro-Wilk test indicate that such values are not normally distributed.

Supplement 2.2. | *Reliability and normality distribution of study measures*

Questionnaires/Scales	*α*	*Shapiro-Wilk*	*Skewness*	*Kurtosis*
Interpersonal Callousness (SRP-SF)	.80	$.967^{***}$	.495	-.431
Social Deviance (SRP-SF)	.79	$.942^{***}$	.856	.468
Reactive Aggression (RPQ)	.82	$.944^{***}$	.890	.991
Proactive Aggression (RPQ)	.84	$.421^{***}$	4.834	27.903
Physical Aggression (BPAQ)	.84	$.968^{***}$	.480	-.319
Perspective-Taking (IRI)	.79	$.978^{***}$	.475	.167
Empathic Concern (IRI)	.78	$.971^{***}$	.319	-.516
Emotion Understanding (ACME)	.89	$.986^{**}$	.414	.275
Affective Resonance (ACME)	.84	$.931^{***}$	.907	.502
Affective Dissonance (ACME)	.85	$.847^{***}$	1.375	1.718
Moral Disengagement (MDS)	.87	.992	.220	-.089

Note. SRP-SF = Self-Report Psychopathy-Short Form; RPQ = Reactive-Proactive Aggression; IRI = Interpersonal Reactivity Index; ACME = Affective and Cognitive Measure of Empathy; MDS = Moral Disengagement Scale.

$^{**}p < .01$, $^{***}p < .001$. Statistically significant Shapiro-Wilk values indicate non-normal distribution.

Supplement 2.3. | *Conditional indirect effects of affective dissonance on proactive aggression at different levels of moral disengagement using the Johnson-Neyman Method*

Moderator values	***Mediator values***	***SE***	***t***	***p***
-1.05	-.06	.05	-1.24	.213
-.92	-.04	.05	-.93	.352
-.79	-.03	.04	-.56	.573
-.65	-.01	.04	-.13	.898
-.52	.01	.04	.39	.694
-.38	.03	.03	1.02	.310
-.25	.05	.03	1.76	.079
-.21	**.06**	**.03**	**1.97**	**.050**
-.11	**.07**	**.03**	**2.63**	**.009**
.02	**.09**	**.03**	**3.62**	**< .001**
.15	**.11**	**.02**	**4.67**	**< .001**
.29	**.13**	**.02**	**5.69**	**< .001**
.42	**.15**	**.02**	**6.54**	**< .001**
.56	**.17**	**.02**	**7.15**	**< .001**
.69	**.19**	**.03**	**7.51**	**< .001**
.83	**.21**	**.03**	**7.64**	**< .001**
.96	**.23**	**.03**	**7.64**	**< .001**
1.10	**.25**	**.03**	**7.54**	**< .001**
1.23	**.27**	**.04**	**7.41**	**< .001**
1.36	**.29**	**.04**	**7.25**	**< .001**
1.50	**.31**	**.04**	**7.10**	**< .001**
1.63	**.33**	**.05**	**6.96**	**< .001**

Note. Significant conditional effects at $p < .05$ are in **bold.**

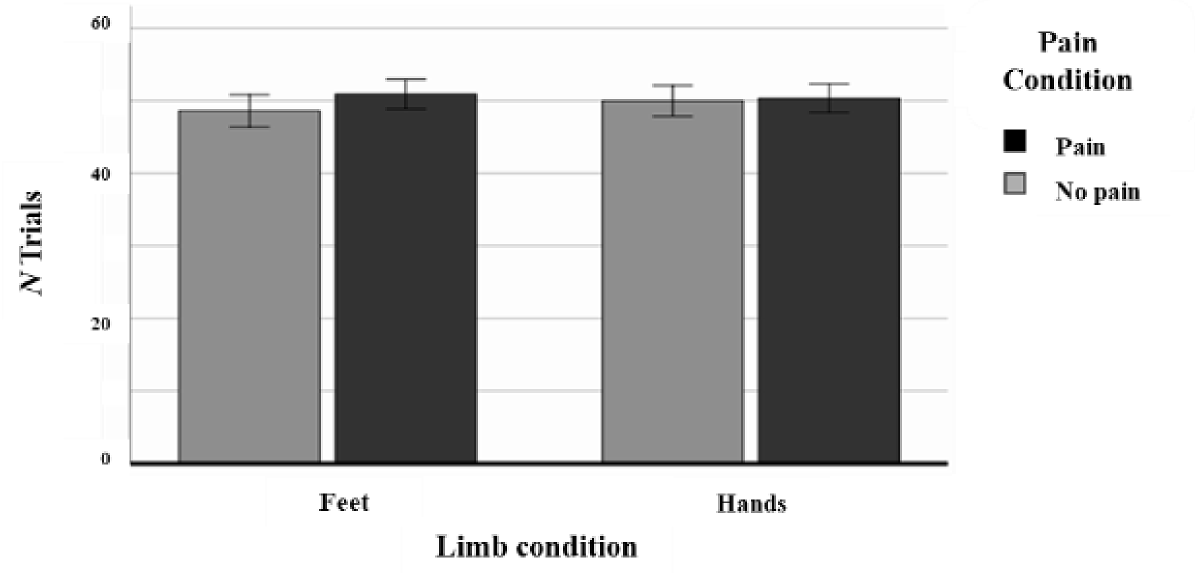
Supplement 3.1. | *Model fit comparisons for the mediation of perceived perpetrator meanness in intentional and accidental conditions in Study 3*

	Intentional	Accidental
Chi	.001	.001
CFI	1.000	.968
TL1	1.031	.805
RMSEA	.000	.257
SRMR	.004	.065
AIC	2799.757	2739.329
BIC	2821.361	2760.932
SABIC	2796.081	2735.652

Supplement 3.2. | *Model fit comparisons for the mediation of perceived perpetrator meanness in intentional and accidental conditions in Study 4*

	Intentional	**Accidental**
Chi	.011	.003
CFI	.985	.980
TL1	.949	.934
RMSEA	.093	.107
SRMR	.035	.053
AIC	9662.473	9771.480
BIC	9703.576	9812.582
SABIC	9668.688	9777.694

Supplement 4.1. | *Averaged number of trials per condition in the passive viewing task across participants*



Note. Error bars represent ±2.5 standard error.

On average, participants completed approximately 50 trials per condition overall (including catch trials). Analysis of within-subject contrasts showed no significant difference in the number of trials between the different limb and pain conditions ($F_{(1,39)}$ = 1.15, p = .290).

Supplement 4.2. | *Summary of preprocessing steps*

Participant	Initial trials	Components	Channels	Final trials
1	160	2	'Cz', 'TP9'	125
2	161	2	'Cz', 'T7', 'TP7'	109
3	153	2	'TP9'	109
4	156	4	*None*	124
5	162	3	'TP9', 'Cz'	108
6	156	3	'T7', 'T8', 'O2', 'Cz'	107
7	161	3	'O2', 'TP9', 'AF8', 'Cz'	125
8	154	3	*None*	123
9	165	2	'CP2'	136
10	158	4	'CP2', 'TP9', 'O2', 'P6'	124
11	156	3	'TP8', 'FT8', 'TP10', 'AF8'	132
12	163	3	'TP9'	133
13	159	3	'TP9'	137
14	161	4	'TP9', 'FP1'	114
15	160	2	'Cz'	137
16	169	3	'CP2', 'TP10', 'FT8', 'T7', 'FP1', 'AF7'	125
17	156	4	'Cz', 'TP9'	102
18	162	4	'T8', 'TP9'	123
19	161	3	'Cz'	109
20	157	2	'Cz'	123
21	162	3	'CP2','TP9'	127
22	156	3	'TP9'	128
23	154	3	'TP9', 'T8'	137
24	163	5	'FT7', 'AF8', 'AF7', 'F8'	118
25	162	4	'TP9'	140
26	160	3	'TP9'	112
27	156	3	'TP9'	113
28	156	3	'T7', 'TP9', 'TP7'	118
29	166	3	None	137
30	164	3	'T8', 'TP9'	114
31	155	5	*None*	99
32	163	3	*None*	121
33	164	4	*None*	125
34	149	2	'TP9'	106
35	160	3	'T7'	124
36	160	3	'TP8'	134
37	163	3	'FP1', 'TP9'	140

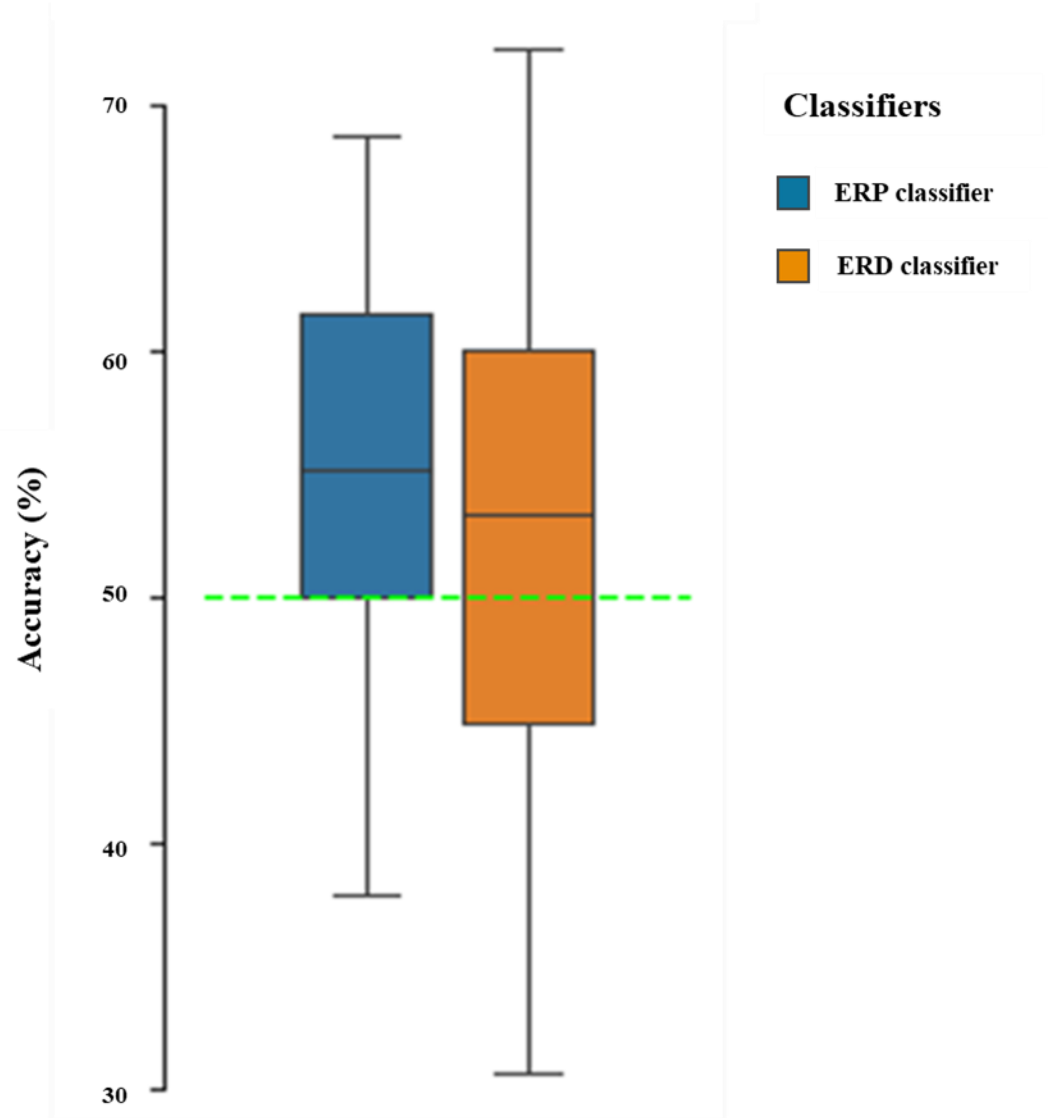
Supplement 4.3. | *Correlations between pain perception and socio-affective traits*

	ERPs			ERDs*		Rating
	C	***CP***	***P***	***α***	***θ***	***Pain-Neutral***
COG	.11 [-.23, .42]	-.01 [-.33, .32]	-.07 [-.39, .26]	.10 [-.25, .41]	.06 [-.27, .38]	.13 [-.21, .44]
RES	.20 [-.13, .50]	.23 [-.11, .52]	.19 [-.15, .49]	-.28 [-.57, .04]	.21 [-.13, .50]	.17 [-.17, .47]
DIS	.00 [-.33, .33]	-.11 [-.43, .23]	-.07 [-.39, .26]	-.03 [-.36, .30]	.10 [-.23, .42]	.07 [-.27, .39]
Cal	-.30 [-.57, .03]	-.35 [-.61, -.02]	-.22 [-.51, .12]	.21 [-.13, .50]	-.03 [-.35, .31]	-.18 [-.48, .16]
Unc	-.34 [-.60, -.02]	-.36 [-.62, -.03]	-.33 [-.59, .00]	.24 [-.09, .53]	-.17 [-.48, .16]	-.12 [-.43, .22]
Une	.03 [-.30, .36]	-.18 [-.48, .16]	-.24 [-.53, .10]	.22 [-.12, .51]	.39 [.07, .64]	.03 [-.30, .36]
React	-.02 [-.35, .31]	-.12 [-.43, 0.22]	-.08 [-.40, .25]	.03 [-.30, .36]	-.23 [-.52, .10]	.01 [-.32, .34]
Proact	-.16 [-.47, .18]	-.20 [-.50, .13]	-.12 [-.43, .22]	.04 [-.29, .37]	-.36 [-.61, -.03]	-.14 [-.45, .20]

Note. ERPs (event-related potentials) reflect brain responses recorded at central (C), centroparietal (CP), and parietal (P) electrodes, and ERDs (event-related dynamics) represent spectral power changes in the alpha (α) and theta (θ) bands. All correlation coefficients are presented with their 95% confidence intervals in brackets. The behavioural variables are defined as follows: COG = Cognitive Empathy, RES = Affective Resonance, DIS = Affective Dissonance, Cal = Callousness, Unc = Uncaring, Une = Unemotional, React = Reactive Aggression, and Proact = Proactive Aggression. “Pain-Neutral” represents the difference in subjective pain ratings between painful and non-painful images.

*Significant changes in ERDs were observed only over centroparietal electrodes.

Supplement 4.4. | *Comparisons of machine learning classifiers*



Note. Event Related Potential features = ERP, Event Related Dynamic features = ERD. The green dotted line indicates a baseline level for more optimal predictions.

Supplement 5.1. | *Search terms per database*

Database	Search terms	Filter	Date	Results
PubMed	(transcranial brain stimulation[Title/Abstract] OR transcranial magnetic stimulation[Title/Abstract] OR theta burst stimulation[Title/Abstract] OR transcranial direct current stimulation[Title/Abstract] OR transcranial electrical stimulation[Title/Abstract] OR transcranial alternating current stimulation[Title/Abstract]) AND (callous[Title/Abstract] OR psychopathy[Title/Abstract] OR empathy[Title/Abstract] OR emotional reactivity[Title/Abstract] OR guilt[Title/Abstract] OR prosocial [Title/Abstract] OR altruis*[Title/Abstract] OR cooperation[Title/Abstract] OR helping behavior [Title/Abstract])	English, Humans, Child: birth-18 years, Adult: 19+ years, Young Adult: 19-24 years, Adult: 19-44 years, Middle Aged + Aged: 45+ years, Middle Aged: 45-64 years.	25/11/2024	93
Web of Science Core Collection	(TS=(transcranial brain stimulation or transcranial magnetic stimulation or theta burst stimulation or transcranial direct current stimulation or transcranial electrical stimulation or transcranial alternating current stimulation)) AND (TS=(callous OR psychopathy OR empathy OR emotional reactivity OR guilt OR prosocial OR altruis* OR cooperation OR helping behavior) AND (ALL=((19-44 years OR Adult) OR (45-64 years OR Middle Aged) OR (45+ years OR Middle Aged + Aged))) NOT (TS=(animal))	Refine terms: Languages (English) + Document Type (Article)	28/11/2024	39
Scopus	((TITLE-ABS-KEY ("transcranial brain stimulation" OR "transcranial magnetic stimulation" OR "theta burst stimulation" OR "transcranial direct current stimulation" OR "transcranial electrical stimulation" OR "transcranial alternating current stimulation") AND TITLE-ABS-KEY ("callous" OR "psychopathy" OR "empathy" OR "emotional reactivity" OR "guilt" OR "prosocial" OR "altruis*" OR "cooperation" OR "helping behavior"))) AND (adult) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English"))	N/A	28/11/2024	210

Supplement 5.2. | *Description of behavioural measures across included studies*

	Empathy
Affective Resonance	
Measures	*Description*
Emotional images	Participants commonly view images that are categorised as neutral, positive, negative, and emotionally disturbing. They rate these images based on arousal, valence, and specific emotions they evoke (e.g., sadness, fear, discomfort).
Emotional videos	Participants watch video clips designed to induce specific moods and then reporting their emotional reactions using standardised measures.
Music	Participants rate their emotional reactions to a piece of sad music.
Word-fragment completion task	Participants complete disgust-related words with missing letters. Faster word completion indicates more attention (reactivity) to emotional content.
Handgrip-force task	Participants hold a handgrip while watching videos of crying infants. Increased pressure on the handgrip indicates more reactivity.
Empathy Quotient	A self-report questionnaire designed to measure empathy in adults, including a subscale to assess emotional reactivity.
Empathic Concern	
Measures	*Description*
Pain empathy task	Participants assess their emotional responses to images or videos that depict individuals in distressing situations (e.g., injured children). Ratings often include levels of empathic concern, vicarious pain, and distress.
Interpersonal Reactivity Index	A self-report questionnaire designed to measure dispositional empathy, including a subscale to assess empathic concern.
Affect rating	Participants rate their levels of pity before and after stimulation.
	Prosociality
Charitable giving/altruism	
Measures	*Description*
Dictator Game	In various studies using the Dictator Game, participants are asked to divide a sum of money (e.g., $10) between themselves and another player or recipient, with different variations to explore factors influencing altruistic behaviour and decision-making processes.
Costly helping	Participants choose how much of their own tokens to donate to either reduce someone else's distress or increase others' tokens at a personal cost.
Donation task	Participants decide how to allocate resources – such as money or credits – to various causes or organisations. Their decisions were influenced by the conditions of the task, such as whether they were making donations publicly or privately.
Helping/Cooperation	
Measures	*Description*
Helping tasks	Participants in these tasks are required to make decisions about helping others at no personal cost in various scenarios.
Cooperation and social dilemma tasks	Participants engage in decision-making scenarios where cooperation or defection influenced both personal and group outcomes. These tasks explore

	how individuals balance personal benefits against collective good in various settings.
Prosocial intervention	Participants watch video clips showing different types of interactions (i.e., cooperative, noncooperative, conflictual, and neutral) and then rate how much they empathised with the actors and their intention to intervene in each scenario. Prosocial behaviour is determined when participants intervene more in conflictual/noncooperative scenarios.
Reciprocity	
Measures	*Description*
Trust Game	Trust games involve a trustor who sends resources to a trustee, who then decides how much to return. Some variations explore how the visibility of the trustee's decisions (whether anonymous or revealed) affects their willingness to reciprocate.
Holdup Game	Participants receive money from an investor and decide how much they will give back in return. The level of reciprocity is indicated by the benefit returned relative to the initial investment.
Redistribution Game	Participants are given the opportunity to rectify an unfair distribution made by another party.
Guilt	
Measures	*Description*
Guilt Knowledge Test	Participants rate their feelings of guilt on a scale from 0 (no guilt) to 5 (maximum guilt) after deception while being asked about a crime.
Calgary Depression Scale for Schizophrenia – Factor II	Clinician rated outcome measure that assesses the level of depression in people with schizophrenia. Factor II in this questionnaire evaluates pathological guilt.
Affect rating	Participants rated their levels of altruistic and deontological guilt before and after stimulation.
Callous-Unemotional Traits	
Measures	*Description*
Psychopathic Personality Inventory	Self-report measure of both global psychopathy and the component traits of psychopathy, comprising eight subscales, including one measuring cold-heartedness (i.e., callousness).
Self-Report Psychopathy Scale-Short Form	Self-reported inventory designed to measure psychopathy subdivided into four different facets, two of which include callous traits and interpersonal manipulative style.

Supplement 5.3. | *Considerations for outcome measures*

In analysing affective empathy data, we ensured that emotional responses were compatible with the stimuli's valence, as accurate assessments of affective empathy require congruence between the individual's response and the emotion conveyed by the stimulus. Emotional measures of empathy were thus defined as capturing emotional contagion, reactivity, or concern for others' feelings. Studies focusing on other aspects of emotional processing, such as emotion recognition, theory of mind, or emotional regulation, were excluded for specific reasons. Firstly, emotional understanding and theory of mind relate more closely to the cognitive dimensions of empathy, which are often intact in individuals exhibiting callous-unemotional traits. This distinction is crucial because individuals with callous-unemotional traits can recognise and understand emotions in others yet may still lack the affective response that characterises true empathy (Bird & Viding, 2014). Secondly, while emotional regulation is indeed necessary for responding appropriately to the emotional states of others, the literature highlights that empathic individuals frequently experience emotional dysregulation due to their heightened sensitivity to the feelings of others. This emotional resonance often leads to an overwhelming experience of others' distress, making it difficult for them to regulate their own emotional responses. Thus, studies indicating improved emotional regulation – evidenced by more positive emotions when confronted with negative stimuli – may inadvertently reflect reduced emotional resonance and, consequently, lower affective empathy.

Furthermore, we excluded prosocial tasks that included aggressive elements, such as punishment, recognising that the presence of aggression could confound interpretations of behaviour as genuinely prosocial (Hu et al., 2015; Rodrigues et al., 2024). The literature suggests that aggression can inherently distort empathetic responses by introducing conflicting emotional dynamics, thereby complicating the understanding of the underlying motivations for prosocial behaviour (Eisenberg & Miller, 1987). By focusing solely on non-aggressive contexts, we aimed to clarify the relationship between affective empathy and prosocial behaviour, ensuring a more accurate representation of how empathetic responses manifest in genuinely supportive actions.

Supplement 5.4. | *Exploratory analyses*

A subset of studies combined anodal and cathodal stimulation over ROI either via HD-tDCS (J. Hu et al., 2017; F. Li et al., 2020; Y. Long et al., 2023; Sergiou et al., 2022; Wu et al., 2018; Q. Zhang et al., 2023) and bilateral bipolar tDCS (Brunoni et al., 2013; S. Chen et al., 2019; Fecteau et al., 2013; Lisoni et al., 2024; Rêgo et al., 2015; Snowdon & Cathcart, 2018; Vanderhasselt et al., 2016; G. Wang et al., 2016; J. Wang et al., 2014; Zheng et al., 2021). These protocols involve the simultaneous use of both anodal and cathodal electrodes over functionally relevant brain regions, meaning the direction of stimulation effects cannot be easily categorised as purely excitatory or inhibitory. Bipolar bilateral tDCS targets regions in both brain hemispheres, while HD-tDCS uses a central electrode with anodal or cathodal stimulation and return surrounding electrodes with opposite polarity. Therefore, we used these data for exploratory analyses.

Studies using bipolar bilateral tDCS ($k = 10$, $n = 23$) exhibited some consistency in stimulation parameters and ROI, typically targeting the DLPFC or OFC at stimulation intensities between 1.5–2 mA. However, heterogeneity analyses indicated high between-study variability ($Tau^2 = 1.84$, $I^2 = 95.01\%$, PI [-2.50, 2.94]). Random effects models revealed a non-significant mean effect size of $g = 0.22$ ($CI_{95\%}$ [-0.36, 0.80], $p = .462$). Subgroup analyses revealed significant positive effects for studies using left hemisphere anodal stimulation ($g = 2.69$; $CI_{95\%}$ [0.56, 4.81], $p = .013$) and for those employing stimulation over the OFC ($g = 1.34$; $CI_{95\%}$ [0.06, 2.62]; $p = .039$). Heterogeneity assessments with HD-tDCS studies ($k = 6$, $n = 14$) also indicated high between-study variability ($Tau^2 = 0.50$, $I^2 = 87.36\%$, PI [-1.54, 1.34]). In these studies, the central electrode delivered anodal stimulation in 10 trials and cathodal stimulation in 4 trials. Neither anodal ($g = 0.10$; $CI_{95\%}$ [-0.34, 0.53], $p = .667$) nor cathodal ($g = -0.57$; $CI_{95\%}$ [-1.26, 0.11], $p = .101$) stimulation produced significant pooled effect sizes.

Supplement 5.5. | *Characteristics of TMS studies*

Normative and healthy sample							
Reference	**Design**	**Sample**	**ROI**	**Montage**	**Stimulation**	**Outcome**	**Intervention effect**
Balconi and Bortolotti, 2012	Crossover, Sham and active control	18 (8 men, 10 women), 23.4±2.60	SMA	LF-rTMS	Offline, 1Hz, 120%rMT, 400pulses, 3 sessions	Emotional faces task	Stimulation reduced the affective response to emotional faces
Balconi and Canavesio, 2014	Crossover, Sham and active control	25 (14 men, 11 women), 23.78±1.16	middle DLPFC	HF-rTMS	Online, 10Hz, 120%rMT, 2400pulses, 3 sessions	Helping behaviour	Stimulation increased prosocial intervention in conflictual scenarios
Berger et al. 2017	Crossover, Sham, Single-blind	20 (all women), 23.55±2.58	right DLPFC	HF-rTMS LF-rTMS	Offline, 10Hz/1Hz, 110%rMT, 900pulses, 2 sessions	Emotional reactivity	No significant effects overall
de Wit et al. 2015a	Parallel, Active control, Single-blind	38 (18 men, 20 women), 39.60±11.40	left DLPFC	LF-rTMS	Offline, 1Hz, 110%rMT, 3000pulses, 1 session	Emotional reactivity	No significant effects of LF-rTMS on expressed distress to negative pictures
Gaesser et al. 2019	Crossover, Active control	17 (7 men, 10 women)	right TPJ	LF-rTMS	Offline, 1Hz, 60%MSO, 1020pulses, 2 sessions	Helping intentions	No significant effects
Gallo et al. 2018	Crossover, Sham	18 (12 men, 6 women), 25±7	left S1	HF-rTMS	Online, 6Hz, 90%rMT, 1440pulses, 1 session	Charitable giving	TMS reduced participants' decision to give away reward money

He et al. 2023	Parallel, Sham	117 (57 men, 60 women), 20.38±0.23	right VLPFC	HF-rTMS	Offline, 10Hz, 90%rMT, 438pulses, 2 sessions	Emotional reactivity	Reduced negative feelings to social exclusion scenarios after stimulation
Jansen et al. 2019a	Parallel, Sham, Single-blind	36 (20 men, 16 women), 43.75±10.90	right DLPFC	HF-rTMS	Offline, 10Hz, 110%rMT, 1 session	Emotional reactivity	TMS intensified experienced emotions in response to positive and neutral images
Knoch et al. 2009	Parallel, Sham, Single-blind	87 (all men), 22.6±.31	right DLPFC left DLPFC	LF-rTMS	Offline, 1Hz, 900pulses, 1 session	Reciprocity	Right DLPFC stimulation reduced willingness to reciprocate; no effect in left DLPFC stimulation
Miller et al. 2020	Parallel, Active control	34 (9 men, 25 women), 20.86±2.75	right TPJ	LF-rTMS	Offline, 1Hz, 100%rMT, 1200pulses, 1 session	Emotional reactivity	Stimulation reduced compassion and increased irritation/annoyance to sad video
Möbius et al. 2017	Crossover, Sham	23, 21.5±3.0	left DLPFC	HF-rTMS	Offline, 10Hz, 110%rMT, 1500pulses, 2 sessions	Emotional reactivity	No changes in affective response to sad videos after stimulation
Müller-Leinß et al. 2018	Crossover, Sham	47 (21 men, 26 women), 24.59±3.47	right DLPFC left DLPFC	LF-rTMS	Offline, 1Hz, 110%rMT, 1200pulses, 1 session	Charitable giving	Right DLPFC stimulation decreased fairness; no significant effects in left DLPFC stimulation
Notzon et al. 2018	Parallel, Sham, Single-blind	40 (17 men, 23 women), 26.525±4.75	right DLPFC	LF-rTMS	Offline, 1Hz, 120%rMT, 1800pulses, 1 session	Emotional reactivity	No significant effects
Soutschek et al. 2015	Parallel, Active control	56 (29 men, 27 women), 26.67±4.53	left DLPFC right DLPFC	LF-rTMS	Offline, 1Hz, 110%rMT, 480pulses, 1 session	Cooperative behaviour	Reduced cooperation rates after stimulation to both hemispheres

Strang et al. 2015	Crossover, Sham, Double-blind	17 (all men), 23.5±1.23	right DLPFC left DLPFC	LF-rTMS	Offline, 1Hz, 110%rMT, 900pulses, 3 sessions	Charitable giving	Right DLPFC LF-rTMS reduced transfers; no effect of left DLPFC
W Yu et al. 2023	Parallel, Active control	108 (54 men, 54 women), 20.43±0.32	right VLPFC	HF-rTMS LF-rTMS	Offline, 10Hz/1Hz, 90%rMT/110%rMT, 1170pulses/900pulses, 1 session	Charitable giving Emotional reactivity	LF-rTMS reduced charitable giving; HF-rTMS increased it and led to positive emotions
Christov-Moore et al. 2017	Parallel, Active control	58 (28 men, 30 women), 21.31±0.29	right DLPFC DMPFC	cTBS	Offline, 5 Hz/50 bursts, 80%aMT, 600pulses, 1 session	Charitable giving	Right DLPFC and DMPFC cTBS increased offers
Holbrook et al. 2021	Parallel, Active control	95 (35 men, 60 women), 20±1.41	MPFC (right DLPFC & pre-SMA)	cTBS	Offline, 5 Hz/50 bursts, 80%aMT, 600pulses, 1 session	Sympathy	Increased reported sympathy for both adversarial and affiliative students
Keuper et al. 2018	Parallel, Active control	48 (23 men, 25 women), 21.46±4.25	right DLPFC	cTBS	Offline, 5 Hz/50 bursts, 50%MSO, 600pulses, 1 session	Emotional reactivity	Stimulation reduced negative resonance
Obeso et al. 2018	Parallel, Active control	32 (15 men, 17 women), 23±0.34	right TPJ	cTBS	Offline, 5 Hz/50 bursts, 80%aMT, 600pulses, 2 sessions	Charitable giving	Stimulation reduced monetary self-interest and increased offers
Soutschek et al. 2016	Parallel, Active control	exp1: 43 (24 men, 19 women), 23.10±2.30 exp2: 38 (8 men, 30 women), 24.10±2.90	right TPJ	cTBS	Offline, 5 Hz/50 bursts, 80%aMT, 600pulses, 1 session	Charitable giving	Increased prosocial reward after stimulation

Tei et al. 2021	Crossover, Sham	25 (all men), 26.50±3.90	right TPJ	cTBS	Offline, 5 Hz/50 bursts, 80%aMT, 600pulses, 2 sessions	Cooperative behaviour	No effect on cooperation ratio
Zinchenko et al. 2021	Parallel, Active control	46 (23 men, 23 women), 21.70±2.10	right DLPFC	cTBS	Offline, 5 Hz/50 bursts, 80%rMT, 600pulses, 1 session	Charitable giving	Stimulation increased charitable giving
Clinical sample							
Reference	**Design**	**Sample**	**ROI**	**Montage**	**Stimulation**	**Outcome**	**Intervention effect**
de Wit et al. 2015b (OCD)	Parallel, Active control, Single-blind	43 (21 men, 22 women), 38.4±10	left DLPFC	HF-rTMS	Offline, 10 Hz, 110%rMT, 3000 pulses, 1 session	Emotional reactivity	No significant effects
Jansen et al. 2019b (Alcoholism)	Parallel, Sham, Single-blind	39 (26 men, 13 women), 41.64±8.63	right DLPFC	HF-rTMS	Offline, 10Hz, 110%rMT, 1 session	Emotional reactivity	Stimulation reduced emotional reactivity
Light et al. 2019 (Depression)	Parallel, Sham, Double-blind	19 (7 men, 12 women), 45.21±11.21	left DLPFC	HF-rTMS	Offline, 10Hz, 120%rMT, 3000 pulses, 20 sessions	Emotional reactivity	Stimulation increased empathic happiness
Enticott et al. 2014 (Autism)	Parallel, Sham, Double-blind	28 (23 men, 5 women), 32.20±10.25	Bilateral DMPFC	HF-rTMS	Offline, 5Hz, 110%rMT, 1500 pulses, 10 sessions	Social relatedness Empathic concern	Stimulation reduced social relatedness

Note. Region of Interest (ROI), supplementary motor area (SMA), dorsolateral prefrontal cortex (DLPFC), temporo-parietal junction (TPJ), primary somatosensory cortex (S1), ventrolateral prefrontal cortex (VLPFC), dorsomedial prefrontal cortex (DMPFC), low-frequency repetitive transcranial magnetic stimulation (LF-rTMS), high-frequency repetitive transcranial magnetic stimulation (HF-rTMS), continuous theta burst stimulation (cTBS), Hertz (Hz), resting motor threshold (rMT), active motor threshold (aMT), maximum stimulator output (MSO).

Values in *cursive* represent approximations calculated/extracted from the article information.

Supplement 5.6. | *Characteristics of tDCS Studies*

Normative and healthy sample							
Reference	**Design**	**Sample**	**ROI**	**Montage**	**Stimulation**	**Outcome**	**Intervention effect**
Boggio et al. 2009	Crossover, Sham/active, Double-blind	23 (11 men, 12 women), 21.3 ± 5.6	left M1 left DLPFC	Anodal	Offline, 2mA, 5min, 4 sessions	Emotional reactivity	Stimulation to left DLPFC reduced emotional responses
Clarke et al. 2020a	Parallel, Sham, Double-blind	37, (12 men, 25 women), *23.17±6.77*	left DLPFC	Anodal	Online, 2mA, 20min, 1 session	Emotional reactivity	Reduced emotional reactivity after stimulation
Clarke et al. 2020b	Parallel, Sham, Single-blind	116 (36 men, 80 women), 23.03±7.43	left DLPFC	Anodal	Online, 2mA, 20min, 1 session	Emotional reactivity	tDCS attenuated reactions to negative emotional content
Colombo et al. 2021	Parallel, Sham, Single-blind	40, 19.80±1.56	left PMv	Cathodal	Online, 1.5mA, 20min, 1 session	Emotional reactivity Empathy	Stimulation reduced arousal and increased self-reported empathy levels
Di Bello et al. 2023	Parallel, Sham, Single-blind	93 (17 men, 76 women), 23.98±8.13	right FTL	Anodal	Online, 2mA, 14min, 1 session	Altruistic behaviour	No effects on altruism
Feeser et al. 2014	Parallel, Sham, Double-blind	42 (20 men, 22 women), 28.45±6.65	right DLPFC	Anodal	Online, 1.5mA, 20min, 1 session	Emotional reactivity	No significant effects
Gao et al. 2023	Parallel, Sham	91 (17 men, 74 women), 21.22±2.28	right MFG	Anodal Cathodal	Online, 1.5mA, 20min, 1 session	Emotional reactivity	Increased negativity after anodal; decreased after cathodal

Hao et al. 2021	Parallel, Sham, Single-blind	90 (40 men, 50 women), 20.1±0.07	right TPJ	Anodal Cathodal	Offline, 1.5mA, 20min, 1 session	Charitable giving	Higher offer after anodal tDCS, but no changes after cathodal tDCS
Hao et al. 2022	Parallel, Sham, Single-blind	90 (37 men, 53 women), 21.46±0.10	left TPJ	Anodal Cathodal	Online, 1.5mA, 20min, 1 session	Charitable giving	Higher investment after cathodal tDCS, but no changes anodal tDCS
H Zhang et al. 2022	Parallel, Sham, Single-blind	107 (39 men, 68 women), 20.07 ± 1.55	DMPFC right TPJ	Anodal	Offline, 1.5mA, 20min, 1 session	Altruistic behaviour	tDCS on the DMPFC, but not to the right TPJ, increased altruism
H Zhang et al. 2023	Parallel, Active, Single-blind	71 (33 men, 38 women), 20.77±1.88	right DLPFC	Anodal	Offline, 1.5mA, 20min, 1 session	Altruistic behaviour	No significant effects
H Zheng et al. 2016a	Parallel, Sham, Single-blind	60 (29 men, 31 women), 21.5±0.23	VMPFC	Anodal Cathodal	Online, 2mA, 20min, 1 session	Altruistic behaviour Reciprocity	Anodal tDCS increased altruistic behaviour
H Zheng et al. 2016b	Parallel, Sham, Single-blind	60 (28 men, 32 women), 21.55±0.23	right DLPFC	Anodal Cathodal	Online, 2mA, 20min, 1 session	Altruistic behaviour Reciprocity	No significant effects
J Li et al. 2018	Parallel, Sham	83 (42 men, 41 women), 24.04±2.75	right DLPFC	Anodal Cathodal	Offline, 1mA, 15min, 1 session	Charitable giving	Higher compliance after anodal; lower after cathodal
J Yu et al. 2022	Parallel, Sham	90 (38 men, 52 women), 20.66±0.06	VMPFC	Anodal Cathodal	Online, 1.5mA, 20min, 1 session	Altruistic behaviour	No significant effects

Karim et al. 2010	Crossover, Sham, Double-blind	22 (13 men, 9 women), 25.6±4.9	anterior PFC	Cathodal	Online, 1mA, 13min, 1 session	Guilt	Lower feelings of guilt
Karim et al. 2010b	Crossover, Sham, Double-blind	22 (9 men, 13 women), 24.8±3.9	anterior PFC	Anodal	Online, 1mA, 13min, 1 session	Guilt	Not significant effects
Liao et al. 2018	Parallel, Sham, Single-blind	60 (30 men, 30 women), 20.80±2.56	MPFC	Anodal Cathodal	Online, 2mA, 20min, 1 session	Helping behaviour	Anodal tDCS increased helping behaviour
Maeoka et al. 2012	Crossover, Sham, Single-blind	15 (10 men, 5 women), 22.2±1.4	left DLPFC	Anodal	Offline, 1mA, 20min, 2 sessions	Emotional reactivity	Emotional reactivity task: affective valence
Nihonsugi et al. 2015	Crossover, Sham	22 (13 men, 9 women), 20.5±1.5	right DLPFC	Anodal	Online, 2mA, 15min, 2 sessions	Reciprocity	tDCS increased cooperation and guilt aversion
NTM Chen et al. 2017	Parallel, Sham, Single-blind	48 (15 men, 33 women), 19.58±3.23	left DLPFC	Anodal	Online, 2mA, 20min, 1 session	Emotional reactivity	No significant effect on state anxiety
Ottaviani et al. 2018	Crossover, Sham, Single-blind	37 (12 men, 25 women), 26.78±5.04	left Ins	Anodal	Online, 2mA, 15min, 2 sessions	Emotional reactivity Guilt	tDCS increased reported disgust and pity, but not guilt
Repetti et al. 2022	Parallel, Sham, Single-blind	102, *19.81±2.36*	right TPJ	Cathodal	Online, 2mA, 20min, 1 session	Pain empathy	No significant changes in vicarious pain and empathic concern

Salvo et al. 2022	Crossover, Sham	36 (18 men, 18 women), 22.44±3.3	left Ins	Anodal Cathodal	Online, 2mA, 15min, 3 sessions	Emotional reactivity Guilt	Anodal tDCS increased disgust ratings; Cathodal tDCS decreased disgust
S Chen et al. 2019a	Parallel, Sham, Single-blind	162 (54 men, 108 women), 20.78±0.04	right DLPFC left DLPFC	Anodal Cathodal	Online, 1.5mA, 20min, 1 session	Cooperative behaviour	No significant effects
Szeremeta et al. 2023	Parallel, Sham, Single-blind	101 (34 men, 67 women), 22.57±5.6	left DLPFC	Anodal	Offline, 1.5mA, 20min, 1 session	Emotional reactivity	tDCS increased arousal for positive content and reduced it for negative content
S Wu et al. 2023	Parallel, Sham/active, Single-blind	106 (40 men, 66 women), 20.92±1.65	right TPJ	Anodal	Offline, 1.5mA, 20min, 1 session	Altruistic behaviour	Anodal tDCS increased altruistic propensity
Xu et al. 2021	Parallel, Sham, Single-blind	80 (40 men, 40 women), 19.7±1.68	left DLPFC	Anodal	Online, 1.5mA, 20min, 1 session	Emotional reactivity	Stimulation increased empathic responses
Yang et al. 2021	Parallel, Sham	96 (24 men, 72 women), 21.23±0.10	right TPJ	Anodal Cathodal	Online, 1.5mA, 20min, 1 session	Charitable giving	Anodal increased donation, cathodal decreased it
Y Chen et al. 2021a	Parallel, Sham, Single-blind	180 (78 men, 102 women), 20.3±0.04	VMPFC	Anodal Cathodal	Offline, 1.5mA, 20min, 1 session	Cooperative behaviour	Anodal stimulation decreased cooperation; no effect with cathodal
Y Chen et al. 2021b	Parallel, Sham, Single-blind	189 (92 men, 97 women), 20.2±0.07	VMPFC	Anodal Cathodal	Online, 1.5mA, 20min, 1 session	Helping behaviour	No significant changes in help degree

Liu et al. 2020	Parallel, Sham, Double-blind	55 (30 men, 25 women)	Right LPFC	Anodal Cathodal	Offline, 1mA, 15min, 1 session	Normative behaviour	Anodal tDCS improved normative judgement; cathodal tDCS reduced it
Yuan et al. 2017	Parallel, Sham	64 (38 men, 26 women), 23.57±2.1	MPFC	Anodal	Offline, 1.5mA, 30min, 1 session	Emotional reactivity	tDCS increased emotional arousal
Brunoni et al. 2013	Crossover, Sham	20 (3 men, 17 women), 24.9±3.8	bilateral DLPFC	left and right anodal/cathodal	Online, 1.5mA, 33min, 3 sessions	Emotional reactivity	No significant effects
Fecteau et al. 2013	Parallel, Sham, Double-blind	36 (11 men, 25 women), 21.6±3.8	bilateral DLPFC	left and right anodal/cathodal	Online, 2mA, 20min, 1 session	Psychopathy	No significant effects
G Wang et al. 2016	Parallel, Sham	60 (25 men, 35 women), 22.37±0.08	right OFC, right DLPFC	anodal OFC, cathodal DLPFC	Offline, 2mA, 15min, 1 session	Reciprocity	tDCS increased money transfer
J Wang et al. 2014	Parallel, Sham, Single-blind	27 (9 men, 18 women), 23.6±2.9	right OFC, left DLPFC	left and right anodal/cathodal	Online, 2mA, 5min, 1 session	Pain empathy	No significant effects on self-discomfort
Rêgo et al. 2015	Parallel, Sham, Double-blind	24 (12 men, 12 women), 23±2.57	bilateral DLPFC	left and right anodal/cathodal	Online, 2mA, 15min, 1 session	Pain empathy	Left anodal tDCS decreased negative feelings and arousal
S Chen et al. 2019b	Parallel, Sham, Single-blind	162 (54 men, 108 women), 20.78±0.04	bilateral DLPFC	left and right anodal/cathodal	Online, 1.5mA, 20min, 1 session	Cooperative behaviour	Left anodal tDCS increased cooperation rates

Snowdon and Cathcart, 2017	Parallel, Sham, Single-blind	103, 23.07 ± 5.36	bilateral DLPFC	left and right anodal/cathodal	Online, 1.5mA, 20min, 1 session	Charitable giving	No significant effects
Brunoni et al. 2013	Crossover, Sham	20 (3 men, 17 women), 24.9±3.8	bilateral DLPFC	left and right anodal/cathodal	Online, 1.5mA, 33min, 3 sessions	Emotional reactivity	No significant effects
F Li et al. 2020	Parallel, Sham, Double-blind	102 (55 men, 47 women), 22.64±7.19	right TPJ	Anodal Cathodal HD-tDCS	Online, 2mA, 11.79min, 1 session	Charitable giving	Increased donations after anodal, but not cathodal, tDCS
Hu et al. 2017	Crossover, Sham, Double-blind	114 (39 men, 75 women), 20.77±2.11	right DLPFC right IPL	Cathodal HD-tDCS	Online, 2mA, 18min, 3 sessions	Helping behaviour	Helping behaviour decreased after stimulation
Long et al. 2023	Crossover, Sham/active	30 (all women), 21.38 ± 2.40	right ATL	Anodal HD-tDCS	Offline, 1mA, 20min, 3 sessions	Empathy	Reduced reported emotional empathy
X Wu et al. 2018	Crossover, Sham, Single-blind	23 (6 men, 17 women), 24.39±3.47	right IFG	Anodal HD-tDCS Cathodal HD-tDCS	Online, 1.5mA, 20min, 3 sessions	Affective sharing	No significant changes in self-reported affective sharing
Q Zhang et al. 2023	Parallel, Sham, Double-blind	63 (32 men, 31 women), 19.83±1.16	right DLPFC right VLPFC	Anodal HD-tDCS	Online, 2mA, 20min, 10 sessions	Emotional reactivity	Only right DLPFC HD-tDCS reduced responses to social exclusion
Clinical sample							
Reference	**Design**	**Sample**	**ROI**	**Montage**	**Stimulation**	**Outcome**	**Intervention effect**
Wilson et al. 2021 (Autism)	Crossover, Sham, Double-blind	7 (5 men, 2 women), 26.1±5.71	right TPJ	Anodal	Online, 2mA, 30min, 2 sessions	Empathy	tDCS increased self-reported levels of empathy

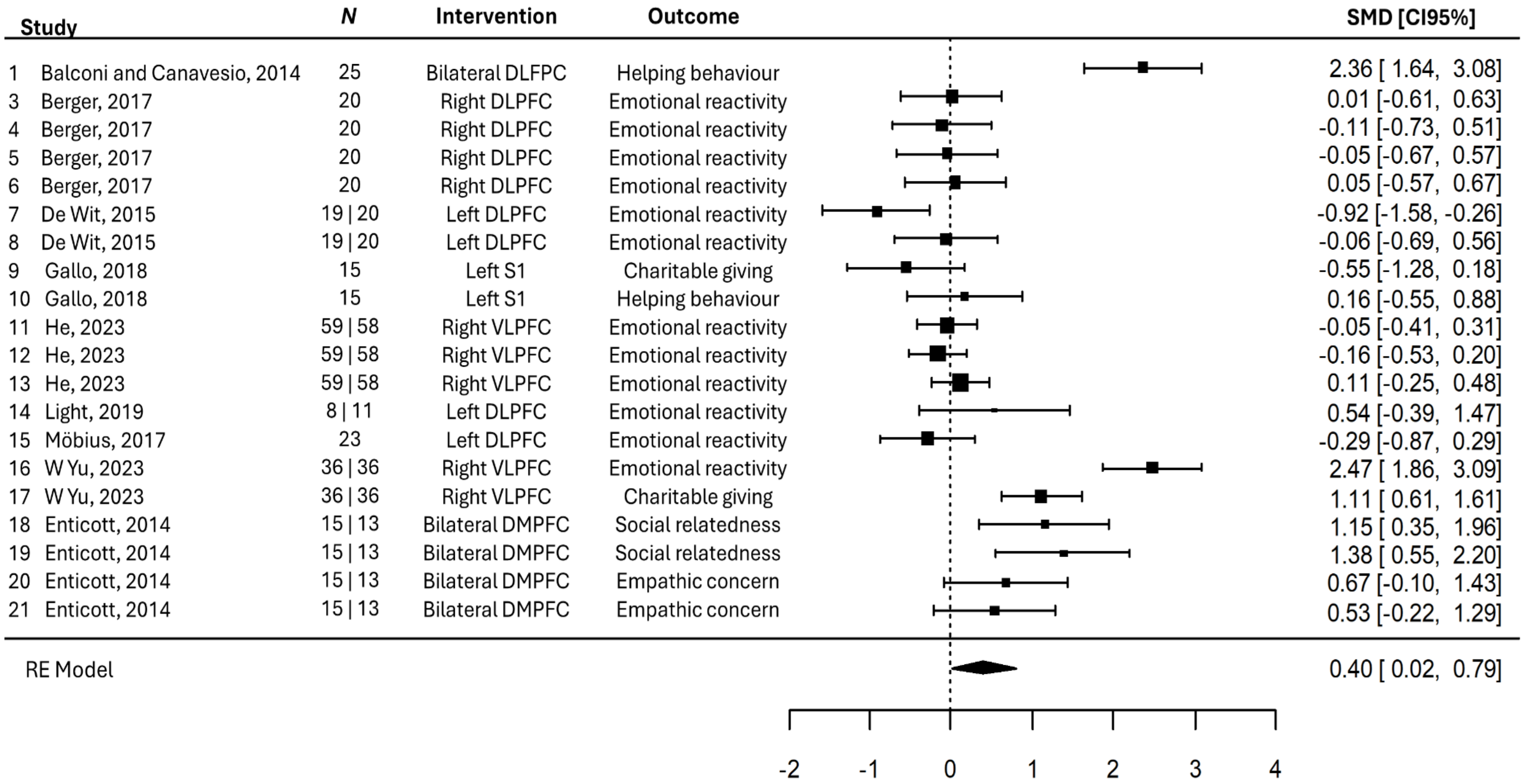
W Zheng et al. 2021	Parallel, Sham, Single-blind	90 (36 men, 54 women), 20.46±0.09	bilateral DLPFC	left and right anodal/cathodal	Online, 1.5mA, 20min, 1 session	Altruistic behaviour	No significant effects
Lisoni et al. 2024 (Schizophrenia)	Parallel, Sham, Double-blind	50 (39 men, 11 women), 42.7±12.17	left DLPFC, right OFC	left anodal/right cathodal	Offline, 2mA, 20min, 15 sessions	Guilt	Stimulation reduced reported guilt
Vanderhasselt et al. 2016 (Depression)	Parallel, Sham	37 (26 men, 11 women), 44.03± 10.75	bilateral DLPFC	left anodal/right cathodal	Offline, 2mA, 30min, 10 sessions	Emotional reactivity	Stimulation increased positive affect and decrease negative affect
Sergiou et al. 2022 (Forensic patients with addiction)	Parallel, Sham, Double-blind	50 (all men), 37.4±9.19	VMPFC	Anodal HD-tDCS	Offline, 2mA, 20min, 10 sessions	Emotional reactivity Empathy Psychopathy	No significant effects overall

Note. Region of Interest (ROI), Prefrontal cortex (PFC), dorsolateral prefrontal cortex (DLPFC), ventral premotor cortex (PMv), medial prefrontal cortex (MPFC), ventromedial prefrontal cortex (VMPFC), ventrolateral prefrontal cortex (VLPFC), orbitofrontal cortex (OFC), temporo-parietal junction (TPJ), medial frontal gyrus (MFG), inferior frontal gyrus (IFG), inferior parietal lobule (IPL), insula (Ins), anterior temporal lobe (ATL), frontal temporal lobe (FTL), primary motor cortex (M1), supplementary motor area (SMA), ventral premotor cortex (VPMC), transcranial direct current stimulation (tDCS), high-definition transcranial direct current stimulation (HD-tDCS), milliampere (mA).

Values in *cursive* represent approximations calculated/extracted from the article information.

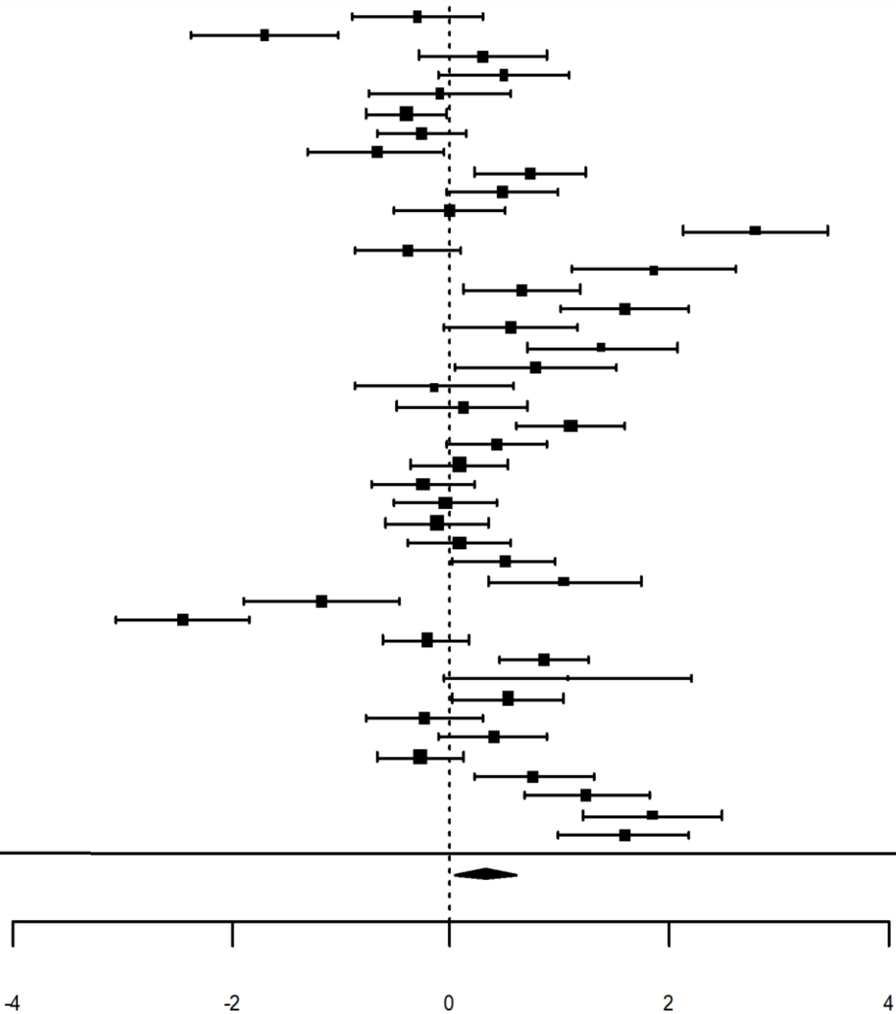
Supplement 5.7. | *Forest plots*

A. HF-rTMS

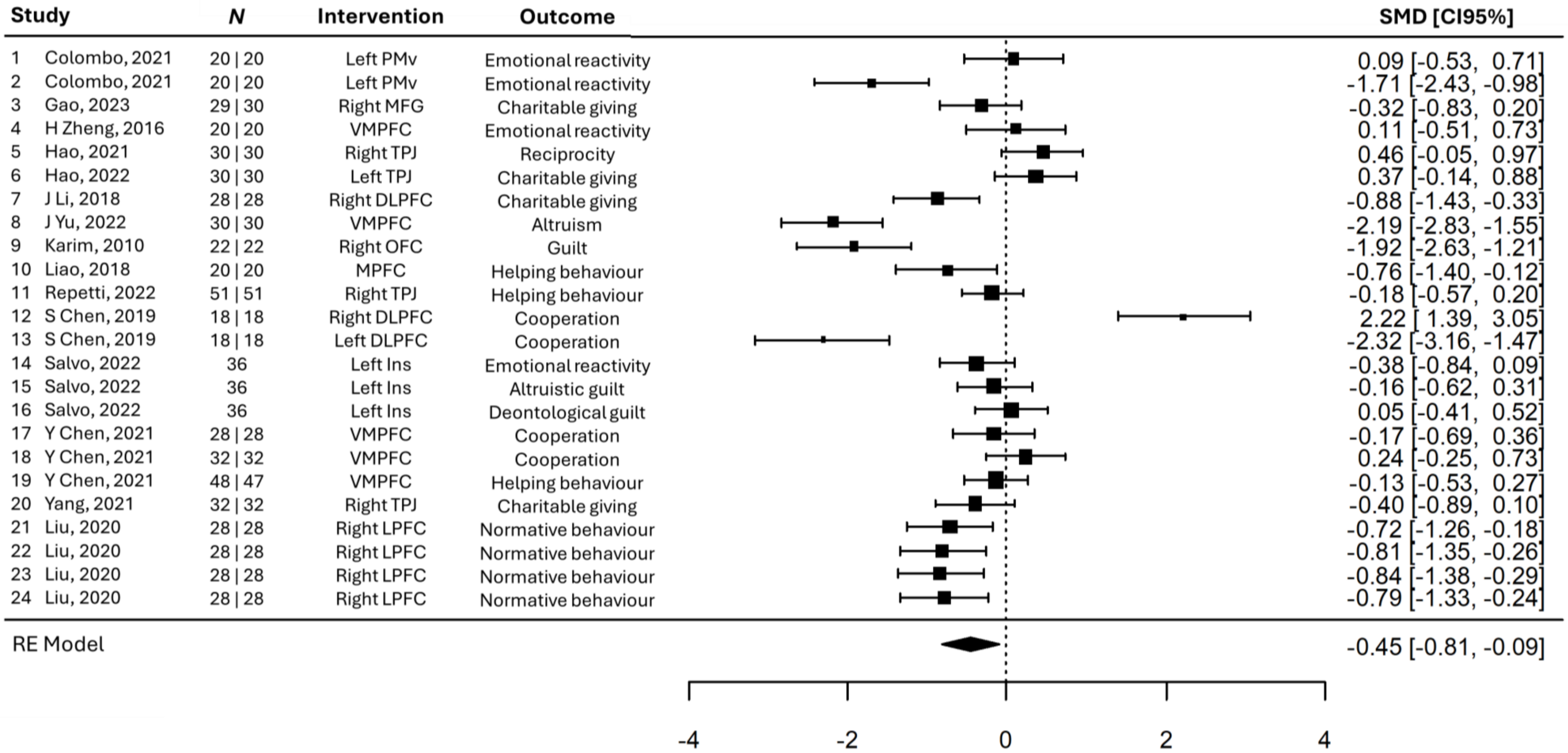


B. A-tDCS

	Study	*N*	Intervention	Outcome	SMD [CI95%]
1	Boggio, 2009	23	Left DLPFC	Emotional reactivity	-0.30 [-0.89, 0.28]
2	Boggio, 2009	23	Left DLPFC	Emotional reactivity	-1.70 [-2.38, -1.02]
3	Boggio, 2009	23	Left M1	Emotional reactivity	0.29 [-0.30, 0.89]
4	Boggio, 2009	23	Left M1	Emotional reactivity	0.49 [-0.11, 1.08]
5	Clarke, 2020	19 | 18	Left DLPFC	Emotional reactivity	-0.10 [-0.74, 0.55]
6	Clarke, 2020	59 | 57	Left DLPFC	Emotional reactivity	-0.41 [-0.78, -0.04]
7	Di Bello, 2023	51 | 46	Right Ins	Emotional reactivity	-0.26 [-0.66, 0.14]
8	Feeser, 2014	21 | 21	Right DLPFC	Emotional reactivity	-0.68 [-1.30, -0.06]
9	Gao, 2023	32 | 30	Right MFG	Emotional reactivity	0.72 [0.21, 1.23]
10	Hao, 2021	30 | 30	Right TPJ	Charitable giving	0.48 [-0.04, 0.99]
11	Hao, 2022	30 | 30	Left TPJ	Charitable giving	-0.02 [-0.53, 0.48]
12	H Zheng, 2022	36 | 33	DMPFC	Altruism	2.78 [2.12, 3.45]
13	H Zheng, 2022	35 | 33	Right TPJ	Altruism	-0.39 [-0.87, 0.09]
14	H Zheng, 2016	20 | 20	VMPFC	Altruism	1.86 [1.11, 2.60]
15	J Li, 2018	27 | 28	Right DLPFC	Charitable giving	0.65 [0.10, 1.19]
16	J Yu, 2022	30 | 30	VMPFC	Altruism	1.59 [1.01, 2.17]
17	Karim, 2010	22	Right OFC	Guilt	0.55 [-0.06, 1.15]
18	Liao, 2016	20 | 20	MPFC	Helping behaviour	1.38 [0.69, 2.07]
19	Maeoka, 2012	15	Left DLPFC	Emotional reactivity	0.78 [0.03, 1.52]
20	Maeoka, 2012	15	Left DLPFC	Emotional reactivity	-0.15 [-0.87, 0.57]
21	Nihonsugi, 2015	22	Right DLPFC	Cooperation	0.11 [-0.48, 0.70]
22	Ottaviani, 2018	37	Left Ins	Emotional reactivity	1.10 [0.61, 1.58]
23	Ottaviani, 2018	37	Left Ins	Emotional reactivity	0.43 [-0.03, 0.89]
24	Ottaviani, 2018	37	Left Ins	Guilt	0.08 [-0.38, 0.53]
25	Ottaviani, 2018	37	Left Ins	Guilt	-0.25 [-0.71, 0.20]
26	Salvo, 2021	36	Left Ins	Emotional reactivity	-0.05 [-0.52, 0.41]
27	Salvo, 2021	36	Left Ins	Emotional reactivity	-0.13 [-0.59, 0.34]
28	Salvo, 2021	36	Left Ins	Guilt	0.08 [-0.38, 0.54]
29	Salvo, 2021	36	Left Ins	Guilt	0.49 [0.02, 0.96]
30	S Chen, 2019	18 | 18	Left DLPFC	Cooperation	1.04 [0.34, 1.73]
31	S Chen, 2019	18 | 18	Right DLPFC	Cooperation	-1.18 [-1.89, -0.47]
32	S Wu, 2023	37 | 35	Right TPJ	Altruism	-2.45 [-3.06, -1.84]
33	Szeremeta, 2023	51 | 50	Left DLPFC	Emotional reactivity	-0.22 [-0.61, 0.17]
34	Szeremeta, 2023	51 | 50	Left DLPFC	Emotional reactivity	0.85 [0.44, 1.25]
35	Wilson, 2021	7	Right TPJ	Dispositional empathy	1.07 [-0.05, 2.19]
36	Yang, 2021	32| 32	Right TPJ	Charitable giving	0.52 [0.02, 1.02]
37	Y Chen, 2021	28 | 28	VMPFC	Cooperation	-0.24 [-0.76, 0.29]
38	Y Chen, 2021	32 | 32	VMPFC	Cooperation	0.39 [-0.10, 0.89]
39	Y Chen, 2021	48 | 47	VMPFC	Cooperation	-0.28 [-0.68, 0.13]
41	Liu, 2020	28 | 28	Right LPFC	Normative behaviour	0.76 [0.22, 1.30]
42	Liu, 2020	28 | 28	Right LPFC	Normative behaviour	1.24 [0.67, 1.82]
43	Liu, 2020	28 | 28	Right LPFC	Normative behaviour	1.84 [1.22, 2.47]
44	Liu, 2020	28 | 28	Right LPFC	Normative behaviour	1.58 [0.98, 2.18]
	RE Model				0.32 [0.05, 0.60]



C. C-tDCS



Note. Forest plots depict intervention effects in trials involving high-frequency rTMS (A), anodal tDCS (B) and cathodal tDCS (C). Targeted brain area for intervention and behavioural outcome for each intervention are presented in the two contiguous columns, with numerical values of the effects (SMD) presented at the right side of each forest plot. References of studies for each extracted effect are presented at the left side of each forest plot. N: number of participants; SMD: standardised mean difference; CI95%: 95% confidence interval; RE model: random effects model.

Supplement 5.8. | *Leave-one-out analyses and influence diagnostics*

HF-rTMS Trials								
Study	**Estimate**	**Pval**	**CI.lb**	**CI.ub**	**Tau2**	**I2**	**Rstudent**	**Cook.d**
1	0.49	0.052	0.00	0.98	1.15	92.86	1.60	0.12
2*	0.40	0.042	0.02	0.79	0.67	88.51	4.06	0.48
3	0.61	0.023	0.08	1.14	1.33	93.67	-0.49	0.01
4	0.62	0.022	0.09	1.14	1.32	93.63	-0.60	0.02
5	0.61	0.022	0.09	1.14	1.33	93.65	-0.54	0.02
6	0.61	0.024	0.08	1.14	1.33	93.69	-0.46	0.01
7	0.66	0.011	0.15	1.16	1.22	93.15	-1.34	0.08
8	0.61	0.022	0.09	1.14	1.32	93.66	-0.55	0.02
9	0.64	0.016	0.12	1.15	1.27	93.50	-0.98	0.05
10	0.60	0.025	0.07	1.13	1.34	93.78	-0.35	0.01
11	0.62	0.022	0.09	1.14	1.33	93.14	-0.55	0.02
12	0.62	0.021	0.10	1.15	1.32	93.09	-0.66	0.02
13	0.61	0.025	0.08	1.14	1.34	93.20	-0.41	0.01
14	0.58	0.030	0.06	1.11	1.34	93.88	-0.04	0.00
15	0.63	0.019	0.10	1.15	1.30	93.52	-0.76	0.03
16	0.48	0.053	-0.01	0.97	1.12	92.54	1.76	0.15
17	0.55	0.040	0.03	1.08	1.33	93.52	0.46	0.01

18	0.55	0.039	0.03	1.08	1.33	93.76	0.48	0.01
19	0.54	0.042	0.02	1.07	1.31	93.69	0.67	0.02
20	0.58	0.033	0.05	1.11	1.35	93.83	0.07	0.00
21	0.58	0.031	0.05	1.11	1.35	93.83	-0.04	0.00

A-tDCS Trials

Study	Estimate	Pval	CI.lb	CI.ub	Tau2	I2	Rstudent	Cook.d
1	0.59	0.04	0.02	1.16	3.50	97.92	-0.47	0.01
2	0.61	0.02	0.08	1.13	2.99	97.58	-1.29	0.03
3	0.58	0.05	0.01	1.16	3.56	97.95	-0.15	0.01
4	0.58	0.05	0.00	1.15	3.56	97.95	-0.05	0.00
5	0.59	0.04	0.02	1.16	3.53	97.94	-0.36	0.01
6	0.60	0.04	0.03	1.16	3.48	97.84	-0.53	0.02
7	0.59	0.04	0.02	1.16	3.51	97.88	-0.45	0.01
8	0.60	0.04	0.04	1.16	3.41	97.87	-0.67	0.02
9	0.57	0.05	0.00	1.15	3.55	97.94	0.08	0.00
10	0.58	0.05	0.00	1.15	3.56	97.94	-0.05	0.00
11	0.59	0.04	0.02	1.16	3.54	97.93	-0.32	0.01
12	0.49	0.06	-0.02	1.01	2.82	97.44	1.32	0.06
13	0.60	0.04	0.03	1.16	3.48	97.89	-0.52	0.02
14	0.54	0.06	-0.02	1.09	3.32	97.82	0.70	0.01

15	0.57	0.05	0.00	1.15	3.56	97.94	0.04	0.00
16	0.55	0.06	-0.02	1.11	3.41	97.86	0.55	0.00
17	0.58	0.05	0.00	1.15	3.56	97.95	-0.02	0.00
18	0.55	0.06	-0.01	1.12	3.46	97.90	0.43	0.00
19	0.57	0.05	0.00	1.14	3.54	97.96	0.11	0.00
20	0.59	0.04	0.02	1.16	3.52	97.94	-0.38	0.01
21	0.59	0.05	0.01	1.16	3.55	97.95	-0.25	0.01
22	0.56	0.05	-0.01	1.13	3.51	97.91	0.28	0.00
23	0.58	0.05	0.00	1.15	3.56	97.93	-0.08	0.00
24	0.59	0.04	0.01	1.16	3.55	97.92	-0.27	0.01
25	0.59	0.04	0.02	1.16	3.51	97.90	-0.44	0.01
26	0.59	0.04	0.02	1.16	3.54	97.91	-0.34	0.01
27	0.59	0.04	0.02	1.16	3.53	97.91	-0.37	0.01
28	0.59	0.04	0.01	1.16	3.55	97.92	-0.26	0.01
29	0.58	0.05	0.00	1.15	3.56	97.93	-0.04	0.00
30	0.56	0.05	-0.01	1.13	3.52	97.94	0.24	0.00
31	0.60	0.03	0.06	1.15	3.25	97.77	-0.96	0.03
32	0.59	0.01	0.12	1.06	2.34	96.92	-1.93	0.01
33	0.59	0.04	0.02	1.16	3.51	97.88	-0.43	0.01
34	0.57	0.05	0.00	1.14	3.55	97.90	0.15	0.00

35	0.56	0.05	-0.01	1.13	3.50	97.94	0.26	0.00
36	0.58	0.05	0.00	1.15	3.56	97.94	-0.03	0.00
37	0.59	0.04	0.02	1.16	3.51	97.91	-0.43	0.01
38	0.58	0.05	0.01	1.15	3.56	97.94	-0.10	0.01
39	0.59	0.04	0.02	1.16	3.50	97.88	-0.46	0.01
40*	0.32	0.02	0.05	0.60	0.78	91.39	10.14	0.75
41	0.57	0.05	0.00	1.14	3.55	97.94	0.10	0.00
42	0.56	0.05	-0.01	1.13	3.49	97.91	0.36	0.00
43	0.54	0.06	-0.02	1.09	3.32	97.81	0.70	0.01
44	0.55	0.06	-0.02	1.11	3.41	97.86	0.55	0.00

C-tDCS Trials

Study	Estimate	Pval	CI.lb	CI.ub	Tau2	I2	Rstudent	Cook.d
1	0.17	0.31	-0.15	0.49	0.28	76.34	0.06	0.00
2	0.12	0.43	-0.18	0.43	0.24	73.48	1.22	0.10
3	0.14	0.38	-0.17	0.46	0.27	75.43	0.68	0.03
4	0.15	0.36	-0.17	0.47	0.28	75.85	0.50	0.02
5	0.14	0.38	-0.18	0.47	0.28	74.32	0.59	0.03
6	0.15	0.37	-0.17	0.47	0.28	74.61	0.47	0.02
7	0.24	0.10	-0.04	0.53	0.20	69.06	-1.84	0.20
8	0.23	0.12	-0.06	0.53	0.22	71.10	-1.55	0.16

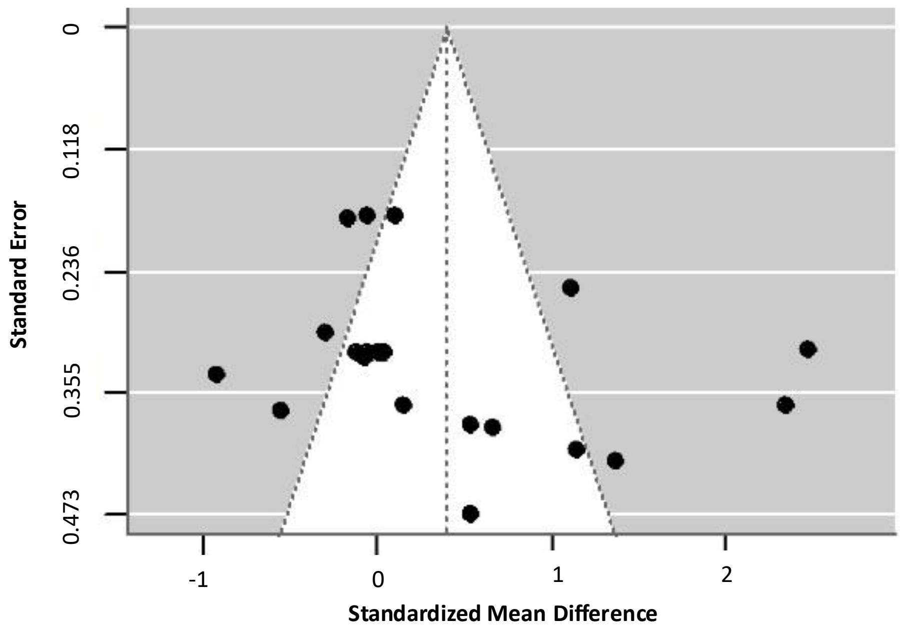
9	0.19	0.24	-0.13	0.51	0.28	76.01	-0.49	0.02
10	0.24	0.11	-0.05	0.52	0.21	70.38	-1.84	0.18
11	0.14	0.39	-0.18	0.45	0.26	75.07	0.87	0.05
12	0.10	0.48	-0.18	0.38	0.19	69.26	2.04	0.21
13	0.19	0.26	-0.14	0.51	0.28	75.86	-0.37	0.01
14	0.17	0.29	-0.15	0.50	0.29	76.21	-0.07	0.00
15	0.19	0.25	-0.13	0.51	0.28	75.86	-0.42	0.01

Note. HF-rTMS: high frequency repetitive transcranial magnetic stimulation; Pval: p value; CI.lb: 95% confidence interval (lower bound); CI.ub: 95% confidence interval (upper bound); Tau2: tau squared value; I2: I squared value; Cook.d: cook distance; A-tDCS: anodal transcranial direct current stimulation; C-tDCS: cathodal transcranial direct current stimulation.

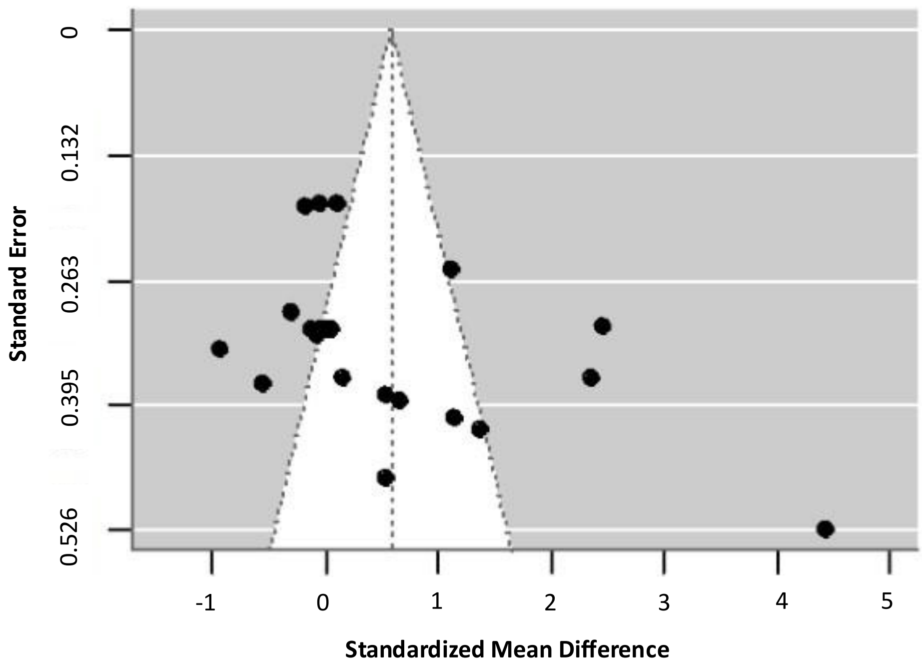
* indicating significant single-study influence on overall effects.

Supplement 5.9. | *Funnel plots*

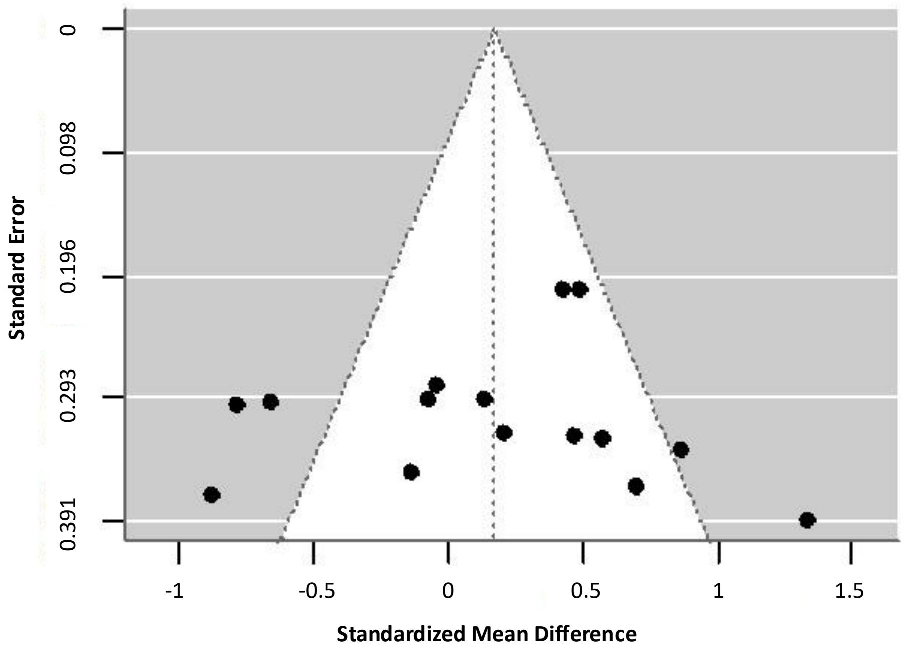
HF-rTMS Trials (After Outlier Exclusion)



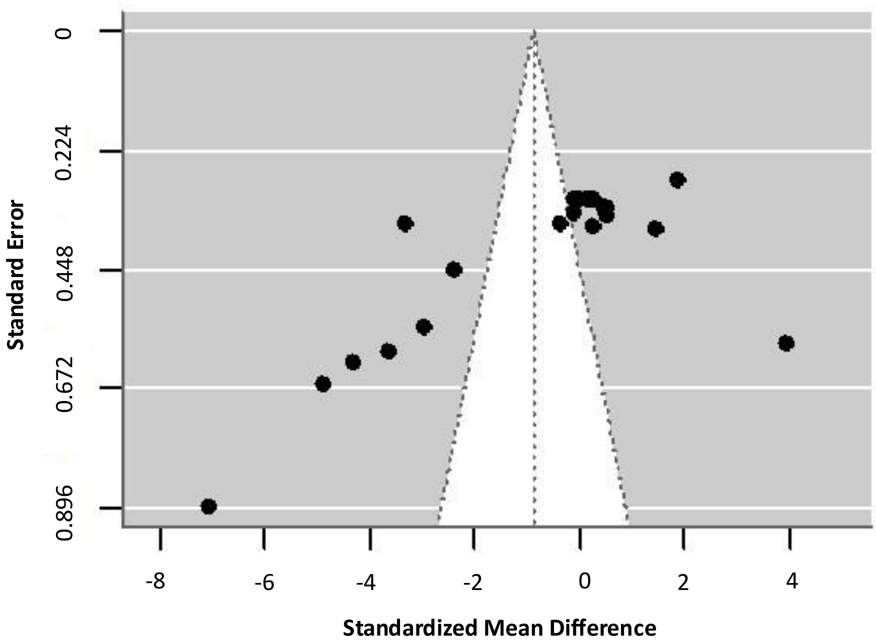
HF-rTMS Trials (Before Outlier Exclusion)



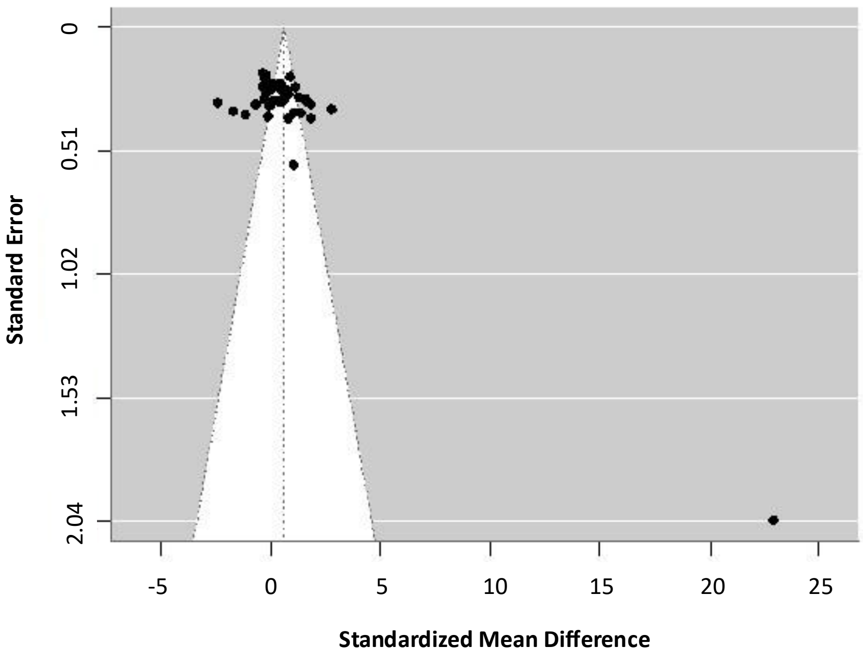
cTBS Trials



LF-rTMS Trials



A-tDCS Trials (Before Outlier Exclusion)



A-tDCS Trials (After Outlier Exclusion)