UNIL | Université de Lausanne

Faculté des sciences sociales et politiques

A Systematic Review of Theory of Mind Abilities in Psychopathy

Autumn Session 2021

Thesis of:
Master of Science in Clinical Psychology and Psychopathology

Presented by: Daulina Ismaili
Supervisor: Valentino Pomini
Expert: Ahmad Abu-Akel

Table of contents

Abstract

Psychopathy is a mental condition characterized by a lack of concern for the well-being of others, manipulativeness, and antisocial behaviours. A controversy in the field is whether the antisocial behaviour of people with psychopathy might be associated with a dysfunction in Theory of Mind (ToM), the ability to understand the cognitive and affective mental states of others. The purpose of the current thesis was thus to bring to sharp focus the state-of-the-art knowledge about the nature of ToM abilities in individuals with psychopathy. To this end, a systematic review between March 2000 and April 2021 was carried for of all peer-reviewed articles published in English that investigated the relationship between ToM abilities and psychopathy in adults. The search yielded 24 articles that met the inclusion criteria. The majority of the studies were almost exclusively in male clinical populations and from Western countries. The results predominantly provide support for intact affective and cognitive ToM abilities in psychopathy. However, further appraisals of the results revealed the importance of considering the different personality and behavioural facets of psychopathy in relation to ToM abilities. Specifically, individuals with higher antisocial lifestyle traits or 'secondary psychopathy' tend to exhibit poorer ToM abilities, whereas those with higher interpersonal/affective traits or 'primary psychopathy' tend to show intact and even superior ToM abilities. This underscores the importance of examining different facets of psychopathy in the assessment of ToM abilities, and that people with psychopathy present considerable individual differences in ToM abilities. Additionally, neuroimaging studies reveal that while the behavioural findings suggest that people with psychopathy have largely intact ToM abilities, the neural findings point to the existence of a distinct neural architecture in individuals with psychopathy compared to controls. This suggests that people with psychopathy rely on a different neural mechanism to represent mental states. These findings have implications for interventions that target ToM abilities in people with psychopathy.

Acknowledgement

I would like to first express my sincere appreciation to Prof. Valentino Pomini for his supervision and directorship of the thesis. I would also like to offer my special thanks to Dr. Ahmad Abu-Akel whose expertise was essential to successfully complete my thesis. My utmost gratitude goes to you for your patience, guidance and insightful feedbacks. Last but not least, I am grateful to my parents, Agim and Flora, for all the unconditional support and encouragement during this intense academic year; without you none of this would be possible.

Introduction

The psychopathy concept refers to a mental condition characterized by a set of co-occurring behavioural and affective traits such as a lack of concern for the well-being of others, callousness, thrill-seeking, manipulative and antisocial behaviours (Hare, 2006). Accordingly, psychopathy is generally conceptualized as a condition that consists of disturbances in interpersonal-affective processes as well as destructive antisocial behaviour. This relatively uncommon condition, making up about 1% of the general population (Hare, 2006), is considered very harmful and costly to society (Neumann & Hare, 2008; Coid et al., 2009). For example, in the United States, it has been reported that psychopathic criminals are responsible for a yearly cost of $2.34 trillion linked to crime, and represent about 16% of the total prison population (Kiehl & Hoffman, 2011 as cited in Drayton et al., 2018). Consequently, considerable attention has been paid to understand the mechanisms of their destructive and antisocial behaviour.

A longstanding notion suggests that this behaviour results from deficits in social-affective processing (Decety et al., 2015). Within this domain, two social-cognitive constructs have received much attention. The first is affective empathy and the second is cognitive empathy. Affective empathy refers to one's ability to recognize the feelings of others and to respond to these feelings with appropriate emotion, while cognitive empathy (also known as *mentalizing*, and most commonly *Theory of Mind* (or ToM) refers to one's ability to represent and understand the affective (i.e., emotions, feelings and desires) and cognitive (i.e., beliefs, intention and knowledge) mental states of others (Feschbach, 1989; Dvash & Shamay-Tsoory, 2014; Vonk et al., 2015). With respect to affective empathy, there appears to be a consensus in the literature that psychopathic individuals present emotional poverty and deficits in affective empathy (Cleckley, 1941/1976, as cited in Sandvik et al., 2014), which appear to originate from impairments in the ability to recognize specific emotions, namely, fear,

sadness and surprise (Marsh and Blair, 2008; Sandvik et al., 2014, Seidel et al., 2013; Deming et al., 2020).

In contrast, a key focus for controversy is whether individuals with psychopathy have intact cognitive empathy/ToM abilities. The lack of consensus regarding their ToM ability might be attributed to the paradoxical nature in which psychopathic individuals behave. On the one hand, psychopathic individuals have been characterized with an enhanced ability to charm and manipulate other people for personal gains (Hare, 2003, as cited in Nentjes et al., 2015b; Sandvik et al., 2014), which is indicative of intact ToM abilities. In agreement with this are several studies demonstrating that psychopathic individuals can detect vulnerable individuals based on affective cues (Book et al., 2007; Wheeler et al., 2009 as cited in Sandvik et al., 2014) and perform well on classic ToM tasks (e.g., Blair et al., 1996; Richell et al., 2003; Dolan & Fullam, 2004; Nentjes et al., 2015a). On the other hand, they are also characterized by traits deemed detrimental for social adjustment such as a lower impulse control and increased irritability in social interactions (Hare, 2003; Patrick et al., 2007 as cited in Nentjes et al., 2015b), which seems to be consistent with another set of studies showing that psychopathic individuals also fail ToM tasks (e.g., Deming et al., 2020; Drayton et al., 2018). With this paradox in mind, it is currently unclear whether individuals with psychopathy have impaired, intact, or superior social understanding capacities.

The present thesis reports on a systematic review of empirical research spanning two decades, from the year 2000 to 2021, with the aim to elucidate whether individuals with psychopathy exhibit ToM deficiencies. Towards this aim, the thesis first provides an overview of psychopathy and ToM, in order to more precisely and fully appreciate their relationship. In Section 1, the thesis presents an overview of psychopathy, including its history, the different terminologies, etiology, and sex differences. In section 2, an overview of ToM is presented, including definitions and terminologies, its development and sex differences. Section 4

provided a detailed description of the employed method for conducting the present systematic review, including the search strategy, study selection and eligibility criteria and the overall results of the search. Subsequently, major findings from the studies included for qualitative analysis are reported in Section 5. Finally, the discussion and conclusion of the results are reported in Section 6 and 7, respectively. Taken together, in this thesis, the systematic review brings to sharp focus the state-of-the-art knowledge about the nature of ToM abilities in individuals with psychopathy in order to shed light on the limitations of existing evidence, and to formulate recommendations that can be instrumental in informing further research on the nature of ToM abilities in people with psychopathy.

1. Psychopathy

1.1 The historical context and evolution of the term psychopathy

Although psychopathy was the first recognized personality disorder in psychiatry (Millon et al., 1998 as cited in Hare, 2006), it remains misconceived at different levels. Indeed, there is ongoing debate regarding the conceptualization, measurement and treatment of psychopathy. It is thus important to explore the history of psychopathy for a better understanding of this disorder. The following paragraphs provide a succinct description of various works that has contributed to our current understanding of the disorder.

Philippe Pinel (1745–1826) is considered among the first to have perceived psychopathy as a mental disorder (Smith, 1978, as cited in Arrigo & Shipley, 2001). In 1801, Pinel noticed that while some of his patients would repeatedly engage in impulsive and self-destructive behaviours (Davies & Feldman, 1981; Millon et al., 1998, as cited in Arrigo & Shipley, 2001), they had their reasoning abilities intact. He concluded that such individuals were suffering from what he called *manie sans délire* (insanity without delirium) (Dinges et al., 1998; Millon et al., 1998, as cited in Arrigo & Shipley, 2001). The postulation of Pinel was very thought-provoking, since it was unusual to consider people with an intact intelligence

and no symptoms of psychosis as 'insane' or 'sick' (Stevens, 1993 as cited in Arrigo & Shipley, 2001).

Pinel's student, Jean Etienne Dominique Esquirol, proposed the two following concepts: *instinctive monomania* and *homocidal monomania* to describe these individuals. He maintained the position that individuals who committed delinquencies and even crimes were afflicted by a disorder of will or more precisely by *instinctive monomania* (Postel & Postel, 1988, as cited in Araujo Gonçalves, 2018). He further proposed that individuals who 'kill' are driven by *homocidal monomania*, meaning that they are influenced by a sort of impulsivity that leads them to commit senseless acts of violence (Postel & Postel, 1988, as cited in Araujo Gonçalves, 2018).

In the early 1800s, the American psychiatrist Benjamin Rush used the term *moral derangement* to describe this population (Rush, 1812, as cited in Kiehl & Hoffman, 2011). He noted that some individuals with intact reasoning skills are capable of demonstrating behaviours that are devoid of any morality (Millon et al., 1998, as cited in Arrigo & Shipley, 2001). In addition, he postulated that the flawed morality of such individuals was due to either a birth defect or caused by some sort of disease (Rush, 1812, as cited in Arrigo & Shipley, 2001). Notably, Rush is known as one of the first in the history of psychopathy to have opted for the social condemnation of this population (Arrigo & Shipley, 2001).

J. C. Prichard is another important scholar who contributed to our modern comprehension of the concept of psychopathy, and particularly through his introduction of the expression *moral insanity*. Prichard defined moral insanity as « a morbid perversion of the natural feelings, affections, inclinations, temper, habits, moral dispositions, and natural impulses, without any remarkable disorder or defect of the intellect or knowing and reasoning faculties, and particularly without any insane illusion or hallucination » (Prichard, 1837/1973, p. 16, as cited in Arrigo & Shipley, 2001). Prichard's view was that an individual suffering from moral

insanity should not be held fully accountable for their crimes or antisocial acts; he deemed them unable to fully control their impulses (Arrigo & Shipley, 2001).

It is thus clear that the aforementioned authors attempted to describe what is known today as psychopathy. It was, however, until late 19th century that the German psychiatrist Julius L. A. Koch (1888) used for the first time the German word *psychopatische,* which means suffering soul, to describe this population (Hervé, 2007, as cited in Kiehl & Hoffman, 2011). He suggested the replacement of Prichard's concept of moral insanity with *psychopathic inferiority* (Sandvik, 2014). With *psychopathic* referring to traits caused by an organic/biological dysfunction and with *inferiority* denoting a deviation from the norm (Millon et al., 1998 as cited in Sandvik, 2014). This concept gained great popularity in early 20th century before being overshadowed for some time by the term sociopathy. The term sociopathy was considered as more suitable and less confusing. It was more suitable since the condition was believed, by the majority at the time, to be a consequence of the environment and less confusing since the lay public would at times confound the term psychopathy with psychosis (Hare, 1999, as cited in Kiehl & Hoffman, 2011). In 1980, when the broader diagnosis of antisocial personality disorder (ASPD) was introduced in the third edition of the Diagnostic Statistical Manual (DSM-III), an ASPD diagnosis was preferred over psychopathy and sociopathy (Kiehl & Hoffman, 2011). It should be noted, however, that psychopathy and sociopathy remain in use to date both in formal and *gray* literature and as such their use remains a source of confusion for the lay public.

Many psychiatrists at the turn of the 20th century were concerned by the fact that psychopathy was viewed through judgmental lenses rather than scientific ones (Kiehl & Hoffman, 2011). However, despite efforts by scholars like Koch to bring to light a more objective description of psychopathy (Hervé, 2007, as cited in Kiehl & Hoffman, 2011), the so called German School of psychopathy did not take into consideration Koch's diagnostic

criteria of psychopathy, and decided to include in the diagnosis of psychopathy anyone presenting symptoms considered as *hors normes* (Kiehl & Hoffman, 2011), and, as a result, the term psychopathy became synonymous with the concept of mental disorder. However, this conception of psychopathy began to change in 1930-1940s with the work of psychiatrists David Henderson and Hervey Cleckley (Kiehl & Hoffman, 2011). Henderson's description of psychopaths as normal and rational, in his 1939 book, *Psychopathic States,* gave impetus to the re-evaluation of the meaning of the word psychopath. In the same manner, Cleckley's book, *Mask of Sanity,* set off an analysis of the existing conception of psychopathy (Kiehl & Hoffman, 2011).

Nevertheless, the attempts of the DSM community to establish a diagnosis of psychopathy was complicated by concerns for diagnostic reliability, and specifically, regarding the ability of clinicians to objectively identify the affective traits in afflicted individuals (Kiehl & Hoffman, 2011). In the first and second edition of the DSM, both affective and behavioural traits were retained (Kiehl & Hoffman, 2011). Indeed, Cleckley was supportive of the description of psychopathy as portrayed in the DSM-II (Cleckley, 1982, as cited in Blackburn, 2007). In DSM-III, however, the category of ASPD retained solely the behavioural components (Kiehl & Hoffman, 2011). The main issue with the ASPD category in the DSM-III was that the disorder, as described, could be easily attributed to individuals who were not psychopaths.

In order to address this issue, the American Psychiatric Association revised the ASPD category in the fourth version of the DSM (DSM-IV) to include certain affective components to the diagnosis. Nonetheless, the DSM-IV has been criticized for underemphasizing the affective features and overemphasizing the behavioural components (Cunningham & Reidy, 1998; cited in Anderson & Kiehl, 2012). Besides, the DSM-IV does not provide any guidance on how to integrate the affective and behavioural components of the disorder (Kiehl &

Hoffman, 2011). Speaking to this issue, Hare (2006) maintains that most psychopathic offenders meet the criteria for an ASPD diagnosis, but very few offenders with ASPD are psychopaths. Indeed, the PCL-R and ASPD are characterized by an asymmetrical association as the prevalence for ASPD among criminal offenders is much higher than the prevalence of psychopathy (Hare, 2006). It seems, therefore, important to avoid using these concepts interchangeably.

It is evident from this historical review that the conceptualization of psychopathy is an evolving subject. It is Cleckley's book, *The Mask of Sanity,* and his characterization of psychopaths, however, that remains most influential for the conceptualization of psychopathy. Indeed, the majority of researchers continue to draw inspiration from Cleckley's book to describe and understand psychopathy (Araujo Gonçalves, 2018).

1.2 Towards a working definition of psychopathy

Given the ambiguity inherent in the DSM, some clinicians set out to develop a working definition of psychopathy. As stated above, Cleckley's characterization of psychopaths provided the main framework for the conceptualization of psychopathy. Cleckley's characterization was inspired by his involvement with several patients in Georgia University Hospital (Sandvik, 2014), which he formulated into 16-item criteria (see Table 1).

Table 1. Cleckley's criteria	
1. Superficial charm and good intelligence	9. Pathological egocentricity and incapacity for love
2. Absence of delusions and other signs of irrational thinking	10. General poverty in major affective reactions
3. Absence of "nervousness" or psychoneurotic manifestations	11. Specific loss of insight
4. Unreliability	12. Unresponsiveness in general interpersonal relations
5. Untruthfulness and insincerity	13. Fantastic and uninviting behaviour with drink and sometimes without
6. Lack of remorse or shame	14. Suicide rarely carried out
7. Inadequately motivated antisocial behaviour	15. Sex life impersonal, trivial, and poorly integrated
8. Poor judgment and failure to learn by experience	16. Failure to follow any life plan

From these criteria, we see that even though Cleckley maintains that psychopaths suffer from a profound emotional deficit (Araujo Gonçalves, 2018), they are characterized with good intelligence, suggesting that their profound emotional deficit does not interfere with the capacity to maintain a "mask of sanity". He used the metaphor of colour blindness to exemplify psychopath's inability to feel with other people. Just as colour-blinded people may appear from the outside as totally capable of viewing the world in colours, psychopathic individuals seem completely capable of experiencing compassion like any other typical individual (Domes et al., 2013). It may be that the characteristics like superficial charm, good intelligence, absence of delusions and nervousness aid psychopaths to hide or camouflage their *insanity* from their surroundings. Psychopathy is thus a condition that can be quite easily hidden compared to most other psychiatric disorders (Cleckley, 1941/1976; Skeem et al., 2011, as cited in Sandvik, 2014). Moreover, it can also be noted from the above criteria that psychopaths were not necessarily perceived by Cleckley as dangerous, violent individuals. In fact, he maintains the following stance concerning the link between psychopathy and criminality:

[t]he typical psychopath, as I have seen him, usually does not commit murder or other offenses that promptly lead to major prison sentence. ... Many people, perhaps most, who commit violent and serious crimes fail to show the chief characteristics which so consistently appear in the cases we have considered (Cleckley, 1941/1976, p. 262 as cited in Sandvik, 2014).

Influenced and inspired by Cleckley's conception of psychopathy as well as those of other clinicians (e.g., McCord and McCord, 1964), the Canadian psychologist Robert Hare recognized the need to develop a reliable assessment tool of psychopathy. In 1980, he developed a 22-item Psychopathy Checklist (PCL; Hare, 1980 as cited in Skeem et al., 2011), which was later revised into the 20-item Psychopathy Checklist-Revised (PCL-R; Hare, 1991

as cited in Skeem et al., 2011). Hare's PCL-R (see Appendix A) provided an alternative view to that of the DSM-III, since it demonstrated that psychopathy could be measured reliably with the inclusion of items account for personality characteristics as well (Skeem et al., 2011).

The PCL-R is a clinical rating scale that uses a semi-structured interview, collateral information (including criminal records and institutional files), and specific scoring criteria to assess whether a person possesses psychopathic traits (Hare, 2006; Bo et al., 2015). The maximum score of the PCL-R is 40, and a cut off score of 30 is typically used for diagnosing someone as a psychopath (Skeem et al., 2011). However, some countries (e.g., European countries) consider a score over 25 as the "cut off score" for the diagnosis of psychopathy (Wynn et al. 2012; Bo et al., 2015). Although the PCL-R measures psychopathy as a unitary construct (Hare & Neumann, 2008, as cited in Skeem et al., 2011), it is primarily composed of two factors: Factor 1 and Factor 2.

Factor 1 (*or primary psychopathy*) is divided into interpersonal and affective facets (Cooke & Michie, 2001; Hare, 2003, as cited in Skeem et al., 2011), whereas Factor 2 (*or secondary psychopathy)* is separated into impulsive-irresponsible lifestyle and antisocial behaviour facets (Hare, 2003, as cited in Skeem et al., 2011). Factor 1 (interpersonal/affective traits) denotes individuals with a lack of remorse, with a tendency to selfishly exploit others, whereas Factor 2 (antisocial lifestyle traits) captures social deviance (e.g., criminality, impulsivity, etc.) (Blackburn, 2007). Factor 1 traits capture relatively well psychopathy as conceptualized by Cleckley, but Factor 2 adds to the initial description. It can therefore be said that Hare did not entirely rely in Cleckley's conception of psychopathy (Blackburn, 2005; Cooke et al., 2004, as cited in Blackburn, 2007). It should also be noted that there have been debates on whether a three or four factor model of psychopathy would be more legitimate (Cooke and Michie, 2001; Cooke et al., 2006 as cited in Wynn et al., 2012).

There is also a shorter version of the PCL-R which was developed as a result of high clinical demands, the Psychopathy Checklist: Screening Version (PCL-SV; Hart, 1995), which requires a score of 18 or more for a diagnosis of psychopathy (as cited in Wynn et al., 2012). Nonetheless, the PCL-R is now the most widely used measure of psychopathy and is considered to offer the best definition of psychopathy to date (Majois et al., 2011, as cited in Araujo Gonçalves, 2018). The initial purpose of the PCL-R was to measure the construct of psychopathy rather than predict or assess the risk of criminal recidivism (Hare, 2006). However, the PCL-R has demonstrated relatively good predictive power of recidivism and violence, making it suitable for forensic assessments (Hare, 2006).

Despite advancement in the conceptualization of psychopathy, popular portrayals of psychopathic individuals continue to result in an erroneous representation of the typical psychopath. The lay public usually think of the psychopath as a cold-blooded, manipulative, sadistic killer, such as the infamous John Wayne Gacy ("the Killer Clown"), a paedophilic serial killer, who is known for the assault and murder of young men and boys. Such portrayal does not capture the complexity of the psychopathy construct, and leads to the stigmatization of this population as serial killers. Thus, it is important to highlight that although psychopathy presents a high risk for violent and unethical modes of conduct (Lykken, 1996, Kiehl & Hoffman, 2011 as cited in Drayton et al., 2018), psychopathic individuals do not necessarily become criminals or present a history of violence (Lilienfeld, 1994), and only a subgroup of criminals fit current standard criteria of psychopathy (Viding, 2019).

1.3 Sex differences

Prior research has reported that males are much more likely than females to receive a diagnosis of psychopathy (e.g., Salekin et al., 1997; Jackson et al., 2002). However, it remains debatable whether such a difference between females and males is of intrinsic nature or a consequence of the existent instruments and conceptualizations of the construct of psychopathy (Wynn et al., 2012), particularly when instruments for the measurement of psychopathy were mainly developed through research conducted on North American Caucasian males (Skeem et al., 2011 as cited in Bueno, 2018).

As pointed above, the diagnostic tools were developed mainly through research on males. Likewise, most individuals diagnosed with psychopathy are males, raising questions whether the origin of such results comes from the failure of existent diagnostic tools to fully capture female psychopathy. Jackson et al. (2002) attempted to elucidate this phenomenon through the investigation of the underlying dimensions of female psychopathy. The authors conducted a series of confirmatory factor analyses of different psychopathy models. The results suggested that Hare's (1991) two-factor model does not capture appropriately female psychopathy.

Instead, Jackson et al. (2002) found that the three-factor model developed of Cooke and Michie (2001), which consisted of the arrogant and deceitful interpersonal style (ADI), deficient affective experience (DAE), and impulsive and irresponsible behaviour factors (IIB) was more sensitive in distinguishing psychopathic and non-psychopathic females. Specifically, Jackson et al. (2002) indicated that the deficient affective experience factor was key in distinguishing a female psychopath from a female non-psychopath, suggesting that an impaired ability to experience empathy was a key feature for the distinction of psychopathic and non-psychopathic females. These findings point out to a potential failure of the two-factor

model of Hare (1991) to account for female psychopathy, which given its widespread use, might be responsible for the underrepresentation of females with a diagnosis of psychopathy.

Extant literature also suggests that there are differences in how psychopathy might manifest itself behaviourally in men and women. For instance, several scholars maintain that psychopathy manifests itself in men with more of an antisocial pattern whereas psychopathy in women present with features that are akin to a histrionic personality disorder (C. G. Cloninger, 1978; Lilienfeld et al., 1986 as cited in Skeems et al., 2011), such that they are presumed to be seductive, theatrical, self-centred and vain. Forouzan and Cooke (2005) also support the notion of differential manifestations of psychopathy in males and females. The authors claims that female manipulativeness is expressed with flirtatious attitudes whereas men will much more likely exhibit conning behaviours.

Understanding female psychopathy is particularly important given indications for an increase in the number of women involved with the criminal justice system. For instance, the number of women arrested in the United States between 1996 and 2005 has increased by 7%, which intriguingly is paralleled with an 8% decrease in the number of men arrested (Federal Bureau of Investigation, FBI, 2005, as cited in Skeems et al., 2011). The results from Robins et al. (2003) further highlight that while there are important contextual differences in relation to the violence committed by men and women, women and men are equally likely to commit violent crimes. These findings call for a necessity to develop robust instruments that can reliably identify psychopathy in women, for clinicians to not ignore women's capacity for violence, and to also take into consideration different contextual elements when dealing with violent women (Robbins et al., 2003) within which psychopathy can be a key construct.

1.4 The etiology of psychopathy

Currently, there is no universally accepted cause of psychopathy. As Viding (2019) best put it: "We should not ask whether a disorder is biological or environmental, we should seek to understand a disorder as a developmental progression of a particular hand of endowments—genetic and environmental—dealt to an individual" (p.155-156). Indeed, several lines of evidence have linked psychopathy to an interaction of multiple factors (i.e., social, environmental and neurobiological) rather than to a single ultimate cause.

Both cross-section and longitudinal studies have demonstrated a link between the emergence of psychopathy and specific social and environmental factors (Viding, 2019). The development of psychopathy has been particularly linked to peer relationships, socioeconomic status and the experience of early adverse, traumatic events such as child abuse, exposure to family conflicts and disorganized attachment style with primary caregivers (Viding, 2019). For instance, a study conducted by Weiler and Widom (1996) demonstrated a link between childhood victimization and higher psychopathy scores as well as between childhood victimization and violence. The authors proposed that psychopathy may serve as a mediator between childhood victimization and violence (as cited in Lang et al., 2002). In contrast, Lang et al. (2002) did not find significant correlation between psychopathy and childhood victimization. Rather, their results suggested that the link between childhood victimization and psychopathy is mediated by psychosocial components. The authors accordingly proposed that there is a specific group of children with enhanced vulnerability for adult psychopathy. That is, victimized children with psychopathic traits are more at risk for developing adult psychopathy as well as violent behaviours than other victimized children.

As pointed above, victimized children do not necessarily develop adult psychopathy. It thus seems plausible to question whether there is also a genetic contribution regarding the emergence of this condition. In addressing this question, Viding (2019) concludes based on

evidence from twin studies that psychopathy has a substantial heritable component, with psychopathic personality features being moderately to strongly heritable and that *psychopathic* antisocial behaviours display a greater heritability than antisocial behaviours alone (i.e., not coupled with psychopathic tendencies).

This genetic vulnerability may contribute to the well-documented brain abnormalities of individuals with psychopathy or psychopathic tendencies. Indeed, with the advent of brain imaging methods, accumulating evidence consistently implicates the amygdala, orbitofrontal cortex, the anterior and posterior cingulate, the parahippocampal gyrus, temporal pole and the insula (for a review see Anderson & Kiehl, 2012). However, due to its well-established role in emotional reactions, instrumental learning, aversive conditioning, aggression and decision-making (Blair, 2003, as cited in Bueno, 2018), the amygdala has been at the centre of inquiry of psychopathy (Anderson & Kiehl, 2012). Numerous studies investigating the role of the amygdala in psychopathy have found reduced amygdala volumes (e.g., Kiehl et al., 2001; Yang et al., 2010; Ermer et al., 2012), as well as decreased amygdala activation in response other people's distress or fear (e.g., Viding et al., 2012).

The evidence for amygdala abnormalities in terms of volume and reactivity is consistent with the low-fear hypothesis (Lykken, 1978, 1995, 2006), which proposes that individuals with psychopathy suffer from what he termed "low fear IQ", as evident by their inability to inhibit responses that led to electric shocks (as cited in Skeem et al., 2011). This is consistent with findings showing that individuals with psychopathy, as compared to healthy controls, display a weaker increase in skin-conductance when anticipating an electric shock (as cited in Skeem et al., 2011). However, when considering the different co-occurring traits of psychopathy, it becomes less clear whether all psychopathic traits can be linked amygdala dysfunction (Bueno, 2018). It could be that amygdala dysfunction accounts for aggressive behaviours in psychopathic individuals, but not for callousness-like traits. In that case,

amygdala dysfunction would fail to explain the etiology of nonviolent psychopathic offenders (Bueno, 2018).

The dual-process model of psychopathy provides a somewhat more complete account of psychopathy than the "low fear hypothesis" (Fowles & Dindo, 2006, 2009; Patrick, 2007; Patrick & Bernat, 2010 as cited in Skeem et al., 2011). This model postulates for the existence of two different processes underlying the interpersonal-affective and antisocial features of psychopathy. More specifically, the model hypothesizes that the process underlying the interpersonal-affective traits is *poorer fear reactivity* whereas the one underlying antisocial feature is *externalizing-propensity* (Skeem et al. 2011). The first process of this model is in accordance with the "low fear hypothesis" as it also hypothesizes innate deficits in fear reactivity, that are putatively linked to abnormalities in amygdala and related brain regions (Skeem et al. 2011). As for the second process—*externalizing-propensity*— it gives insight regarding the origins of the antisocial and reckless behaviours often seen in psychopathic individuals (Patrick et al., 2005 as cited in Skeem et al. 2011), which has been hypothesized to be linked to abnormalities in fronto-cortical brain regions, and particularly the orbital frontal cortex (Skeem et al., 2011).

Taken together, psychopathy is most likely to result from a combination of environmental, genetic and neurobiological factors, which should be considered together to more fully account for the apparent heterogeneity in psychopathy. It should be noted that while the abovementioned findings have highly contributed to our understanding of the disorder, there is still a need for studies to illuminate the current conflicting findings which can be accomplished by conducting research with a combination of diverse assessment methods (Viding, 2019). Consequently, this may help us find a way to manage and treat more effectively this mental condition (Anderson & Kiehl, 2012)

2. Theory of Mind

Theory of Mind (also known as *cognitive empathy* or *mentalizing*) refers to the ability to represent and understand both cognitive and affective mental states of other people as well as our own (e.g., intentions, beliefs, feelings) without necessarily demonstrating or experiencing an emotional response (Lockwood et al., 2013; Ibanez & Manes, 2012). It serves a very important function as it permits individuals to use these mental states—processed explicitly or implicitly (Apperly, 2010; Apperly & Butterfill, 2009)—to predict and interpret people's behaviours and to carry out successful everyday social interactions (Happé, 2003; Drayton et al., 2018) (See Figure 1 for a visualization of the different subcomponent of ToM).
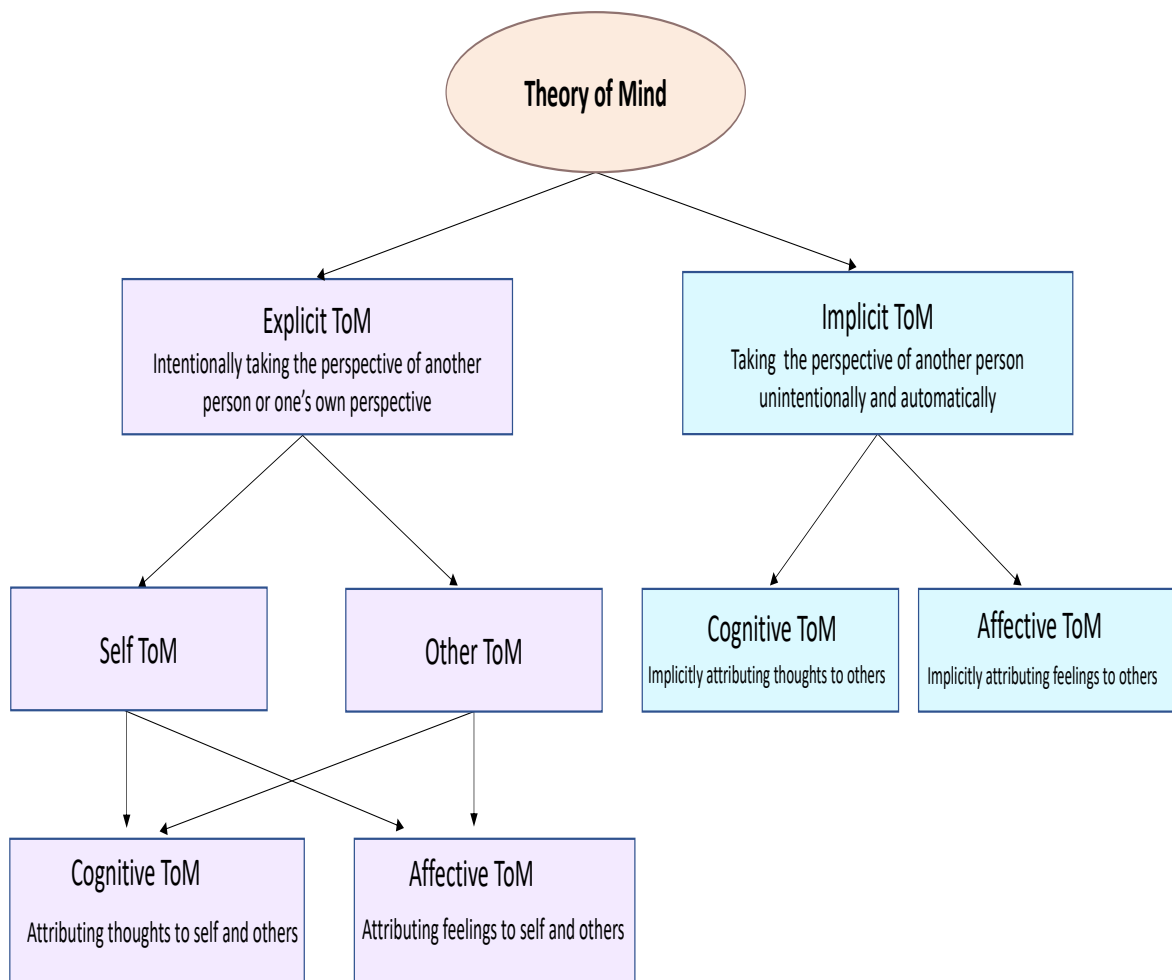

Figure 1 The different subcomponents of Theory of Mind.

2.1 Cognitive and affective theory of mind

A prevailing consensus today is that ToM abilities are no longer considered a monolithic process, but rather a composite of processes that responsible for the representation of cognitive and affective mental states. Indeed, evidence shows that the *cognitive* and *affective components of ToM* are dissociable (Shamay-Tsoory et al., 2009; Shamay-Tsoory et al., 2010; Dvash and Shamay-Tsoory, 2014), and which can be localized to different brain networks (Abu-Akel and Shamay-Tsoory, 2011). According to the theoretical framework of Shamay-Tsoory et al. (2010), "cognitive ToM"—which is synonymous with the cognitive component of empathy—permits individuals to understand and thus make inferences about other people's beliefs whereas "affective ToM" refers to the ability to make inferences about emotions. In contrast, affective ToM seems to integrate the cognitive and to a lesser extent the affective components of empathy. Thus, it is the emotional aspect that aids in the differentiation between cognitive ToM and affective ToM.

Extensive brain research suggests that ToM abilities are subtended by both cortical and subcortical regions that are functionally organized into networks that subserve the ability represent cognitive and affective mental states to both self and other (for a review see Abu-Akel and Shamay-Tsoory, 2011). More specifically, Abu-Akel and Shamay-Tsoory (2011) maintain that the integration of these cortical and subcortical brain regions into a functionally interconnected circuit is necessary for a better understanding of ToM processing and endorsees the following statements: (a) Cognitive and affective ToM are carried out by dissociated yet interacting brain circuits; (b) There are distinct neural bases for representing self and other mental states and the capacity to differentiate between self and other mental states is mediated by an attention system (i.e., the ventral and dorsal attention systems); (c) The dopaminergic- serotonergic (DS) system contributes to our ToM functioning, and specifically in the maintenance and application processes of represented mental states.

According to this model, ToM impairments can thus be of varying degrees, which could be traced to: 1) a disruption in the ability to represent mental states; 2) disruption in the ability to differentiate between self and other mental states; and/or 3) disruption in the ability to apply and execute mental states, which can manifest in the form of either a *hypo-tendency* to apply mental states or as a *hyper-tendency* to attribute erroneous mental states in an uncontrollable manner (Abu-Akel & Shamay-Tsoory, 2011).

2.2 Explicit and implicit theory of mind

From a global processing perspective, ToM can also be distinguished in terms of implicit and explicit processing of mental states (Apperly & Butterfill, 2009; Butterfill & Apperly, 2013). That is, individuals can process the perspective of others in controlled versus automatic ways. Implicit or automatic ToM is executed effortlessly and thus does not require the use of executive functions (e.g., inhibition). In contrast, explicit theory of mind requires executive control processes for successful execution of ToM. For instance, Samson et al. (2010) developed a computerized perspective-taking task which depicts an avatar situated in a room with dotted walls. The task's requirement is for participants to count and report the number of dots they see on the wall. The authors noticed that the participants' responses slow down when the position of the avatar indicates that it sees fewer dots than the participants themselves. Although the participants were not required to consider the avatar's perspective, the results show that the position of the avatar implicitly interfered with their own perspective (as cited in Apperly, 2012).

The findings from Samson et al. (2010) clearly indicate the existence of an automatic, implicit and less cognitively demanding type of ToM processing. A "two-system" account of ToM has been developed as a result of this distinction between implicit and explicit ToM (Apperly, 2010; Apperly & Butterfill, 2009 as cited in Apperly, 2012). This "two-system"

account maintains that humans from every group age (i.e., infants, children, adults) have the capacity for implicit ToM, which is cognitively efficient but inflexible (i.e., utilized only for simple issues) (Surtees et al., 2012 as cited in Apperly, 2012), and another in which only older children and adults exhibit the ability of more cognitively demanding and flexible ToM.

2.3 The developmental stages of theory of mind

Theory of mind follows a developmental trajectory starting from early childhood—with the first traces of ToM abilities already observed during early infancy in the form of, for example, shared attention and finger pointing to indicate to another individual an object of interest (Baron-Cohen, 1995 as cited in Shamay-Tsoory & Aharon-Peretz, 2007)—and well into adolescence and even early adulthood Dummontheil et al., 2010; Apperly et al., 2011). Developmental studies suggest that ToM develops in stages and with increasing level of complexity. It identifies three main developmental levels/stages: first-order ToM (Wellman, 1990), second order ToM (Perner and Wimmer, 1985), and understanding and recognition of faux pas (FP) (Stone et al., 1998). First-order ToM is defined as the capacity to understand that other people can have false beliefs concerning the world. This capacity starts developing around 4-5 years of age (as cited in Abu-Akel & Abushua'leh, 2004). Around 6-7 years of age, second-order ToM or "belief about belief", which refers to the ability of a person to realize that one can have a false belief about the belief of another person, begin to emerge (Abu-Akel & Abushua'leh, 2004). This is followed by the ability to recognize a FP, which develops around the age of 9-11. At this level, individuals can recognize and understand whenever another person lacks awareness about the inadequacy of his/her statement (i.e., saying something out of place) (Stone et al., 1998), and reflect the ability of the individual to integrate both affective and cognitive type of reasoning, that is, knowledge about emotions and knowledge about beliefs.

In addition, substantial research suggests that ToM abilities continue to develop in later adolescence (Dummontheil et al., 2010; Apperly et al., 2011) and that adults present considerable individual differences in ToM abilities and performance (Apperly et al., 2012). The development of ToM abilities well into adulthood is consistent with a substantial body of neuroimaging evidence regarding key ToM brain regions (i.e., medial prefrontal cortex and lateral tempo-parietal regions), which continue to develop into adolescence and adulthood (see Dummontheil et al., 2010). Moreover, Apperly et al. (2012) emphasize the role of individual differences in ToM abilities, which can be explained by distinct executive function abilities (e.g., the ability to resist interference from self-perspective).

Taken together, ToM abilities are not nothing-or-all capacity that follow a protracted developmental trajectory and with considerable individual differences.

2.4 Sex differences in theory of mind

A commonly held stereotype regarding social cognition is that females exhibit better ToM abilities than males. In this context, Baron-Cohen (2009) emphasizes the importance of considering sex differences for a better understanding of social behaviours (as cited in Ibanez et al., 2012) and is a topic of considerable debate and controversy. For example, it has been proposed that sex differences in ToM abilities might be mediated by empathic capacities, suggesting that female's tendency to possess higher empathic abilities could account for their better performance. To address this, Ibanez et al. (2012) investigated whether ToM performance would be better in females and individuals with higher empathic capacities and fluid intelligence, and found that while females were superior to males in ToM abilities, this link was mediated by empathy, suggesting that sex alone was not sufficient to explain higher ToM abilities in females. To explain their findings, the authors invoke the empathizing-systemizing theory according to which, females tend to engage considerably more in empathic thinking whereas males show more interests in systems.

In contrast, Russell et al. (2007) found no support for the stereotype regarding female superiority in ToM abilities. To the contrary, they found superior ToM abilities for men as measured with the Happé's cartoon task (Happé et al., 1999). Here took the authors utilized the empathizing-systemizing theory to elucidate the mechanisms underlying this discrepancy. The male's superior performance may have resulted from a task-specific "systemizing" strategy since Happé's task requires from individuals to understand the rules about how the system functions. However, the authors proceed to question whether male's advantage is dependent on type of ToM. It could be that females are better in tasks measuring affective ToM whereas males in cognitive ToM tasks. Speculatively, this is possible given the study of Ibanez et al. (2012) used the Reading the Mind in the Eyes test, a measure affective ToM, while Russell et al. (2007) employed Happé's task, a measure of cognitive nature. Taken together, there appear indeed differences between males and females in ToM performance, but these differences seem to depend on the component of ToM that is measured by the task.

3. Aim of the study

The aim of the present systematic review is to synthesize and critically appraise existing empirical evidence concerning the relationship between psychopathy and ToM abilities. The conceptualizations of both constructs have been subject of considerable debate, possibly influencing the observed lack of consistency in results. Therefore, the results of this review will bring to light important limitations of existing empirical evidence and hopefully will serve as a resource to inform future research directions.

4. Method

4.1 Search Strategy

A systematic review was conducted following the recommendations outlined in the PRISMA guide (Moher et al., 2009). This systematic review was carried out using empirical studies exploring the relationship between ToM abilities and psychopathy based on all peer-reviewed studies published between March 2021 and April 2021 that met the eligibility inclusion criteria (see Section 4.2 for details on eligibility criteria). To identify relevant studies, the systematic literature search was conducted in PubMed (Medline), Google Scholar, PsycINFO (Ovid) and Wiley Online Library. The search strategy included the term *psychopathy* coupled with the following keywords: *theory of mind, mentalization, mentalizing, social perception, affective theory of mind, cognitive theory of mind, perspective taking,* and *perspective-taking* (see Appendix B for details of search terms and filters).

4.2 Study Selection and Eligibility Criteria

The studies were included if they met the following eligibility criteria: (1) empirical studies that provides original data; (2) the document is peer reviewed; (3) written in English; (4) uses clinical or non-clinical populations; (5) studies that explore ToM in relation to psychopathy. The studies were excluded if: (1) were not empirical studies (e.g., reviews, meta-analyses, book chapters, letters, theoretical articles, commentary etc.); (2) written in another language than English; (3) studies using interchangeably the concept of Psychopathy, Antisocial Personality Disorder, Dissocial personality disorder and Dark triad; (4) did not make use of psychopathy and ToM measures; (5) used all-encompassing and unspecific measures of psychopathy (e.g., Dark Triad tasks); (6) unavailable full texts or abstract-only papers; (7) included participants under the age of 18. There were no imposed restrictions on: (1) sex of participants; (2) study size; or (3) study design.

5. Results

The PRISMA guidelines were followed for the present systematic review (Moher et al., 2010). Initially, the search yielded a total of 2044 articles. After removing duplicates, a total of 1196 records were then subject to the eligibility criteria. Of 1196 articles, 1160 articles were deleted as they did not meet the eligibility criteria. In total, 24 articles were identified as eligible for the qualitative analysis (see Figure 2). A summary of the characteristics and main findings of the studies in clinical populations (N = 18) are summarized in Table 2 and those of the studies in non-clinical populations (N=6) are summarized in Table 3.
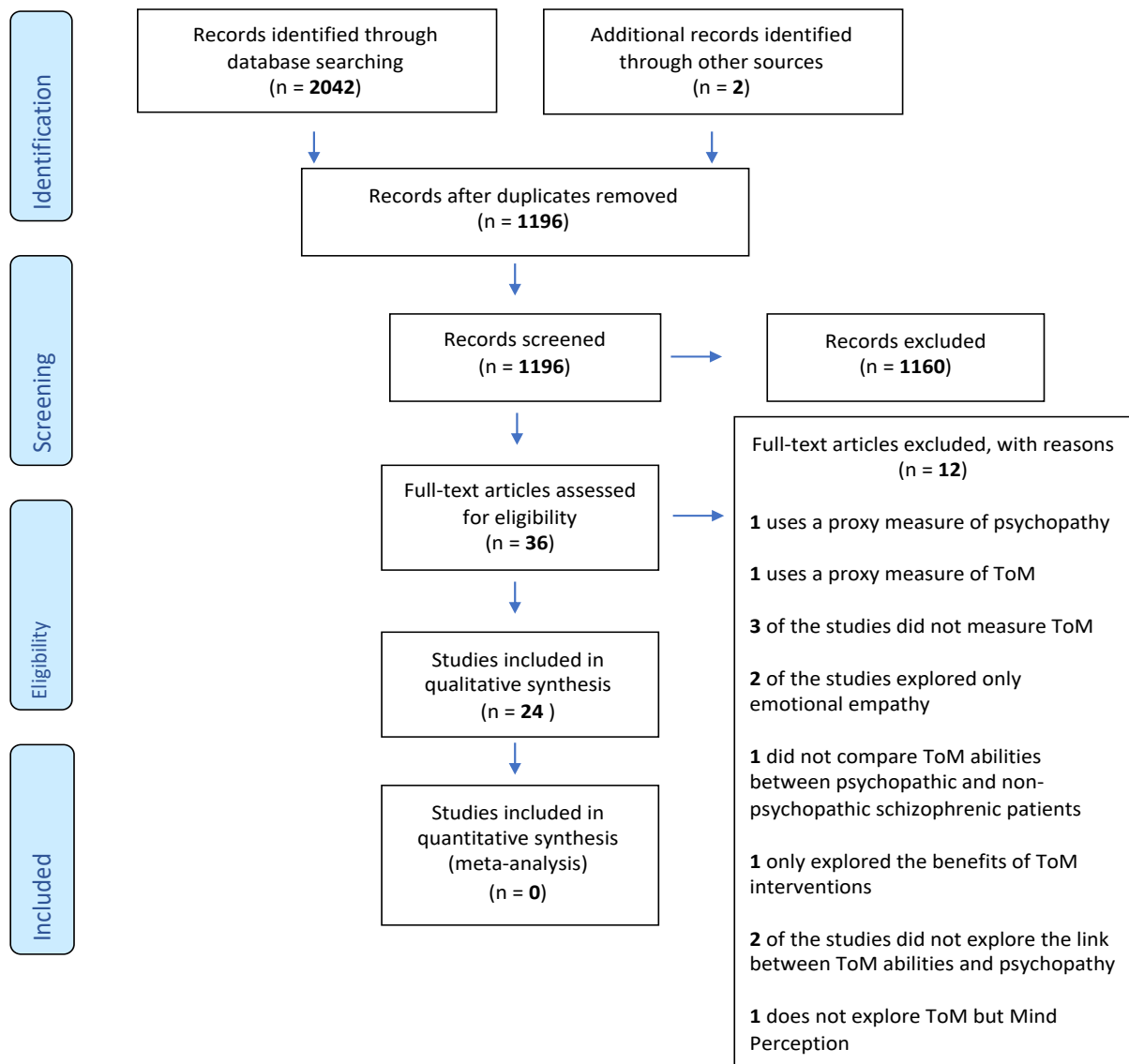

Figure. 2. PRISMA flow diagram (Based on Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009)

Table 2. The association between psychopathy and ToM abilities in clinical populations (2000-2021)

Author (year)	Country	Diagnostic/categories	Sample characteristics: (N, Age, Sex distribution)	Measure of psychopathy	Measure of ToM (affective/cognitive; explicit/implicit)	Main findings
Richell, R. A., Mitchell, D. G., Newman, C., Leonard, A., Baron-Cohen, S., & Blair, R. J. R. (2003)	United Kingdom	Clinical population: Psychopathic group (n= 19) Nonclinical population: Healthy control group(n=18)	a) N=37 b) Age: M=32.7 (SD= 7.4) c) Males (N=37)	PCL-R	**Explicit affective ToM task:** Reading the Mind in the Eyes Test (RMET) - revised	The results indicated intact affective ToM abilities in psychopathic participants
Dolan, M., & Fullam, R. (2004)	United Kingdom	Clinical population (n = 89) : Incarcerated psychopathic ASPD and non-psychopathic ASPD groups Nonclinical population (n=20): Healthy control group	a) N= 109 b) Age: M= 31.8 (SD=) An age range between 18 and 55 years old c) Males (N=109)	PCL-SV	**Explicit cognitive ToM task:** First-order belief tasks **Explicit cognitive ToM task:** Second-order false belief tasks **Explicit cognitive & affective ToM task:** Faux as tasks **Explicit cognitive & affective empathy questionnaire:** The Interpersonal Reactivity Index (IRI) **Explicit affective ToM task:** Facial emotional expression task	-The results revealed no cognitive ToM/perspective-taking differences between psychopathic, non-psychopathic and healthy participants -This study found no deficit in the recognition of Faux Pas in psychopathic ASPDs. However, there was a marked deficit in the ability to understand the emotional impact such a Faux Pas can have on a person -Psychopathic ASPDs did not display difficulties compared to healthy controls in reading basic and complex mental states from facial expressions (e.g., distress, guilt, arrogant)
Sommer, M., Sodian, B., Döhnel, K., Schwerdtner, J., Meinhardt, J., & Hajak, G. (2010)	Germany	Clinical population: Criminal psychopathic patients and criminal non-psychopathic patients	a) N= 28 b) Age Criminal psychopathic patients: M= 31.4 (SD= 7.9) Criminal non-psychopathic patients: M= 29.2 (SD= 5.8) c) Males (N= 28)	PCL-R	**Explicit affective ToM task:** Cartoon stories task (3 conditions: intention fulfilled, intention not fulfilled, non-mentalizing reality condition) fMRI scanning during the cartoon stories task	-Psychopathic participants showed no affective ToM deficits in the behavioural task -The fMRI task revealed that psychopathic and non-psychopathic individuals showed different brain activities during an affective ToM task. Psychopathic participants showed increased activity in the mirror neuron system, whereas non-psychopathic participants showed increased activity in regions linked to attention and monitoring

Author (year)	Country	Diagnostic/categories	Sample characteristics: (N, Age, Sex distribution)	Measure of psychopathy	Measure of ToM (affective/cognitive; explicit/implicit)	Main findings
Shamay-Tsoory, S. G., Harari, H., Aharon-Peretz, J., & Levkovitz, Y. (2010)	Israel	Clinical population: 1) Psychopathy group 2) Lesion groups (OFC group, dlPFC group, NF group) Nonclinical population: healthy control group	a) N= 64 b) Age -Psychopathy group (N=17): M= 29.82 (SD= 10.09) -Lesion groups (N=27): 1) OFC group: M=39.22 (SD= 14.87) 2) dlPFC group: M= 35.55 (SD= 8.56) 3) NF group: M= 40.50 (SD=17.89) -Healthy group (N=20): M= 27.70 (8.36) c) Males (N=64)	LSRP III SRP-II	**Explicit cognitive & affective ToM task:** First- and second-order cognitive and affective ToM task **Explicit cognitive & affective empathy questionnaire:** The Interpersonal Reactivity Index (IRI)	Psychopathic individuals as well as those with OFC lesions exhibited impaired 'affective ToM' but intact cognitive 'ToM'. Therefore, psychopathic individuals are similar to patients with OFC lesions in their ToM capacities
Brook, M., & Kosson, D. S. (2013)	USA	Adult male inmates	a) N= 103 b) Age: M=31.31 (SD=8.00) c) Males (N=103)	PCL-R	**Explicit affective ToM task:** Empathic accuracy task **Explicit cognitive & affective empathy questionnaire:** The Interpersonal Reactivity Index (IRI)	-Psychopathy scores were inversely related with empathic accuracy task performance. That is, individuals with higher psychopathy scores exhibited poorer affective ToM abilities. - Poor performance on the empathic accuracy task was most strongly linked to the Factor 2 (Antisocial Lifestyle) traits of psychopathy -Factor 1 PCL-R traits were correlated with poor empathic accuracy for the emotion of joy whereas Factor 2 PCL-R for negative emotions
Decety, J., Chen, C., Harenski, C., &. Kiehl, K. A. (2013)	USA	Incarcerated males: 1) High psychopaths (n=37): PCL-R > 30 2) Medium psychopaths (n= 44): PCL-R 21-29 3) Low psychopaths (n=40): PCL-R < 20	a) N= 121 b) Age (between ages 18 & 50) - High psychopathy group: M=32.5 (SD= 7.8) - Medium psychopathy group: M= 34.1 (SD=7) - Low psychopathy group: M= 34.6 (SD= 6.9) c) Males (N=121)	PCL-R	**Explicit affective ToM task:** -Participants underwent fMRI scanning while viewing stimuli depicting bodily injuries -Participants received perspective-taking instructions (imagine-self & imagine- other) during the fMRI task	Compared to controls, individuals with psychopathy showed typical activity in the pain-affective neural circuits during the imagine-self perspective condition. However, psychopaths compared to controls exhibited an atypical brain activity during imagine-other perspective in effective connectivity between frontal regions and the amygdala and anterior insula
Domes, G., Hollerbach, P., Vohs, K., Mokros, A., & Habermeyer, E. (2013)	Germany	Clinical population (n= 90): 1) Low psychopathy (n=29): -PCL-R < 15 2) Medium psychopathy (n=33) -PCL-R 15- 21 3) High psychopathy (n=28) -PCL-R > 21 Non-clinical population (n=28): non-criminal healthy control	a) N= 118 males b) Age -PCL-R < 15: M= 48.8 (SD= 14.1) -PCL-R 15-21: M= 47.8 (SD= 10.4) -PCL-R > 21: M= 49.5 (SD= 11.4) c) Males (N=118)	PCL-R	**Explicit cognitive & affective empathy questionnaire:** The Interpersonal Reactivity Index (IRI) **Explicit cognitive empathy questionnaire:** Empathy Quotient (EQ) **Explicit cognitive empathy questionnaire:** Multifaceted Empathy Test (MET) **Explicit affective ToM task:** The Reading the Mind in the Eyes Test	Compared to healthy controls, criminal offenders exhibited impaired emotional and cognitive empathy regardless of psychopathy levels

Author (year)	Country	Diagnostic/categories	Sample characteristics: (N, Age, Sex distribution)	Measure of psychopathy	Measure of ToM (affective/cognitive; explicit/implicit)	Main findings
Seidel, E. M., Pfabigan, D. M., Keckeis, K., Wucherer, A. M., Jahn, T., Lamm, C., & Derntl, B. (2013)	Austria	Clinical population (n=30): incarcerated violent offenders Non-clinical population (n=30): healthy controls	a) N= 60 b) Age -Offender group: M=35.6 (SD= 12.5) -Control group: M=34.8 (SD= 10.2) c) Males (N=60)	PCL-R	**Explicit affective ToM task:** Emotional perspective-taking task **Explicit cognitive & affective empathy interview:** The Interpersonal Reactivity Index (IRI) **Emotion recognition task**	-There was no correlation between higher psychopathy scores, IRI scores and emotion recognition performance. That is, cognitive perspective-taking and emotion recognition abilities were not related to psychopathy levels. However, higher psychopathy scores correlated with reduced accuracy in affective responsiveness - The violent offenders exhibited reduced accuracy in affective responsiveness (i.e., emotionl empathy) with increasing Factor 1 scores. In contrast, violent offenders with higher Factor 2 scores displayed poorer affective perspective taking abilities (i.e., ToM)
Bo, S., Abu-Akel, A., Kongerslev, M., Haahr, U. H., & Bateman, A. (2014)	Denmark	Patients meeting the ICD-10 diagnostic criteria for schizophrenia	a) N= 108 b) Age: M=36.8 (SD=11.2) c) Males (N=87) Females (N=21)	PCL-R	**General explicit ToM interview:** The Metacognitive Assessment Scale abbreviated version (MAS-A)	The results show that higher PCL-R scores can explain the propensity of patients with schizophrenia to engage in premeditated aggression, and that this relationship is mediated by a mentalizing profile of intact cognitive and deficient emotional mentalizing abilities
Mier, D., Haddad, L., Diers, K., Dressing, H., Meyer-Lindenberg, A., & Kirsch, P. (2014)	Germany	Clinical population: Male inmate subjects with psychopathy (n=11) Nonclinical population: Male control subjects (n=18)	a) N=29 b) Age: M= 44.55 (SD=8.97) c) Males (N=29)	PCL-R PPI (German version)	**Explicit task of emotion recognition, affective ToM & neutral face processing:** A social cognitive task of neutral face processing, emotion recognition, and affective ToM during fMRI scanning	-The behavioural findings demonstrate an intact affective ToM in psychopaths -The neural findings demonstrate reduced activity in inferior prefrontal gyrus, superior temporal sulcus and amygdala in psychopaths during affective ToM, and reduced connectivity between left amygdala and right STS. The reduced amygdala activation during affective ToM in psychopaths may originate from reduced input from the mirror neuron system (MNS)
Möller, C., Falkenström, F., Holmqvist Larsson, M., & Holmqvist, R. (2014)	Sweden	Clinical population: Male offenders	a) N= 42 b) Age: M= 20.1 (SD= 0.7; range= 18 to 21) c) Males (N= 42)	PCL-SV	**General explicit ToM interview:** Mentalizing ability was assessed by rating 42 Adult Attachment Interviews with the Reflective Functioning (RF) scale and with Crime-Specific Reflective Functioning Interview	Reflective functioning scores were not related to the level of psychopathy as measured by the PCL-SV
Sandvik, A. M., Hansen, A. L., Johnsen, B.H., & Laberg, J. C. (2014)	Norway	Clinical population: Male offenders	a) N= 92 b) Age: M= 33.47 (SD= 10.77) c) Males (N=92)	SRP-III PCL-R	**Explicit affective ToM task:** A computerized version of the Reading the Mind in the Eyes Test- revised (RMET)	-Inmates with higher Factor 2 (antisocial and impulsive lifestyle) exhibit poorer affective ToM abilities those with higher Factor 1 (interpersonal/affective) traits. -Inmates with Factor 1 psychopathic traits exhibit an enhanced affective ToM
Decety, J., Chen, C., Harenski, C. L., & Kiehl, K. A. (2015)	USA	Criminal male offenders: 1. High psychopaths (n=38): PCL-R > 30 2. Medium psychopaths (n= 67): PCL-R 21-29 3. Low psychopaths (n=50): PCL-R < 20	a) N= 155 b) Age High psychopaths M=32.4 (SD= 6.6) Medium psychopaths M=33.1 (SD=7.6) Low psychopaths M=31.6 (SD=7) c) Males (N=155)	PCL-R	**Explicit affective ToM task:** fMRI scanning while viewing short clips and identifying the emotional states of the protagonists	-Individuals who scored higher in the PCL-R exhibited superior affective ToM abilities -Individuals with high psychopathy compared to those scoring low exhibited an increased neural response in the right pSTS/TPJ, dmPFC, dlPFC, temporal pole, and ACC during an affective ToM task

Author (year)	Country	Diagnostic/categories	Sample characteristics: (N, Age, Sex distribution)	Measure of psychopathy	Measure of ToM (affective/cognitive; explicit/implicit)	Main findings
Nentjes, L., Bernstein, D. P., Arntz, A., Slaats, M. E., & Hannemann, T. (2015a)	Netherlands	Clinical population: 1. Psychopathic offenders (n=39) PCL-R > 25 2. Nonpsychopathic offenders (n=37) PCL-R < 25 Nonclinical population: Nonoffender controls (n=26)	a) N= 102 b) Age -Psychopathic offenders: M= 38.6 (9.7) -Nonpsychopathic offenders: M= 39.3 (SD=10.2) Nonoffenders: M= 35.6 (13.5) c) Males (N=102)	PCL-R LSRP	**Explicit & implicit measure of affective ToM:** Two different Reading the Mind in the Eyes Task (RMET) conditions: -RMET short (i.e., time restriction) -RMET long (i.e., no time restriction)	-Psychopathic individuals exhibited intact ToM during the short RMET condition as well as in presence of hostile answering alternatives - During the long RMET condition, psychopathic individuals were better at distinguishing hostility in eyes in comparison to nonoffender controls. This superior ability to distinguish hostility was related to both Factor 1 (interpersonal and affective) and Factor 2 (antisocial lifestyle) traits of psychopathy
Nentjes, L., Bernstein, D., Arntz, A., van Breukelen, G., & Slaats, M. (2015b)	Netherlands	Clinical population: 1. Psychopathic offenders (n=42) PCL-R > 25 2. Non-psychopathic offenders (n=40) PCL-R < 25 Nonclinical population: Healthy control group (n=26)	a) N= 108 b) Age Psychopathic offenders: M= 39.1 (SD= 9.5) Non-psychopathic offenders: M=38.8 (SD=9.9) Healthy control group: M=35.6 (SD=13.5) c) Males (N=108)	PCL-R	**General explicit ToM task:** Happé's Advanced Test of ToM	-The results indicate that psychopathic offenders, non-psychopathic offenders, and healthy individuals do not differ in terms of performance in Happé's Advanced Test of ToM -The results indicate that ToM deficiencies were not associated with Factor 1 traits nor Factor 2 traits of psychopathy
Abu-Akel, A., Heinke, D., Gillespie, S. M., Mitchell, I. J., & Bo, S. (2015)	Denmark	Patients meeting the ICD-10 diagnostic criteria for schizophrenia from a forensic facility	a) N=79 b) Age: M=36.86 (SD= 10.37) c) Males (64) Females (15)	PCL-R	**General explicit ToM interview:** The Metacognitive Assessment Scale abbreviated version (MAS-A)	Metacognition deficiencies of schizophrenia patients tend to be less pronounced among those who exhibit high psychopathy scores
Drayton, L.A., Santos, L. R., & Baskin-Sommers, A. (2018)	USA	1. Psychopathic offenders: PCL-R > 30 2. Non-psychopathic offenders: PCL-R < 20	a) N= 106 b) Age: M=36.32 (SD=10.70) c) Males (N=106)	PCL-R	**Explicit & implicit cognitive ToM task:** A computer-based perspective-taking task	Psychopathic individuals may lack the ability to automatically take the perspective of others. However, they possess an intact capacity to deliberately take another's perspective
Deming, P., Dargis, M., Haas, B. W., Brook, M., Decety, J., Harenski, C.,… & Kosson, D.S. (2020)	USA	Adult male inmates	a) N= 94 b) Age (between ages 18 & 55): M= 32.9 (SD=8.2) c) Males (N=94)	PCL-R	**Explicit affective ToM task:** -Affective perspective-taking task -fMRI scanning during affective perspective-taking	-The behavioural findings indicate a diminished capacity in psychopaths during affective perspective-taking for fear, happiness, and sadness -The neural findings indicate diminished activity in key empathy brain regions solely during fear-perspective taking

Notes. PCL-R = Psychopathy Checklist-Revised; PCL-SV= Psychopathy Checklist: Screening Version; LSRP III = Levenson Self-Report Scale (Version III); SRP-II = Self-Report Psychopathy Scale; SRP-III Self-Report Psychopathy Scale; PPI (German version) = Psychopathy Personality Inventory (German version); N= Number of participants; M= Mean; SD= Standard Deviation

Table 3: The association between psychopathy and ToM abilities in non-clinical populations (2000-2021)

Author (year)	Country	Population	Sample characteristics: (N, Age, Sex distribution)	Measure of psychopathy	Measure of ToM (affective/cognitive; explicit/implicit)	Main findings
Mullins-Nelson, J. L., Salekin, R. T., & Leistico, A. M. R. (2006)	USA	Undergraduate students	a) N= 174 b) Age: M= 19.34 (SD=1.81) c) Males (N=44) Females (N=139)	PPI-SF	**Explicit cognitive & affective empathy questionnaire:** The Interpersonal Reactivity Index	-Psychopathy as a unitary construct was characterized by intact perspective-taking abilities and deficient affective empathy - Higher scorers of interpersonal/affective facets of psychopathy had intact perspective-taking and affective empathy abilities whereas higher scorers of the antisocial lifestyle facets of psychopathy had deficient perspective-taking and affective empathy modules
Ali, F., & Chamorro-Premuzic, T. (2010)	United Kingdom	Non-clinical population	a) N =112 b) Age: - M= 20.74 (SD= 5.85) c) Females (N= 92) Males (N=20)	LSRP	**Explicit affective ToM task:** Faces Test **Explicit affective ToM task:** Reading the Mind in the Eyes Test (RMET)- revised **Explicit affective ToM task:** Reading the Mind in the Voice Test (RMVT)	-Primary and secondary psychopathy were associated with general ToM deficits as well as ToM deficits to specific emotions (i.e., neutral and positive mental states) - Primary psychopathy exhibited more ToM impairments than secondary psychopathy
Lockwood, P. L., Bird, G., Bridge, M., & Viding, E. (2013)	United Kingdom	Healthy participants	a) N= 110 b) Age: M= 21.9 (SD= 3.7) c) Males (N=55) Females (N=55)	SRP-4-SF	**Explicit cognitive ToM task:** Theory of mind animations task	-High psychopathic traits were related to reduced affective resonance but not cognitive perspective-taking -The results indicate intact cognitive perspective-taking when a purely cognitive task is used
Gillespie, S. M., Mitchell, I. J., & Abu-Akel, A. M. (2017)	United Kingdom	Healthy student population	a) N= 55 b) Age: M= 20 (SD=2.59) c) Males (N=16) Females (N=39)	LSRP	**General explicit ToM task:** The Movie for the Assessment of Social Cognition (MASC)	- Higher psychopathic tendencies coupled with autism traits were associated with a decline in cognitive ToM performance -Higher psychopathic tendencies coupled with positive psychotic experiences were associated with a better cognitive ToM performance -Psychopathic tendencies were positively associated with worse affective ToM performance

Author (year)	Country	Population	Sample characteristics: (N, Age, Sex distribution)	Measure of psychopathy	Measure of ToM (affective/cognitive; explicit/implicit)	Main findings
Jameel, L., Vyas, K., Bellesi, G., Crawford, S., & Channon, S. (2019)	United Kingdom	Sub-clinical population: 1) high-scoring and low-scoring autistic trait group 2) high-scoring and low-scoring psychopathic trait group	a) N= 79 b) Age Psychopathic trait group: M=21.53 (SD= 3.96) Autistic trait group: M=20.53 (SD=2.70) c) -20 high-scoring (10 male) and 19 (9 male) low-scoring autistic trait group -21 high-scoring (11 male) and 19 low-scoring (9 male) psychopathic trait group	PPI-SF	**Explicit cognitive & affective empathy questionnaire:** The Interpersonal Reactivity Index (IRI)	The results provide evidence of intact cognitive empathy but reduced affective empathy with higher psychopathic traits
Lanciano, T., & Curci, A. (2019)	Italy	Healthy participants	a) N=736 b) Age: M=21.79; SD= 5.03 75.7% females	PPI-R (Italian version)	**Explicit cognitive empathy task:** Perspective Taking subscale of Interpersonal Reactivity Index (IRI)	-The results indicate that the relationship between psychopathy and perspective-taking varies as a result of different psychopathic traits. More precisely, the self-centred impulsivity and cold-heartedness factors of the PPI-R were associated with poor cognitive ToM, whereas the fearless dominance factor of the PPI-R was associated with high cognitive ToM abilities. -The findings suggest that individuals with psychopathic traits are characterized by overall cognitive ToM/perspective-taking deficiencies

Notes. PPI-SF = Psychopathic Personality Inventory- Short Form; LSRP = Levenson Self-Report Psychopathy Scale; SRP-4-SF = Self-Report Psychopathy Scale-Short Form; PPI-R (Italian version) = Psychopathic Personality Inventory-Revised (Italian version); N= Number of participants; M= Mean; SD= Standard Deviation

5.1 Characteristics of studies

All 24 studies were published between 2003 and 2020. The majority of the included studies (n=16) reported results exclusively on male participants, while the remaining 8 studies included both male and female participants. The sixteen studies including solely male participants used clinical populations, whereas among the eight studies including both sexes, six of them were conducted in healthy populations. Therefore, a total of eighteen studies were in clinical samples (n= 1453; Mean age (SD) = 34.67 (6.03)) and a total of six studies were in healthy samples (n=1266; Mean age (SD) = 20.80 (1.00)). The sample size across the clinical studies varied and ranged from 29 to 155 (Mean sample size (SD) = 85.47 (36.60)). The sample size across the non-clinical studies varied more widely and ranged from 55 to 736 (Mean sample size (SD) = 211 (260.29)). About half of the studies were conducted in the UK (n=6) and the U.S. (n=6). The remaining studies were conducted in the following countries: Germany (n=3), the Netherlands (n=2), Denmark (n=2), Italy (n=1), Sweden (n=1), Austria (n=1), Israel (n=1), and Norway (n=1). It should be noted that the two studies conducted in Netherlands (Nentjes et al., 2015a; Nentjes et al., 2015b) were both conducted at the same time by the same authors as part of a larger research project and thus utilized the same sample.

In addition, there were several different instruments used for the measure of psychopathy. The most commonly used measure was the Psychopathy Checklist-Revised (PCL-R; Hare, 1991; Hare, 2003) (n=15), followed by the Levenson Self-Report Psychopathy Scale (LSRP; Levenson et al., 1995) (n=5). It should be noted that one of the studies employed the LSRP-III version of the task (Shamay-Tsoory et al., 2010). The included studies also used the Self-Report Psychopathy scale (SRP; Hare, 1985) (n=3). The SRP was utilized in three different versions; That is, SRP-II (Hare, 1991), SRP-III (Jones & Paulhus, 2010; Paulhus, Neumann & Hare, in press) as well as SRP short form (SRP-4-SF; Paulhus et al., in press). Finally, the

remaining studies employed the following tasks: Psychopathy Checklist: Screening Version (PCL-SV; Hart et al. 1995) (n=2), Psychopathic Personality Inventory-Short Form (PPI-SF; Lilienfeld & Hess, 2001; Lilienfeld, 2004) (n=2), the Italian version of the Psychopathic Personality Inventory-Revised (PPI-R; La Marca et al. 2008; Lilienfeld and Widows, 2005) (n=1) and the German version of the Psychopathy Personality Inventory (PPI; Eisenbarth and Alpers, 2007) (n=1).

Moreover, the studies used a variety of ToM measures that varied in complexity and included both explicit (n =33) and implicit (n =2) measures of ToM, as well as measures of cognitive (n = 13) and affective (n = 14) components of ToM. Some studies included ToM instruments tapping both cognitive and affective mental states (i.e., General ToM tasks) (n= 6). They also varied in mode of presentation such that ToM was measured through self-report questionnaires (n = 8), experimental paradigms (n = 19), as well as through guided interviews (n = 3). The most used self-report-questionnaire was the Interpersonal Reactivity Index (n=8) (IRI; Davis, 1980; Davis, 1983), which measures cognitive and affective empathic abilities including cognitive ToM abilities. With respect to experimental paradigms, the most common measure was the Reading the Mind in the Eyes Test (n=6) (RMET; Baron-Cohen et al., 2001), which taps affective ToM. Notably, the version of the employed RMET varied from study to study, and one study (Ali & Chamorro-Premuzic, 2010) employed the Reading the Mind in the Voice Test (RMVT; Golan et al., 2007), in which participants are required to infer mental states from voice. In addition, the medium of presentation of the ToM measures varied widely with some measures requiring the participants to infer the mental states from images, text (e.g., short stories), cartoons and pictures, as well as video clips and animations.

5.2 Findings from studies in clinical populations

Summary of the findings from the studies in clinical populations are presented under three headings. The first (section 5.2.1) summarizes the behavioural findings, the second (section 5.2.2) summarizes findings from neuroimaging studies, and the third (section 5.2.3) summarizes findings from studies that investigated performance in controlled versus automatic mode of ToM processing.

5.2.1 Behavioural findings

As stated above, a total of 18 out of 24 studies were conducted in clinical populations. However, since the study of Decety et al. (2013) did not provide behavioural findings, it is not included in this section. 10 out of the 17 studies (or 59%) report intact ToM abilities in individuals with psychopathy (Richell et al., 2003; Dolan & Fullam, 2004; Sommer et al., 2010; Domes et al., 2013; Seidel et al., 2013; Mier et al., 2014; Decety et al., 2015; Nentjes et al., 2015a; Nentjes et al., 2015b; Drayton et al., 2018). There is also evidence providing support for superior ToM abilities in psychopathy (Sandvik et al., 2014; Decety et al., 2015; Nentjes et al., 2015b).

Two of the seven remaining studies explored the relationship between psychopathy— measured with the PCL-R (Hare, 2003)—and ToM abilities—measured with the metacognition assessment scale-abbreviated (MAS-A; Lysaker et al., 2005)—in schizophrenic patients: Abu-Akel et al. (2015) demonstrating an advantage in ToM abilities among schizophrenic patients with higher psychopathy scores, and Bo et al. (2014) showed that schizophrenic patients with higher psychopathy scores are characterized by intact cognitive but impaired affective mentalizing abilities.

The remaining five studies reported impaired ToM capabilities. Three studies reported affective ToM/perspective-taking deficits in individuals with higher scores of psychopathy (Shamay-Tsoory et al., 2010; Brook & Kosson, 2013; Deming et al., 2020). In Deming et al.

(2020), individuals with higher psychopathy scores exhibited impaired affective perspective-taking abilities for specific emotions, which included fear, happiness, and sadness. In Shamay-Tsoory et al. (2010), the results showed that individuals with psychopathy, as measured by the LSRP III and the SRP-II (Hare, 1991), as well as individuals with OFC lesions were impaired in *affective* but not *cognitive* ToM. It was concluded that the pattern of ToM deficits in psychopathy resembles that seen in participants with lesions of the frontal lobe, particularly with OFC damage. Brook and Kosson (2013) reported an inverse relationship between psychopathy scores (PCL-R; Hare, 1991) and affective ToM performance on a modified version of the Empathic Accuracy Task (Ickes, 1997; Brook & Kosson, 2013).

In addition, using a battery of self-report measures, Domes et al. (2013) reported impaired emotional and cognitive empathy in criminal offenders regardless of psychopathy levels (PCL-R; Hare, 2003), which suggests that impaired cognitive empathy in this population cannot be attributed to the psychopathy construct. Likewise, Möller et al. (2014) reported impaired perspective-taking abilities in a sample of criminal offenders regardless of psychopathy levels (PCL-SV; Hart et al. 1995). Finally, Sandvik et al. (2014) examined the relationship between affective ToM (RMET; Baron-Cohen et al., 2001) and two different types of psychopathy assessment methods, a clinical (PCL-R; Hare, 2003) and a self-report (SRP-III; Jones & Paulhus, 2010; Paulhus, Neumann & Hare, 2009). While the authors found no significant correlation between the total score of the PCL-R and affective ToM, there was an overall negative association between affective ToM and the total score of the SRP-III.

It should be noted that five out of the total of 17 studies took into consideration the multidimensional nature of psychopathy (Seidel et al., 2013; Brook & Kosson, 2013; Sandvik et al., 2014; Nentjes et al. 2015a; Nentjes et al., 2015b). The findings from Seidel et al. (2013) and Sandvik et al. (2014) are consistent as both studies indicated reduced accuracy in

affective ToM with higher PCL-R Factor 2 traits (i.e., antisocial lifestyle traits) but not with higher PCL-R Factor 1 traits (i.e., Interpersonal/Affective traits). Sandvik et al. (2014) even found superior affective ToM capacity for individuals with higher Factor 1 traits. Besides that, Sandvik et al. (2014) showed that while the interpersonal and affective traits of the self-report measure of psychopathy (SRP-III; Jones & Paulhus, 2010; Paulhus, Neumann & Hare, in press) did not predict ToM performance, the antisocial lifestyle traits of the measure predicted worse performance on the accurate identification of negative and neutral mental states but not of positive mental states. As for Brook & Kosson (2013), the authors found that poor performance on the empathic accuracy task was most strongly linked with PCL-R Factor 2 traits. The Factor 2 traits of psychopathy were associated with poorer accuracy for negative emotions whereas Factor 1 personality features for the emotion of joy.

Finally, Nentjes et al. (2015a) investigated the relation between ToM abilities (RMET; Baron-Cohen et al., 2001) and psychopathy (PCL-R; Hare, 2003). In this study, the RMET was administered under two different conditions: In the first condition participants performed the RMET under time restriction (henceforth, the RMET short condition); and in the second the participants performed the RMET with no time restriction (henceforth, the RMET long condition). That is, the stimulus was either presented quickly or for an unlimited amount of time. The results indicated that during the long RMET condition, psychopathic individuals were better at distinguishing hostility in eyes in comparison to non-offender controls. This superior ability to distinguish hostility was carried by both Factor 1 (interpersonal and affective) and Factor 2 (antisocial lifestyle) traits. In a separate study, but using the same sample as in the RMET study (Nentjes et al., 2015a), Nentjes et al. (2015b) administered the Happé's Advanced Test of ToM (Happé, 1994) and found no differences in performance between psychopathic offenders, non-psychopathic offenders, and healthy individuals. The

results also indicated that performance was not associated with either Factor 1 or Factor 2 traits of psychopathy.

Taken together, individuals with psychopathy tend to demonstrate an unimpaired, and even superior capacity to reason about other people's mental and emotional states. Nonetheless, the evidence providing support for ToM impairments in psychopathy (Shamay-Tsoory et al., 2010; Brook & Kosson, 2013; Deming et al., 2020) should not be neglected as it may shed light to the current controversies surrounding the nature of ToM in psychopathy.

5.2.2 Findings from fMRI studies

Of the total 24 studies, five studies (see Table 4) examined the neural responses of psychopathic and non-psychopathic participants while performing ToM/perspective-taking tasks (Sommer et al., 2010; Decety et al., 2013; Mier et al., 2014; Decety et al., 2015; Deming et al., 2020). All five studies were performed in all male clinical populations, except Mier et al. (2014) who included an additional sample of healthy males for comparison. Furthermore, all included studies measured psychopathy with the PCL-R (Hare, 2003); however, Mier et al. (2014) also used the Psychopathic Personality Inventory (PPI; Eisenbarth & Alpers, 2007) for measuring psychopathy in the healthy control group. Four out of the five studies reported behavioural findings, with three of the four studies reporting intact affective ToM/perspective-taking abilities (Decety et al., 2015; Mier et al., 2014; Sommer et al., 2010). In contrast, Deming et al. (2020) reported that psychopathic individuals experienced difficulties in identifying specific emotions (i.e., fear, happiness, sadness) during an affective perspective-taking task (Haas et al., 2015). Decety et al. (2013) gave perspective-taking instructions during fMRI scanning, but no behavioural data was reported. Nonetheless, and as summarized in Table 4, all five studies found neural activity abnormalities in individuals with psychopathy compared to controls.

Table 4: The brain activity of psychopathic participants during affective perspective-taking

Study	Brain regions	The role of brain regions
Sommer et al. (2010)	Increased activity of the temporo-parietal junction, superior temporal sulcus, medial prefrontal cortex, orbitofrontal cortex	Implicated in monitoring and attention
Decety et al., (2013)	*Self-ToM*: Increased activity of the anterior midcingulate cortex, anterior insula, inferior frontal gyrus, somatosensory cortex, right amygdala coupled with orbitofrontal cortex *Other- ToM*: Decreased activity of the anterior insula, ventromedial prefrontal cortex, and the amygdala with the orbitofrontal cortex	Implicated in the capacity to respond to and reflect about mental states
Mier et al. (2014)	Decreased activity of the amygdala, inferior prefrontal gyrus and superior temporal sulcus as well as reduced connectivity between superior temporal sulcus and amygdala.	Implicated in embodied simulation
Decety et al., (2015)	Increased activity of the right posterior superior temporal sulcus/temporo-parietal junction, amygdala, insula, ventral striatum, and prefrontal cortex.	Implicated in intention and emotion understanding
Deming et al. (2020)	Decreased activity in left anterior insula, left posterior orbitofrontal cortex, precuneus and visual cortex	Implicated in empathy

Sommer et al. (2010) asked male participants to indicate the emotional state of a protagonist whose intention was either fulfilled or not fulfilled. The behavioural findings revealed no ToM/perspective-taking deficits in psychopaths as compared to non-psychopaths. The neural findings, however, found differences between the two groups. Non-psychopathic participants exhibited increased activity in brain regions implicated in the mirror neuron

system (MNS) (i.e., bilateral supramarginal gyrus and superior frontal gyrus), which is putatively involved in simulating the mental states of others. Psychopathic participants, on the other hand, presented increased activity in brain regions implicated in monitoring and attention (i.e., temporo-parietal regions (e.g., superior temporal sulcus), medial frontal cortex, and orbitofrontal cortex). The authors conclude that the two groups make use of disparate brain regions when reasoning about other people's emotional states. Affective ToM in psychopaths is ensured using more rational processes whereas in non-psychopaths via the simulation of other's mental states.

Decety et al., (2013) investigated neural responses in psychopathic and non-psychopathic males during an affective perspective-taking task. In this task participants viewed images of bodily injuries and were asked to adopt either an imagine-self or imagine-other perspective. In other words, they were instructed to imagine these painful situations as happening to themselves or to someone else (i.e., self and other ToM). When individuals with high psychopathic scores were asked to adopt an imagine-self perspective, they exhibited typical enhanced activity in the anterior midcingulate cortex, anterior insula, and inferior frontal gyrus, somatosensory cortex, and right amygdala with the orbitofrontal cortex as compared to non-psychopathic individuals. In contrast, when asked to imagine another person in pain (i.e., imagine-other perspective), they exhibited reduced activity in similar brain areas: the anterior insula, ventromedial prefrontal cortex, the orbitofrontal cortex, and the amygdala as compared to the non-psychopathic participants. The low anterior insula activity for psychopathic individuals and the high activity for non-psychopathic individuals during the imagine-other perspective condition is an indicator of higher vicarious experience in non-psychopathic individuals.

Mier et al. (2014) investigated neural responses while male psychopathic and healthy control participants performed a social cognitive task for the assessment of general face

processing, emotion recognition and affective ToM. The behavioural findings revealed no deficits in affective ToM or emotion recognition abilities in psychopathic individuals. However, the neural findings demonstrated abnormal brain activity in the psychopathy group during task performance. Specifically, affective ToM was associated with reduced neural activity in key brain regions for embodied simulation as compared to non-psychopathic individuals, which included the amygdala, inferior prefrontal gyrus and the superior temporal sulcus. Moreover, reduced connectivity between the superior temporal sulcus and the amygdala was also observed.

In a later study by Decety et al., (2015), individuals who obtained higher psychopathy scores exhibited higher accuracy rate compared to low psychopaths in identifying emotional states of the protagonists from short clips. At the neural level, however, there were important neural differences between the groups. Individuals with high psychopathy compared to low psychopaths exhibited increased neural response in brain regions implicated in intention and emotion understanding and included the right posterior superior temporal sulcus/temporo-parietal junction (pSTS/TPJ), amygdala, insula, ventral striatum, and the prefrontal cortex (PFC).

Finally, Deming et al. (2020) investigated the neural activity of psychopathic individuals during affective perspective-taking, during which they were required to identify the emotional state of a target among two people interacting. Behavioural findings revealed that psychopathy was negatively correlated with accuracy in identifying three emotions: fear, sadness and happiness. In addition, neural findings indicated hypoactivity in brain regions implicated in key empathy regions, however, only during fear trials. The brain regions included the left anterior insula (aINS) and left posterior orbitofrontal cortex (pOFC).

Taken together, the finding from the neuroimaging studies show that while for the most part people with psychopathy exhibit intact ToM abilities, their neural responses suggest that

psychopaths represent mental states using different mechanism than people without psychopathy. However, this is some indication for a breakdown in ToM abilities under specific conditions such as when explicitly required to process another's fear (Deming et al., 2020).

5.2.3 Controlled versus automatic ToM processing

Only two of the included studies investigated controlled versus automatic ToM processing in psychopathic individuals (Nentjes et al., 2015a; Drayton et al., 2018). Controlled ToM processes refer to an intentional attempt to consider the perspective of another individual, whereas automatic ToM processes refer to the automatic and unintentional representation of another person's perspective (Drayton et al., 2018). Both studies utilized the PCL-R (Hare, 2003) for the assessment of psychopathy, and included all male incarcerated offenders.

Nentjes et al., (2015a) examined the impact of controlled (slow) versus automatic (fast) processing in psychopathic individuals' ToM performance, using the RMET task. For the study's purposes, the RMET task was divided into the two following conditions: RMET short (i.e., time restriction) and RMET long (i.e., no time restriction), which respectively designed to tap automatic and controlled ToM processes. Furthermore, the authors investigated the influence of hostility biases in ToM performance by the inclusion of more hostile answering options in the RMET task. The results indicated that psychopathic individuals ToM abilities were intact during both RMET conditions as well as in presence of hostile alternatives. Thus, RMET performance was not undermined by the quick presentation of stimuli nor by the presence of hostile response alternatives.

In Drayton et al., (2018), controlled versus automatic ToM abilities was examined in a sample of psychopathic and non-psychopathic male offenders. The authors employed a computer-based perspective-taking task developed by Samson et al. (2010). Participants were shown pictures of a human avatar in a room with dotted walls and asked to identify the

number of dots either from the avatar's perspective (controlled) or from their own perspective (automatic). The results of this study revealed that individuals with higher psychopathy scores were less likely to be affected by the perspective of the avatar when identifying the number of dots from their own perspective than individuals with lower psychopathy scores. These results suggest that psychopathic individuals may lack the ability to automatically take the perspective of others but can deliberately take another's perspective.

Taken together, there is limited literature on the association of psychopathy with processing type required to perform ToM abilities. The findings from the two studies provide conflicting results with Nentjes et al. (2015a) showing no ToM deficits in psychopathic individuals during controlled or automatic ToM processing, while Drayton et al. (2018) report a diminished propensity of psychopathic individuals to automatically take the perspective of others. These differences may be due to employing different tasks to measure controlled vs automatic ToM processing. Nentjes et al. (2015a) utilized an affective perspective-taking task (RMET; Baron-Cohen et al., 2001), whereas Drayton et al. (2018) used a cognitive perspective-taking measure (Samson et al., 2010). Examining differences in controlled vs automatic processing is potentially a useful framework for future research to more fully characterize the association of psychopathy with ToM abilities.

5.3 Findings from non-clinical populations

Summary of the findings from the studies in non-clinical populations are presented under two headings. The first (section 5.3.1) summarizes the behavioural findings, and the second (section 5.3.2) summarizes findings from studies examining sex differences.

5.3.1 Behavioural findings

There were only six studies using healthy samples (Mullins-Nelson et al., 2006; Ali & Chamorro-Premuzic, 2010; Lockwood et al., 2013; Jameel et al., 2019; Lanciano & Curci, 2019). The findings from four studies report ToM impairments in people with psychopathy (Mullins-Nelson et al., 2006; Ali & Chamorro-Premuzic, 2010; Gillespie et al., 2017; Lanciano & Curci, 2019). The two remaining studies demonstrated intact ToM abilities in psychopathic individuals (Jameel et al. 2019; Lockwood et al. 2013).

Mullins-Nelson et al. (2006) found ToM impairments (IRI; Davis, 1983) among higher scorers of psychopathy (PPI-SF; Lilienfeld, 2004) when considering the construct of psychopathy as multidimensional, but intact ToM capacities when considering psychopathy as a unitary construct. The PPI-SF is based on two factors, the first one (PPI-SF-I) corresponds to personality whereas the second one (PPI-SF-II) to behavioural features (Mullins-Nelson et al. 2006). Overall, the findings of this study revealed intact perspective-taking abilities and affective empathy for high scorers on the personality dimension (PPI-SF-I), but impaired perspective-taking and affective empathy for high scorers on behavioural dimension of the (PPI-SF-II). Therefore, only high scorers on the behavioural dimension of the PPI-SF exhibited deficits in perspective-taking abilities and affective empathy.

Ali and Chamorro-Premuzic (2010) utilized the LSRP (Levenson et al., 1995) for the measurement of primary and secondary psychopathy. The authors found that both psychopathy sub-types were associated with general ToM deficits as well as ToM deficits to specific emotions (Faces Test, Baron-Cohen et al., 1997; RMET, Baron-Cohen et al., 2001;

RMVT, Golan et al., 2006). However, primary psychopathy was characterized with more ToM impairments than secondary psychopathy. Individuals with primary psychopathy traits had more difficulties identifying neutral emotional states whereas individuals with secondary psychopathy traits experienced more difficulties with positive emotional states.

Using a sample of healthy students, Gillespie et al. (2017) employed a naturalistic ToM task referred to as the Movie for the Assessment of Social Cognition (MASC; Dziobek et al., 2006). The purpose of the study was to examine whether psychopathic tendencies (LSRP;) coupled with either autistic traits or positive psychotic experiences affect differently ToM abilities. The results revealed that psychopathic tendencies with higher autistic traits were associated with a declining *cognitive ToM* performance. In contrast, psychopathic tendencies coupled with higher positive psychotic experience were linked to improved *cognitive ToM* performance. The authors also found that poorer *affective ToM* abilities were uniquely linked to higher psychopathic tendencies.

Lanciano and Curci (2019) used the Italian version of PPI-R (La Marca et al. 2008; Lilienfeld and Widows, 2005) to measure different factors of psychopathy, namely Self-Centered Impulsivity, Fearless Dominance, and Cold-heartedness. The results showed that perspective-taking abilities (IRI; Davis, 1980, 1983) varied as a result of different psychopathic traits. More precisely, the self-centred impulsivity and cold-heartedness factors of the PPI-R were associated with low perspective-taking abilities, whereas the fearless dominance factor of the PPI-R was associated with high perspective-taking abilities. These results indicate that psychopathic individuals may differ in their perspective-taking abilities when considering different psychopathic trait dimensions

In contrast, the two remaining studies report no ToM impairments in healthy participants with higher scores of psychopathy (Jameel et al. 2019; Lockwood et al. 2013). Jameel et al. (2019) found intact perspective-taking abilities, but impaired emotional empathy (IRI; Davis,

1980) in individuals with higher psychopathy scores (PPI-SF; Lilienfeld & Hess, 2001). The results of Lockwood et al. (2013), who employed a purely cognitive ToM task (Theory of mind animations task; Abell et al., 2000), showed a positive relationship between high psychopathic traits (SRP-4-SF; Paulhus et al., in press) and reduced affective resonance but not for cognitive perspective-taking. Thus, perspective-taking abilities appear to be intact when employing a purely cognitive task.

5.3.2 Sex differences in ToM abilities

Out of the 24 studies included for qualitative analysis, only 8 of them had both female and male participants (Mullins-Nelson et al., 2006; Ali & Chamorro-Premuzic, 2010; Lockwood et al., 2013; Bo et al., 2014; Abu-Akel et al. 2015; Jameel et al., 2019; Lanciano & Curci, 2019), of which only two were in clinical populations (Bo et al., 2014; Abu-Akel et al. 2015). If we take into consideration the outcome of the eight studies using both male and female participants, the studies conducted by Mullins-Nelson et al. (2006), Ali and Chamorro-Premuzic (2010), Gillespie et al. (2017) and Lanciano and Curci (2019) demonstrated impaired ToM/perspective-taking abilities. In contrast, the remaining studies demonstrated intact ToM/perspective-taking abilities (Lockwood et al., 2013; Bo et al., 2014; Abu-Akel et al. 2015; Jameel et al., 2019).

However, only one of the seven studies explored sex differences (Mullins-Nelson et al., 2006). When considering psychopathy as a unitary construct, the authors found no significant difference between psychopathic and non-psychopathic individuals in perspective-taking abilities (IRI; Davis, 1980), but deficiencies in affective empathy. From a dimensional perspective, in the entire sample, the authors found a positive association for perspective-taking with the personality dimension of the PPI-SF (Factor 1), but a negative association with behavioural dimension of the PPI-SF (Factor 2). When splitting the sample by sex, the study did not reveal significant differences between males and females. The data

demonstrated a *significant* positive association between males with higher Factor 1 scores and perspective-taking abilities. With respect to females with higher Factor 1 scores, the results demonstrated a *non-significant* negative association with perspective-taking abilities. Regarding higher scorers of factor 2 PPI-SF, both female and male participants exhibited a *significant* negative correlation with perspective-taking abilities. Therefore, perspective-taking deficits were found only in higher scorers of Factor 2 traits. However, these results should be interpreted with caution as the sample size of males (n = 44) was considerably smaller than the sample size of females (n = 140). It is evident that sex differences regarding the relationship between psychopathy and ToM/perspective taking is massively an underrepresented area of research.

6. Discussion

The association between psychopathy and ToM abilities is an area of considerable interest and debate, particularly in light of the paradoxical nature in which psychopathic individuals behave. Indeed, psychopathic individuals have been characterized with an enhanced ability to charm and manipulate other people for personal gains which might be indicative of operative ToM abilities, and yet they also present traits deemed detrimental for the well-being of other individuals and society, which might be indicative of dysfunctional ToM abilities. In an effort to reconcile this paradox, numerous studies have been carried out, yielding mixed results. In order to improve our current understanding of the association between ToM abilities and psychopathy, I conducted a systematic review of all empirical peer-reviewed studies published in English that investigated the relationship between ToM abilities and psychopathy in adults over last two decades.

In the following pages, I thematically summarize the main empirical findings of the systematic review and their implications to understanding the relationship between ToM abilities and psychopathy in clinical and non-clinical populations. I then discuss the

limitations of the reviewed research and make recommendations for future directions. Finally, I conclude with comments highlighting the implication of these findings to interventions that target ToM abilities in people with psychopathy.

6.1 The association between psychopathy and ToM abilities in clinical populations

Broadly speaking, the results from the studies using clinical populations revealed an overall tendency for intact affective and cognitive ToM abilities in psychopathic individuals. Indeed, most of the studies showed no ToM impairments at both the diagnostic (Richell et al., 2003; Dolan & Fullam, 2004; Sommer et al., 2010; Seidel et al., 2013; Mier et al., 2014; Decety et al., 2015; Nentjes et al., 2015a, Nentjes et al., 2015b) and trait levels (Domes et al., 2013; Möller et al., 2014). Notably, in one study that examined this relationship in schizophrenic patients (Abu-Akel et al., 2015) revealed similar results, such that patients with higher psychopathic traits presented better ToM abilities.

However, in a set of other studies that examined both the cognitive and affective components of ToM revealed a more nuanced picture. Shamay-Tsoory et al. (2010) found intact cognitive ToM but deficient affective ToM in individuals with psychopathy. Similarly, albeit in patients with schizophrenia, Bo et al. (2014) found that with co-occurring psychopathy are characterized by intact cognitive but deficient emotional mentalizing abilities. The findings from Deming et al. (2020) further suggest that impaired affective perspective-taking in psychopathy might be specific to representing other's fear, happiness and sadness. Given these results, it is tempting to speculate for intact cognitive ToM but impaired affective ToM abilities in psychopathy. However, the majority of the studies using clinical populations employed affective perspective-taking tasks and yet revealed no affective ToM deficits in individuals with psychopathy (Richell et al., 2003; Sommer et al., 2010; Seidel et al., 2013; Mier et al., 2014; Decety et al., 2015; Nentjes et al., 2015a).

Taken together, the overwhelming evidence indicates that cognitive ToM abilities are intact. With respect to affective ToM abilities, the findings are mixed and thus firm conclusions cannot be drawn. The findings regarding intact or even superior ToM abilities aligns with the notion of enhanced capacity of some individuals with psychopathy to manipulate and deceive other people for their own personal gains. Indeed, intact ToM is necessary for successfully deceiving other people (Baron-Cohen, 2000; Talwar et al., 2007 as cited in Nentjes et al., 2015b). For Smith (1978), psychopathic individuals are masters when it comes to finding out what other people want or desire. This competence aids them in getting close to others as well as make others feel as if they are getting what they want, while the truth is that the psychopath may be exploiting them (Smith, 1978 as cited in Mullins-Nelson et al., 2006).

However, it is important to consider the inconsistent findings regarding affective ToM abilities in psychopathy, particularly in light of claims suggesting that poorer ToM abilities in people with elevated antisocial lifestyle traits may account for the aggressive and antisocial conducts (e.g., Taubner et al., 2013). If so, one reason for these mixed results might be due the lack of studies considering the different facets of psychopathy, and the reliance instead on a categorical approach of psychopathy (psychopaths vs non-psychopaths), which might obscure the heterogeneity in people with psychopathy. Consistent with this, Blackburn (2007) suggests that it is inaccurate to sharply categorize individuals as psychopathic and non-psychopathic, but rather as individuals possessing varying levels of psychopathic traits. The results of four studies (Dolan & Fullam, 2004, Seidel et al., 2013, Brook & Kosson, 2013; Sandvik et al., 2014) converge to suggest that reduced accuracy in affective ToM/perspective-taking abilities are found in people with psychopathy who express higher antisocial lifestyle traits (Factor 2). However, given the limited number of studies that examined ToM abilities

with the different facets of psychopathy, future research is needed to inform the nature of this relationship.

Another source that might have contributed to these divergent results is the choice of instrument for measuring psychopathy (see Sandvik et al., 2014). Hence, it is important to consider the type of instrument used for psychopathy measurement when comparing findings from distinct studies. Researchers should be particularly careful with self-report instruments as they may yield erroneous findings, particularly regarding the interpersonal/affective traits of psychopathy (Sandvik et al., 2014). Indeed, in the study by Sandvik et al. (2014), the self-report (SRP-III; Hare, 1985) results obtained for the interpersonal/affective traits of psychopathy did not match those obtained from the clinical psychopathy measure (PCL-R; Hare, 2003).

Examining ToM abilities using paradigms that distinguish between controlled versus automatic ToM processing can considerably improve our understanding of how people with psychopathy process mental states. However, the systematic review detected only two studies (Drayton et al., 2018), Nentjes et al., 2015a), reporting conflicting results, possibly due to the use of different tasks. Nentjes et al. (2015a) used an affective ToM task whereas Drayton et al. (2018) a cognitive ToM task. Using a cognitive ToM task, Drayton et al. (2018) pointed out to a failure of psychopathic individuals to automatically take the perspectives of others but intact deliberate perspective-taking. The authors postulate that psychopathic people experience difficulties in automatically taking the perspectives of others due to attention abnormalities. In other words, they select only goal-relevant information and limit the influence of goal-irrelevant cues. This attention dysfunction may be at the heart of the behavioural and affective abnormalities with which present psychopathic individuals. In support of this assumption are the results from the study of Dadds et al. (2006). Using a sample of children with psychopathic traits, the authors demonstrated that the initial fear

recognition deficits disappeared when asking the participants directly to pay attention to the eye region of people's faces. In line with this, Richell et al. (2003) employed the RMET task, and found an intact capacity in a sample of adults with psychopathy to identify emotions from the eye region only.

Nentjes et al. (2015a) did not find any ToM deficits in psychopathic individuals during controlled or automatic ToM processing. The authors find questionable the tendency of psychopathic individuals to perform well in the RMET task, considering the importance of amygdala activation for better RMET performance (Baron-Cohen et al., 1999 as cited in Nentjes et al., 2015a). More precisely, psychopathy is expected to impact negatively RMET performance as it is a condition characterized by amygdala abnormalities. It has been thus suggested that individuals with psychopathy compensate for their amygdala dysfunctions by relying on other cortical neural circuits (Richell et al., 2003). Individuals with psychopathy may achieve the same results as other healthy individuals through different means. To prevent psychopathic individuals from using such compensatory strategies, Nentjes et al. (2015a) modified the RMET task. However, this experimental task may need further modification as it may have been not sufficiently suitable for tapping into early stages of information processing (Nentjes et al., 2015a). The authors suggest that future studies may need to clarify the role of cortical brain regions by employing tasks with shorter stimulus duration in combination with neuroimaging techniques (Nentjes et al., 2015a).

The present systematic review included five studies utilizing a combination of behavioural and neuroimaging tasks (Sommer et al., 2010; Decety et al., 2013; Mier et al., 2014; Decety et al., 2015; Deming et al., 2020), which might help explain potential compensatory mechanisms in people with psychopathy. All five studies demonstrated neural activity abnormalities during affective perspective-taking tasks among individuals with psychopathy. Importantly, are the findings of the four studies reporting behavioural findings, which demonstrated neural

abnormalities in the presence of typical ToM performance (Sommer et al., 2010; Mier et al., 2014; Decety et al., 2015). These findings resonate with Richell and colleagues (2003) who stated that: "Theory of Mind is intact in psychopathic individuals even if the neural architecture mediating Theory of Mind is not equivalent to that used by healthy developing populations (p.525). Collectively, these findings suggest that people with psychopathy rely on a different neural mechanism than non-psychopaths. In this regard, Sommer et al. (2010) suggested that while ToM in individuals with psychopathy is mediated by the activation of brain regions implicated in rational processes, it is mediated by the mirror neuron system, which is implicated in the simulation of other people's mental states. These results support Mealey and Kinner's (2003) proposition with respect to psychopath's emotional functioning. Simply put, individuals with psychopathy do not lack the capacity for inferring other people's emotional and mental states but rather the ability to simulate the emotional states of other people (Mealey and Kinner, 2003, as cited in Sommer et al., 2010). The seeming inability of people with psychopathy to take the perspective of others through simulation processes might help explain their callousness and lack of empathy towards others.

Finally, the studies utilizing clinical populations used mainly all male samples which makes generalizability to clinical psychopathic females difficult and thus prevent us from gaining a better understanding of the potential role of sex in the relationship between psychopathy and ToM abilities. Indeed, the assumption that results achieved from men are generalizable to women may lead to substantial scientific errors (Wynn et al., 2012). Additionally, Russell et al. (2007) proposed that females may have an advantage in affective ToM whereas males in cognitive ToM. It is thus possible that female psychopaths are better at affective ToM tasks than cognitive ToM tasks whereas psychopathic males better at cognitive ToM tasks than affective ones. Therefore, potential differences between the sexes must receive greater attention in future studies.

6.2 The association between psychopathy and ToM abilities in nonclinical populations

There were only six studies examining the relationship between psychopathy and ToM abilities in nonclinical populations. Three of studies revealed a positive association between psychopathy, conceptualized as a unitary construct, and cognitive ToM abilities in healthy individuals (Mullins-Nelson et al., 2006; Lockwood et al., 2013; Jameel et al., 2019). However, none of these studies assessed affective ToM, although they all report negative association between psychopathy and affective empathy, which measures the ability to respond to the distress of others with an appropriate emotion.

The affective component of ToM in relation to psychopathy was assessed by Ali and Chamorro-Premuzic (2010) and Gillespie et al. (2017), and interesting findings emerged since both studies revealed affective ToM impairments. In addition, the relationship between ToM and psychopathy in this population turns out to be more nuanced when considering the different dimensions/facets of psychopathy. In this respect, Mullins-Nelson et al. (2006) found that while people with higher scores on the interpersonal/affective facet of psychopathy presented better perspective-taking abilities, people with higher scores on the antisocial facet of psychopathy showed worse perspective-taking abilities. Ali and Chamorro-Premuzic (2010) showed that ToM impairments were in addition linked to the antisocial facet of psychopathy. However, they also report that the magnitude of the impairment was more pronounced in relation to the interpersonal/affective facet compared to the antisocial facet of psychopathy. Finally, Lanciano and Curci (2019) also report that ToM abilities in psychopathic individuals varied as a result of different psychopathic traits. In this study, higher perspective-taking abilities were associated higher scores in fearless dominance traits (characterized by diminished fear responses and superior manipulative abilities), whereas lower perspective-taking abilities were associated with both higher scores in self-centred impulsivity (characterized by the tendency to engage in reckless and unplanned behaviours

with no consideration of social norms) as well as cold-heartedness (characterized by callousness, guiltlessness as well as lower empathy).

The above findings clearly demonstrate the way results may differ because of the approach taken when conceptualizing psychopathy. Perspective-taking/ToM impairments were brought to light only by studies that took a multidimensional approach to psychopathy. In addition, all the included studies utilized samples with both male and female participants. It should be noted that Mullins-Nelson et al. (2006), Ali & Chamorro-Premuzic (2010), Lanciano & Curci (2019) utilized female dominated samples which may impact the generalizability of findings. Finally, all six studies relied on self-report measures of psychopathy, thus vigilance needs to be taken when trying to generalize the above-mentioned findings to clinical populations.

Self-report measures can be used for gaining better understanding of a disorder (Lilienfeld & Fowler, 2006, as cited in Mullins-Nelson et al., 2006), but it has been advised to use them the least possible and in combination with other types of testing (Mullins-Nelson et al., 2006). Indeed, self-report measures present a great risk for dishonest answering (Allport, 1961; Lilienfeld & Fowler 2006, as cited in Sandvik, 2014), especially when measuring psychopathy as it is a condition composed of diverse socially unacceptable traits (Sandvik, 2014). In addition, it remains unclear whether self-report measures of cognitive empathy are reliable capturers of empathic abilities in psychopathic individuals (i.e., three out of the six studies used self-reported measures of empathy) (Mullins-Nelson et al., 2006). It is uncertain whether individuals are able provide reliable information about such subtle and complex processes (Nisbett & Wilson, 1977 as cited in Brook & Kosson, 2013). Brook & Kosson (2013) maintain that most existent instruments fail to accurately measure both cognitive and emotional aspects of empathy. They do not consider the complexity of empathy in human interactions; That is, human to human contact involves a variety of verbal and non-verbal cues which need to be interpreted for successful social interaction (Brook & Kosson, 2013).

6.3 Limitations of the reviewed research and recommendations for future direction

The majority of the clinical studies were composed of generally small, all-male participants. The inclusion of all-male inmates limits the generalizability of the findings to the psychopathic population at large. For instance, the participant's sex may be a protective factor with respect to general or specific components of Theory of Mind (Russell et al. 2007; Ibanez. Et al., 2012). As pointed above (Section 2.4), there appear indeed differences between males and females in ToM performance (Russell et al. 2007; Ibanez. Et al., 2012), but these differences may depend on the component of ToM (Russell et al., 2007). The potential impact of sex differences is further supported by findings, which suggest varying behavioural manifestations of psychopathy as a function of the participant's sex (Forouzan and Cooke, 2005). It is, therefore, particularly relevant for future clinical studies to include female participants for comparison. With respect to the non-clinical studies, the studies were composed of female-dominated, convenience samples and, therefore, the results may not generalize fully to male participants or to the general population. Moreover, the studies findings were predominantly based on samples from, educated, industrialized, rich and democratic (WEIRD) societies, which may not be applicable to other cultural contexts.

Another important limitation is that most of the studies—with the exception of the study by Lanciano and Curci (2019) (N= 736)—were conducted in relatively small sample sizes. This suggests that some of the studies are underpowered and, therefore, potentially unable to detect group differences, or to uncover potentially meaningful significant correlations between psychopathy and ToM.

The use of different assessment methodology for the measure of psychopathy is another concern that may limit the interpretation of reported findings. As demonstrated by Sandvik et al. (2014), the use of self-report versus clinical measures of psychopathy may yield quite different findings. Self-report measures may omit important aspects of the

interpersonal/affective traits of psychopathy (Sandvik et al., 2014), and present a risk for dishonest answering (Allport, 1961; Lilienfeld & Fowler 2006, as cited in Sandvik, 2014), especially when measuring a condition like psychopathy, which is comprised of various socially unacceptable traits (Sandvik, 2014). In addition, the multifaceted nature of psychopathy appears also majorly neglected by existent research. A further limitation is that there were studies that did not control for confounding variables such as including participants with comorbid personality disorders (e.g., Nentjes et al., 2015a,b), as well as inmates with a variety of offences (e.g., sexual offenses, drug-related sentences, violent-related sentences), and times spent imprisoned (e.g., Sommer et al., 2010).

Likewise, the use of different assessment measures of ToM across the study raises several methodological limitations. First, several studies have relied on self-report than performance-based measures of ToM. Self-report measures of ToM can be problematic, because they may not adequately capture ToM abilities (Nisbett & Wilson, 1977 as cited in Brook & Kosson, 2013) as they are susceptible to social-desirability responding and often are found to only be weakly related to performance-based tasks (e.g., Dodell-Feder et al., 2013; Giordano et al., 2019). Second, another limitation is the use of age-inappropriate ToM measures for adults (e.g., Dolan & Fullam, 2004), which may induce a ceiling effect in task performance. Third, most included studies failed to capture both cognitive and affective components of ToM. Indeed, the vast majority of ToM tasks tapped only on one of the components of ToM, which makes it then difficult to identify whether individuals with psychopathy suffer from cognitive or affective ToM deficiencies. Relatedly, ToM's definition varied from study to study, with some distinguishing the affective from cognitive aspects of ToM, while others were negligent of the multifaceted nature of ToM. Finally, there is even a debate regarding whether some of the performance-based measures of ToM, such as the RMET (Baron-Cohen et al., 2001), are indeed adequate measures of the ToM construct, which might be only tapping lower-level

processes such as emotion recognition, rather than the ability to represent mental states per se (see Quesque and Rossetti (2020) for a useful discussion of selection criteria of appropriate tasks). Therefore, there is an urgent need for future research to have a careful consideration of the operationalization and selection of performance-based ToM measures, and for greater consistency in the criteria used to select ToM tasks in clinical and forensic psychological science. This is necessary to improve our understanding of the relationship between psychopathy and ToM, and more generally the effects of psychopathology on social cognitive abilities.

6.4 Limitations of the systematic review

The present systematic review aimed to provide a sharp focus of the state-of-the-art knowledge about the nature of ToM abilities in psychopathy. However, the current review may have not been exhaustive, and it is, therefore, possible that some of the papers were not captured. For instance, the keywords Cognitive Empathy and Mindreading were not employed during the search, and this possibly omitted some potential studies for inclusion. Moreover, a meta-analytic evaluation of the available studies would be an important complement to the current systematic review, as it can provide a statistical estimate of the degree (small/medium/large) of the association between ToM abilities and psychopathy. However, robust meta-analysis requires similar measurements across studies, which would have limited the number of studies eligible for the present systematic review.

7. Conclusion and implications

As revealed by most results examined in the present systematic review, both the cognitive and affective components of ToM appear to be intact in people with psychopathy. This may account for their ability to successfully manipulate and deceive other people. However, a multidimensional view of psychopathy revealed a more nuanced picture, where antisocial/lifestyle traits or 'secondary psychopathy' appear to be related to poorer ToM abilities, and interpersonal/affective traits or 'primary psychopathy' appear to be associated with no ToM deficits and even superior ToM capabilities. It seems thus important for future research to avoid conceptualizing psychopathy as unitary construct, an approach that might obscure specific associations between different facets of psychopathy and ToM.

Furthermore, neuroimaging showed that while psychopathic individuals perform typically on ToM tasks, they present atypical brain activities. This suggests that psychopathic individuals may be using different neural means to navigate the social world. Moreover, the present systematic review indicated that the majority of available evidence might be challenged by a number of methodological limitations, which considerably limit the generalizability of the findings. Taken together, the exact nature of ToM abilities in psychopathy remains controversial, and it is a necessity for future studies to take into consideration the multidimensional nature of both psychopathy and ToM, task criteria, potential neural abnormalities, as well as the potential influence of the participant's sex on the results.

Notwithstanding these limitations, the findings of the current systematic review may have important implications for intervention programs targeting ToM abilities in psychopathy (e.g., Bateman et al., 2016). It is, therefore, important to be aware of the skills psychopathic individuals may acquire from existent treatment programs. In other words, therapists should avoid implementing treatments capable of further enhancing the ToM abilities of

psychopathic individuals, and particularly in those with higher interpersonal/affective traits. Multiple reports suggest that individuals with psychopathy may later on employ some of the learned treatment tactics for the manipulation of future victims (see Viding, 2019 for a review). A future direction for treatment programs would be to target the affective component of empathy, which has been shown to malfunction in psychopathic individuals (Dolan & Fullam, 2004; Mullins-Nelson et al., 2006; Seidel et al., 2013; Seara-Cardoso et al., 2016; Jameel et al., 2019). Although overly optimistic, the final goal would be to enhance spontaneous empathizing (i.e., affective resonance) in individuals with primary psychopathy (Viding, 2019). In contrast, psychopathic individuals expressing higher antisocial lifestyle traits may benefit from treatment programs designed at enhancing their deficient affective ToM abilities. Enhancing affective ToM abilities in psychopaths with elevated antisocial lifestyle traits may inhibit violent behaviours, which characterize this particular subgroup of psychopaths. The potential of these ToM-related interventions in reducing antisocial behaviours in people with psychopathy should be evaluated in future controlled clinical trials.

Bibliography

Allport, G. W. (1961). Pattern and growth in personality. New York, NY: Holt, Rinehart and Winston.

Abu-Akel, A., & Abushua'leh, K. (2004). 'Theory of mind'in violent and nonviolent patients with paranoid schizophrenia. *Schizophrenia research*, 69(1), 45-53.

Abu-Akel, A., & Shamay-Tsoory, S. (2011). Neuroanatomical and neurochemical bases of theory of mind. *Neuropsychologia*, *49*(11), 2971-2984.

Abu-Akel, A., Heinke, D., Gillespie, S. M., Mitchell, I. J., & Bo, S. (2015). Metacognitive impairments in schizophrenia are arrested at extreme levels of psychopathy: The cut-off effect. *Journal of Abnormal Psychology*, *124*(4), 1102.

Abell, F., Happe, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development*, *15*(1), 1-16.

Ali, F., & Chamorro-Premuzic, T. (2010). Investigating theory of mind deficits in nonclinical psychopathy and Machiavellianism. *Personality and Individual Differences*, *49*(3), 169-174.

Anderson, N. E., & Kiehl, K. A. (2012). The psychopath magnetized: insights from brain imaging. *Trends in cognitive sciences*, *16*(1), 52-60.

American Psychiatric Association. (1980). Diagnostic and statistical manual of mental disorders (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev.). Washington, DC: Author.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5 ed.). Washington, DC: American Psychiatric Publishing.

Apperly IA, Butterfill SA (2009): Do humans have two systems to track beliefs and belief-like states? *Psychol Rev* 116:953–970

Apperly, I. A. (2010). Mindreaders: The cognitive basis of "theory of mind". Hove, UK: *Psychology Press*.

Apperly, I. A., Warren, F., Andrews, B. J., Grant, J., & Todd, S. (2011). Developmental continuity in theory of mind: Speed and accuracy of belief–desire reasoning in children and adults. *Child development*, *82*(5), 1691-1703.

Apperly, I. A. (2012). What is "theory of mind"? Concepts, cognitive processes and individual differences. *Quarterly Journal of Experimental Psychology*, *65*(5), 825-839.

Arrigo, B. A., & Shipley, S. (2001). The confusion over psychopathy (I): Historical considerations. *International Journal of Offender Therapy and Comparative Criminology*, 45(3), 325-344.

Araujo Gonçalves, S. (2018). *La psychopathie et ses différents enjeux: terminologie, sémiologie, facteurs étiologiques et traitements. De la clinique à la fiction* (Doctoral dissertation, Université de Lausanne, Faculté des sciences sociales et politiques).

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: Bradford Books, MIT Press.

Baron-Cohen, S., Wheelwright, S., & Jolliffe, A. T. (1997). Is there a" language of the eyes"? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual cognition*, *4*(3), 311-331.

Baron-Cohen, S., Ring, H. A., Wheelwright, S., Bullmore, E. T., Brammer, M. J., Simmons, A., & Williams, S. C. (1999). Social intelligence in the normal and autistic brain: an fMRI study. *European journal of neuroscience*, *11*(6), 1891-1898.

Baron-Cohen, S. (2000). Theory of mind and autism: A fifteen year review. *Understanding other minds*: *Perspectives from developmental cognitive neuroscience*, *2*(3-20), 102.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 241-251.

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.

Baron-Cohen, S. (2009). Autism: The empathizing–systemizing (E–S) theory. *Annals of the New York Academy of Sciences*, *1156*(1), 68–80.

Blair, J., Sellars, C., Strickland, I., Clark, F., Williams, A., Smith, M., & Jones, L. (1996). Theory of mind in the psychopath. Journal of Forensic Psychiatry, 7(1), 15-25.

Blair, R. J. R. (2003). Neurological basis of psychopathy. *The British Journal of Psychiatry, 182*, 5–7. doi:10.1192/bjp.182.1.5

Blackburn, R. (2005). Psychopathy as a personality construct. *Handbook of personology and psychopathology*, 271-291.

Blackburn, R. (2007). Personality disorder and psychopathy: Conceptual and empirical integration. *Psychology, Crime & Law*, *13*(1), 7-18.

Book, A. S., Quinsey, V. L. & Langford, D. (2007). Psychopathy and the perception of affect and vulnerability. *Criminal Justice and Behavior*, *34*, 531–542.

Butterfill S.A., Apperly I.A. (2013): How to construct a minimal theory of mind. *Mind Lang* 28:606–637.

Brook, M., & Kosson, D. S. (2013). Impaired cognitive empathy in criminal psychopathy: evidence from a laboratory measure of empathic accuracy. *Journal of abnormal psychology*, *122*(1), 156.

Bateman, A., O'Connell, J., Lorenzini, N., Gardner, T., & Fonagy, P. (2016). A randomised controlled trial of mentalization-based treatment versus structured clinical management for patients with comorbid borderline personality disorder and antisocial personality disorder. *BMC psychiatry*, *16*(1), 304. https://doi.org/10.1186/s12888-016-1000-9

Bueno, R. M. (2018). *Wolves in Sheep's Clothing: An Integrative Literature Review on Psychopathy*. California Institute of Integral Studies.

Bo, S., Abu-Akel, A., Kongerslev, M., & Simonsen, E. (2020). Predictors of criminal offending in a clinical sample of patients diagnosed with schizophrenia: A 6-year follow-up study. *Personality Disorders: Theory, Research, and Treatment.* https://doi.org/10.1037/per0000401

Cleckley, H. M. (1941/1976). The mask of sanity (5th ed.). St. Louis, MO: Mosby.

Cleckley, H.M. (1982). The mask of sanity (6th ed.). St Louis, MO: Mosby.

Cloninger, C. G. (1978). The link between hysteria and sociopathy. In H. S. Akiskal & W. Webb (Eds.), *Psychiatric diagnosis*: *Explorations of biological predictors* (pp. 189–218). New York, NY: Spectrum.

Cunningham, M. D., & Reidy, T. J. (1998). Antisocial personality disorder and psychopathy: Diagnostic dilemmas in classifying patterns of antisocial behavior in sentencing evaluations. *Behavioral sciences & the law*, *16*(3), 333-351.

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment*, *13*, 171–188. doi:10.1037/1040-3590.13.2.171

Cooke, D. J., Michie, C., Hart, S. D., & Clark, D. A. (2004). Reconstructing psychopathy: Clarifying the significance of antisocial and socially deviant behavior in the diagnosis of psychopathic personality disorder. *Journal of personality disorders*, *18*(4), 337-357.

Cooke, D. J., Michie, C., Hart, S. D., & Patrick, C. (2006). Facets of clinical psychopathy. *Handbook of psychopathy*, 91-106.

Coid, J., Yang, M., Ullrich, S., Roberts, A., & Hare, R. D. (2009). Prevalence and correlates of psychopathic traits in the household population of Great Britain. *International journal of law and psychiatry*, *32*(2), 65-73. doi: 10.1016/j.ijlp.2009.01.002.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalogue of Selected Documents in Psychology 10*, MS 2124, 85.

Davies, W., & Feldman, P. (1981). The diagnosis of psychopathy by forensic specialists. *British Journal of Psychiatry*, *138*, 329-331.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of personality and social psychology*, *44*(1), 113.

Dinges, N. G., Atlis, M. M., & Vincent, G. M. (1998). Cross-cultural perspectives on antisocial behavior. In D. M. Stoff, J. Breiling, & J. D. Maser (Eds.), *Handbook of antisocial behavior* (pp. 463-473). New York: Wiley.

Dolan, M., & Fullam, R. (2004). Theory of mind and mentalizing ability in antisocial personality disorders with and without psychopathy. *Psychological medicine*, *34*(6), 1093-1102.

Dadds, M. R., Perry, Y., Hawes, D. J., Merz, S., Riddell, A. C., Haines, D. J., ... & Abeygunawardane, A. I. (2006). Attention to the eyes and fear-recognition deficits in child psychopathy. *The British Journal of Psychiatry*, *189*(3), 280-281.

Dziobek, I. et al. Introducing MASC: A movie for the assessment of social cognition. *Journal of Autism and Developmental Disorders 36*(5), 623–636 (2006).

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., & Convit, A. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger

syndrome using the Multifaceted Empathy Test (MET). *Journal of autism and developmental disorders*, *38*(3), 464-473.

Derntl, B., Seidel, E. M., Kryspin-Exner, I., Hasmann, A., & Dobmeier, M. (2009a). Facial emotion recognition in patients with bipolar I and bipolar II disorder. *British Journal of Clinical Psychology*, *48*(4), 363-375.

Derntl, B., Finkelmeyer, A., Toygar, T. K., Hülsmann, A., Schneider, F., Falkenberg, D. I., & Habel, U. (2009b). Generalized deficit in all core components of empathy in schizophrenia. *Schizophrenia research*, *108*(1-3), 197-206.

Dumontheil, I., Apperly, I. A., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental science*, *13*(2), 331-338.

Domes, G., Hollerbach, P., Vohs, K., Mokros, A., & Habermeyer, E. (2013). Emotional empathy and psychopathy in offenders: An experimental study. *Journal of personality disorders*, *27*(1), 67-84.

Dodell-Feder, D., Lincoln, S. H., Coulson, J. P., & Hooker, C. I. (2013). Using fiction to assess mental state understanding: A new task for assessing theory of mind in adults. PLoS One, 8(11), e81279.

Dvash, J., & Shamay-Tsoory, S. G. (2014). Theory of mind and empathy as multidimensional constructs: Neurological foundations. *Topics in Language Disorders*, *34*(4), 282-295.

Decety, J., Chen, C., Harenski, C., & Kiehl, K. A. (2013). An fMRI study of affective perspective taking in individuals with psychopathy: imagining another in pain does not evoke empathy. *Frontiers in human neuroscience*, *7*, 489.

Decety, J., Chen, C., Harenski, C. L., & Kiehl, K. A. (2015). Socioemotional processing of morally-laden behavior and their consequences on others in forensic psychopaths. *Human brain mapping*, *36*(6), 2015-2026.

Drayton, L. A., Santos, L. R., & Baskin-Sommers, A. (2018). Psychopaths fail to automatically take the perspective of others. *Proceedings of the National Academy of Sciences*, *115*(13), 3302-3307.

Deming, P., Dargis, M., Haas, B. W., Brook, M., Decety, J., Harenski, C., ... & Kosson, D. S. (2020). Psychopathy is associated with fear-specific reductions in neural activity during affective perspective-taking. *Neuroimage*, *223*, 117342.

Eisenbarth, H., & Alpers, G. W. (2007). Validierung der deutschen Übersetzung des Psychopathy Personality Inventory (PPI). *Zeitschrift für Klinische Psychologie und Psychotherapie*, *36*(3), 216-224.

Ermer, E., Cope, L. M., Nyalakanti, P. K., Calhoun, V. D., & Kiehl, K. A. (2012). Aberrant paralimbic gray matter in criminal psychopathy. *Journal of abnormal psychology*, *121*(3), 649.

Feshbach, N. D. (1989). Empathy training and prosocial behavior. *Aggression and war: Their biological and social bases,* 101-111. New York, NY: Cambridge University Press.

Fonagy, P., Target, M., Steele, H., & Steele, M. (1998). Reflective-functioning manual. Version 5: For application to Adult Attachment Interviews. London, UK: University College London.

Federal Bureau of Investigation. (2005). Uniform crime reports. Washington, DC: Author.

Forouzan, E., & Cooke, D. J. (2005). Figuring out la femme fatale: Conceptual and assessment issues concerning psychopathy in females. *Behavioral sciences & the law*, *23*(6), 765-778.

Fowles, D. C., & Dindo, L. (2006). A dual-deficit model of psychopathy. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 14–34). New York, NY: Guilford Press.

Fowles, D. C., & Dindo, L. (2009). Temperament and psychopathy: A dual-pathway model. *Current Directions in Psychological Science*, *18*, 179–183. doi:10.1111/j.1467-8721.2009.01632.x

George, C., Kaplan, N., & Main, M. (1985). The Adult Attachment Interview, unpublished protocol (2nd ed.). Berkeley, CA: Department of Psychology, University of California.

Gur, R. C., Sara, R., Hagendoorn, M., Marom, O., Hughett, P., Macy, L., ... & Gur, R. E. (2002). A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *Journal of neuroscience methods*, *115*(2), 137-143.

Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. D. (2007). The 'Reading the Mind in the Voice'test-revised: a study of complex emotion recognition in adults with and without autism spectrum conditions. *Journal of autism and developmental disorders*, *37*(6), 1096-1106.

Gillespie, S. M., Mitchell, I. J., & Abu-Akel, A. M. (2017). Autistic traits and positive psychotic experiences modulate the association of psychopathic tendencies with theory of mind in opposite directions. *Scientific reports*, *7*(1), 1-9.

Giordano, M., Licea-Haquet, G., Navarrete, E., Valles-Capetillo, E., Lizcano-Cortés, F., Carrillo-Peña, A. and Zamora-Ursulo, A., 2019. Comparison between the short story task and the reading the mind in the eyes test for evaluating theory of mind: a replication report. *Cogent Psychology*, *6*(1), p.1634326.

Hart, S. D., Cox, D. N. & Hare, R. D. (1995). *The Hare Psychopathy Checklist: Screening Version (PCL-SV)*. Multihealth Systems Inc.: Toronto.

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

Hare, R. D. (1985). Comparison of procedures for the assessment of psychopathy. *Journal of Consulting and Clinical Psychology*, *53*, 7–16.

Hare, R. D., Harpur, T. J., Hakstian, A. R., Forth, A. E., Hart, S. D., & Newman, J. P. (1990). The revised psychopathy checklist: reliability and factor structure. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *2*(3), 338.

Hare, R. D. (1991). Manual for the Psychopathy Checklist- Revised. Toronto, Ontario, Canada: Multi-Health Systems.

Hare, Robert D. Without Conscience: The Disturbing World of the Psychopaths Among Us. 23– 25 Guilford Publications, Inc:1999.1993; [hereinafter Hare, Without Conscience].

Hare, R. D. (2003). Manual for the revised psychopathy checklist (2nd ed). Toronto: Multi-Health Systems.

Hare, R. D. (2003). The Hare Psychopathy Checklist -Revised (PCL-R) (2 ed.). Toronto: Multi-Health Systems Inc.

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

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of autism and Developmental disorders*, *24*(2), 129-154.

Happé, F., Brownell, H., & Winner, E. (1999). Acquired 'theory of mind' impairments following stroke. *Cognition*, *70*(3), 211-240.

Happé, F. (2003). Theory of mind and the self. *Annals of the New York Academy of Sciences*, *1001*(1), 134-144.

Hervé, H. (2007). Psychopathy across the ages: A history of the Hare psychopath. *The psychopath: Theory, research, and practice*, p.34.

Haas, B. W., Brook, M., Remillard, L., Ishak, A., Anderson, I. W., & Filkowski, M. M. (2015). I know how you feel: The warm-altruistic personality profile and the empathic brain. *PloS one*, *10*(3), e0120639.

Ickes, W. J. (1997). *Empathic accuracy*. New York, NY: Guilford Publications.

Ibanez, A., Huepe, D., Gempp, R., Gutiérrez, V., Rivera-Rei, A., & Toledo, M. I. (2013). Empathy, sex and fluid intelligence as predictors of theory of mind. *Personality and individual differences*, *54*(5), 616-621.

Jackson, R. L., Rogers, R., Neumann, C. S., & Lambert, P. L. (2002). Psychopathy in female offenders: An investigation of its underlying dimensions. *Criminal justice and behavior*, *29*(6), 692-704.

Jones, D. N. & Paulhus, D. L. (2010). Different provocations trigger aggression in narcissists and psychopaths. *Social Psychological and Personality Science*, *1*, 12–18.

Jameel, L., Vyas, K., Bellesi, G., Crawford, S., & Channon, S. (2019). Thinking about other's mistakes: contrasting patterns of performance in groups high or low in autistic or psychopathic traits. *Cognitive neuropsychiatry*, *24*(5), 369-385.

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.

Kiehl, K. A., & Hoffman, M. B. (2011). The criminal psychopath: History, neuroscience, treatment, and economics. *Jurimetrics*, *51*, 355.

Lykken, D. T. (1978). The psychopath and the lie detector. *Psychophysiology*, *15*, 137–142. doi:10.1111/j.1469-8986.1978. tb01349.x

Lykken, D. T. (1995). The antisocial personalities. Hillsdale, NJ: Erl- baum.

Lykken, D. T. (1996). Psychopathy, Sociopathy, and Crime. *Society*, *34*(1), 29-38.

Lykken, D. T. (2006). Psychopathic personality: The scope of the problem. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 3–13). New York, NY: Guilford Press.

Lilienfeld, S. O. (1994). Conceptual problems in the assessment of psychopathy. *Clinical Psychology Review*, *14*(1), 17-38.

Lilienfeld, S. O., VanValkenberg, C., Larntz, K., & Akiskal, H. S. (1986). The relationship of histrionic personality disorder to anti- social personality disorder and somatization disorders. *American Journal of Psychiatry*, *143*, 718–722.

Lilienfeld, S. O., & Hess, T. (2001). Psychopathic personality traits and somatization: Sex differences and the mediating role of negative emotionality. *Journal of Psychopathology and Behavioral Assessment*, *23*, 11–24.

Lilienfeld, S.O. (2004, March). Psychometric Properties of Self- Report Psychopathy Measures. *Paper presented at the meeting of the American Psychology-Law Society, Scottsdale, AZ.*

Lilienfeld, S. O., & Widows, M. R. (2005). Psychopathic personality inventory-revised: Professional manual. Lutz, Florida: *Psychological Assessment Resources*, Inc.

Lilienfeld, S. O., & Fowler, K. A. (2006). The self-report assessment of psychopathy: Problems, pitfalls, and promises. In C. J. Patrick (Ed.), *Handbook of psychopathy*, (pp. 107-132). New York, NY: Guilford Press

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.

Lang, S., Af Klinteberg, B., & Alm, P. O. (2002). Adult psychopathy and violent behavior in males with early neglect and abuse. *Acta Psychiatrica Scandinavica*, *106*, 93-100.

Lysaker, P. H., Carcione, A., Dimaggio, G., Johannesen, J. K., Nicolò, G., Procacci, M., & Semerari, A. (2005). Metacognition amidst narratives of self and illness in schizophrenia:

associations with neurocognition, symptoms, insight and quality of life. *Acta psychiatrica scandinavica*, *112*(1), 64-71.

La Marca, S., Berto, D., & Rovetto, F. (2008). Traduzione ed adattamento italiano del PPI-R: Psychopathic Personality inventory revised-PPI- R. Firenze: OS.

Lockwood, P. L., Bird, G., Bridge, M., & Viding, E. (2013). Dissecting empathy: high levels of psychopathic and autistic traits are characterized by difficulties in different social information processing domains. *Frontiers in human neuroscience*, *7*, 760.

Lanciano, T., & Curci, A. (2019). Psychopathic traits and self-conscious emotions: What is the role of perspective taking ability?. *Current Psychology*, 1-9.

McCord, W., & McCord, J. (1964). The psychopath: An essay on the criminal mind. New Jersey, NJ: D. Van Nostrand Company, Inc.

Millon, T., Simonsen, E., & Birket-Smith, M. (1998). Historical conceptions of psychopathy in the United States and Europe. In T. Millon, E. Simonsen, M. Birket-Smith, & R. D. Davis (Eds.), *Psychopathy: Antisocial, criminal, and violent behavior* (pp. 3-31). New York: Guilford.

Mier, D., Lis, S., Neuthe, K., Sauer, C., Esslinger, C., Gallhofer, B., & Kirsch, P. (2010a). The involvement of emotion recognition in affective theory of mind. *Psychophysiology*, *47*(6), 1028-1039.

Mier, D., Sauer, C., Lis, S., Esslinger, C., Wilhelm, J., Gallhofer, B., & Kirsch, P. (2010b). Neuronal correlates of affective theory of mind in schizophrenia out-patients: evidence for a baseline deficit. *Psychological Medicine*, *40*(10), 1607-1617.

Mier, D., Haddad, L., Diers, K., Dressing, H., Meyer-Lindenberg, A., & Kirsch, P. (2014). Reduced embodied simulation in psychopathy. *The World Journal of Biological Psychiatry*, *15*(6), 479-487.

Mealey, L., & Kinner, S. (2003). Psychopathy, Machiavellianism and theory of mind. *The social brain*: *Evolution and pathology*, 355-372.

Marsh, A. A., & Blair, R. J. R. (2008). Deficits in facial affect recognition among antisocial populations: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, *32*(3), 454-465.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., ... & Tugwell, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement (Chinese edition). *Journal of Chinese Integrative Medicine*, *7*(9), 889-896.

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.

Majois, V., Saloppé, X., Ducro, C., & Pham, T. H. (2011). Psychopathie et son évaluation. *Encyclopédie Médico-Chirurgicale-Psychiatrie*, 37-320.

Möller, C., Falkenström, F., Holmqvist Larsson, M., & Holmqvist, R. (2014). Mentalizing in young offenders. *Psychoanalytic psychology*, *31*(1), 84.

Nisbett, R. E. & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35*, 250–256. doi:10.1037/0022-3514.35.4.250

Neumann, Craig S.; Hare, Robert D. (2008). "Psychopathic traits in a large community sample: Links to violence, alcohol use, and intelligence"(PDF). *Journal of Consulting and Clinical Psychology*. *76*(5): 893–9. doi:10.1037/0022-006X.76.5.893

Nentjes, L., Bernstein, D., Arntz, A., Van Breukelen, G., & Slaats, M. (2015a). Examining the influence of psychopathy, hostility biases, and automatic processing on criminal offenders' Theory of Mind. *International journal of law and psychiatry*, *38*, 92-99.

Nentjes, L., Bernstein, D. P., Arntz, A., Slaats, M. E., & Hannemann, T. (2015b). Theory of mind, social desirability, and unlikely symptom reporting in offenders with and without psychopathy. *The Journal of nervous and mental disease*, *203*(8), 596-603.

Prichard, J. C. (1837/1973). A treatise on insanity and other disorders affecting the mind. New York: Arno.

Perner, J., Wimmer, H., 1985. ''John thinks that Mary thinks that. . .'': attribution of second-order false beliefs by 5- to 10- year-old children. *J. Exp. Child Psychol. 39*, 437–471.

Postel, J., & Postel, M. (1988). Esquirol et la monomanie homicide. *Histoire des sciences médicales*, *22*(2), 181-186.

Paulhus, D. L., Neumann, C. S. & Hare, R. D. (in press). Manual for the Self–Report Psychopathy Scale. Toronto: Multi-Health Systems.

Patrick, C. J., Hicks, B. M., Krueger, R. F., & Lang, A. R. (2005). Relations between psychopathy facets and externalizing in a criminal offender sample. *Journal of Personality Disorders*, *19*, (4), 339-356.

Patrick, C. J. (2007). Antisocial personality disorder and psychopathy. In W. O'Donohue, K. A. Fowler, & S. O. Lilienfeld (Eds.), Personality disorders: Towards the DSM-V (pp. 109–166). New York, NY: Sage Publications. 339–356. doi:10.1521/pedi.2005.19.4.339

Patrick, C. J., Hicks, B. M., Nichol, P. E., & Krueger, R. F. (2007). A bifactor approach to modeling the structure of the psychopathy checklist revised. *Journal of personality disorders*, *21*(2), 118-141.

Patrick, C. J., & Bernat, E. M. (2010). Neuroscientific foundations of psychopathology. In T. Millon, R. F. Krueger, & E. Simonsen (Eds.), Contemporary directions in psychopathology: Scientific foundations of the DSM-V and ICD-11 (pp. 419–452). New York, NY: Guilford.

Quesque, F., & Rossetti, Y. (2020). What Do Theory-of-Mind Tasks Actually Measure? Theory and Practice. *Perspectives on Psychological Science*, *15*(2), 384-396.

Rush, B. (1812). Medical inquiries and observations upon the diseases of the mind. Philadelphia: Kimber & Richardson.

Robbins, P. C., Monahan, J., & Silver, E. (2003). Mental disorder, violence, and gender. *Law and Human Behavior*, *27*(6), 561-571.

Richell, R. A., Mitchell, D. G., Newman, C., Leonard, A., Baron-Cohen, S., & Blair, R. J. R. (2003). Theory of mind and psychopathy: can psychopathic individuals read the 'language of the eyes'?. *Neuropsychologia*, *41*(5), 523-526.

Russell, T. A., Tchanturia, K., Rahman, Q., & Schmidt, U. (2007). Sex differences in theory of mind: A male advantage on Happé's "cartoon" task. *Cognition and emotion*, *21*(7), 1554-1564.

Smith, R. J. (1978). Personality and psychopathology: A series of monographs, texts, and treatises. New York: Academic Press.

Smith, R. J. (1978). The psychopath in society. New York, NY: Academic.

Stevens, G. F. (1993). Applying the diagnosis Antisocial Personality to imprisoned offenders. *Journal of Offender Rehabilitation*, *19*(1/2), 1-26.

Salekin, R. T., Rogers, R., & Sewell, K. W. (1997). Construct validity of psychopathy in a female offender sample: A multitrait–multimethod evaluation. *Journal of abnormal psychology*, *106*(4), 576.

Stone, V.E., Baron-Cohen, S., Knight, R.E., 1998. Frontal lobe contributions to theory of mind. *J. Cogn. Neurosci. 10*, 640–656.

Shamay-Tsoory, S. G., & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. *Neuropsychologia*, *45*(13), 3054-3067.

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*, 617–627.

Shamay-Tsoory, S. G., Harari, H., Aharon-Peretz, J., & Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex*, *46*(5), 668-677.

Sommer, M., Sodian, B., Döhnel, K., Schwerdtner, J., Meinhardt, J., & Hajak, G. (2010). In psychopathic patients emotion attribution modulates activity in outcome-related brain areas. Psychiatry Research: *Neuroimaging*, *182*(2), 88-95.

Samson, D., Apperly, I. A., Braithwaite, J., Andrews, B., & Bodley, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1255–1266.

Skeem, J. L., Polaschek, D. L., Patrick, C. J., & Lilienfeld, S. O. (2011). Psychopathic personality: Bridging the gap between scientific evidence and public policy. *Psychological Science in the Public Interest*, *12*(3), 95-162.

Surtees, A., Butterfill, S., & Apperly, I. A. (2012). Direct and indirect measures of Level-2 perspective- taking in children and adults. *British Journal of Developmental Psychology*, *30*(1), 75–86.

Seidel, E. M., Pfabigan, D. M., Keckeis, K., Wucherer, A. M., Jahn, T., Lamm, C., & Derntl, B. (2013). Empathic competencies in violent offenders. *Psychiatry research*, *210*(3), 1168-1175.

Sandvik, A. M., Hansen, A. L., Johnsen, B. H., & Laberg, J. C. (2014). Psychopathy and the ability to read the "language of the eyes": Divergence in the psychopathy construct. *Scandinavian journal of psychology*, *55*(6), 585-592.

Sandvik, A. M. (2014). Psychopathy-the heterogeneity of the construct.

Seara-Cardoso, A., Sebastian, C. L., Viding, E., & Roiser, J. P. (2016). Affective resonance in response to others' emotional faces varies with affective ratings and psychopathic traits in amygdala and anterior insula. *Social neuroscience*, *11*(2), 140-152.

Talwar, V., Gordon, H. M., & Lee, K. (2007). Lying in the elementary school years: verbal deception and its relation to second-order belief understanding. *Developmental psychology*, *43*(3), 804.

Taubner, S., White, L. O., Zimmermann, J., Fonagy, P., & Nolte, T. (2013). Attachment-related mentalization moderates the relationship between psychopathic traits and proactive aggression in adolescence. *Journal of abnormal child psychology*, *41*(6), 929-938.

Viding, E., Sebastian, C. L., Dadds, M. R., Lockwood, P. L., Cecil, C. A., De Brito, S. A., & McCrory, E. J. (2012). Amygdala response to preattentive masked fear in children with conduct problems: the role of callous-unemotional traits. *American journal of psychiatry*, *169*(10), 1109-1116.

Viding, E. (2019). Psychopathy: A Very Short Introduction (1st ed., p. 67). New York: Oxford University Press.

Vonk, J., Zeigler-Hill, V., Ewing, D., Mercer, S., & Noser, A. E. (2015). Mindreading in the dark: Dark personality features and theory of mind. *Personality and Individual Differences*, *87*, 50-54.

Wellman, H.M., 1990. The Child's Theory of Mind. MIT Press, Cambridge, MA.

Weiler, B. L., & Widom, C. S. (1996). Psychopathy and violent behaviour in abused and neglected young adults. *Criminal behaviour and mental health*, *6*(3), 253-271.

Wheeler, S., Book, A. & Costello, K. (2009). Psychopathic traits and perceptions of victim vulnerability. *Criminal Justice and Behavior*, *36*, 635–648.

Wynn, R., Høiseth, M. H., & Pettersen, G. (2012). Psychopathy in women: theoretical and clinical perspectives. *International journal of women's health*, *4*, 257.

Yang, Y., Raine, A., Colletti, P., Toga, A. W., & Narr, K. L. (2011). Abnormal structural correlates of response perseveration in individuals with psychopathy. *The Journal of neuropsychiatry and clinical neurosciences*, *23*(1), 107-110.

APPENDIX A
The 20 items from the Psychopathy Checklist-Revised
(2 Factor Model; Hare et al., 1990)

Items		Factor (1, 2) *
1	Glib and Superficial Charm	1
2	Grandiose Self Worth	1
3	Stimulation Seeking	2
4	Pathological Lying	1
5	Conning Manipulative	1
6	Lack Remorse or Guilt	1
7	Shallow Affect	1
8	Callous/Lack of Empathy	1
9	Parasitic Orientation-Lifestyle	2
10	Poor Behavioural Control	2
11	Sexual Promiscuity	2
12	Early Behavioural Problems	2
13	Lack of Realistic Goals	2
14	Impulsivity	2
15	Irresponsibility	2
16	Failure to Accept Responsibility	1
17	Many short-term relationships	2
18	Juvenile Delinquency	2
19	Revocation of Conditional Release	2
20	Criminal Versality	2

* Factor 1= Interpersonal/Affective traits; Factor 2= Antisocial lifestyle traits

APPENDIX B

Search Databases and Terms

The search was performed in four major databases and included PubMed, PsycInfo (OVID), Wiley Online Library and Google Scholar. Total number of records identified through all databases was 2042. The search terms and filters used for each of the databases are specified below.

PubMed: *2000-2021. Filter in Body- All words including the following search terms*

1)   (« Theory of mind ») AND Psychopathy

2)   (« Mentalization ») AND Psychopathy

3)   (« Social perception ») AND Psychopathy

4)   (« Affective theory of mind ») AND Psychopathy

5)   (« Cognitive theory of mind ») AND Psychopathy

6)   « Mentalizing » AND Psychopathy:

7)   « Perspective taking » AND Psychopathy:

8)   « Perspective-taking » AND Psychopathy

PsycInfo (OVID): *2000-2021. Filter with the following search terms*

1)   « Theory of mind » AND Psychopathy

2)   « Mentalization » AND Psychopathy

3)   « Social perception » AND Psychopathy

4)   « Affective theory of mind » AND Psychopathy

5)   « Cognitve theory of mind » AND Psychopathy

6)   « Mentalizing » AND Psychopathy

7)   « Perspective taking » AND Psychopathy

8)   « Perspective-taking » AND Psychopathy

Wiley Online Library: *Filter with advanced search: 2000-2021. Here, I selected the «abstract» option, meaning that the search terms will be necessarily mentioned in the abstracts of the results.*

1)   « Theory of mind » in Abstract and Psychopathy in Abstract

2)   « Mentalization » in Abstract and Psychopathy in Abstract

3)   « Social perception » in Abstract and Psychopathy in Abstract

4)   « Affective theory of mind » in Abstract and Psychopathy in Abstract

5)   « Cognitive theory of mind » in Abstract and Psychopathy in Abstract

6)   « Mentalizing » in Abstract and « Psychopathy » in Abstract:

7)   « Perspective taking » in Abstract and Psychopathy in Abstract

8)   « Perspective-taking » in Abstract and Psychopathy in Abstract

Google Scholar: *2000-2021. Filter allintitle with the following search terms*

1)   allintitle: Theory of mind Psychopathy (2000-2021)

2)   allintitle: Mentalization Psychopathy (2000-2021)

3)   allintitle: Mentalizing Psychopathy (2000-2021)

4)   allintitle: Perspective-taking Psychopathy (2000-2021)

5)   allintitle: Perspective taking Psychopathy (2000-2021)

6)   allintitle: Social Perception Psychopathy (2000-2021)

7)   allintitle: Affective theory of mind Psychopathy (2000-2021)

8)   allintitle: Cognitive theory of mind Psychopathy (2000-2021)