The Role of Empathy in Family Violence

By

Elizabeth Fitzmaurice

Centre for Forensic and Criminological Psychology

University of Birmingham

July 2015

Acknowledgements

I would like to thank my academic supervisor, Dr Louise Dixon, for all of her support and pearls of wisdom throughout the project. I will be eternally grateful for the effort both Louise and Professor Erica Bowen put into developing the research project and supporting me throughout. I would like to thank all of the staff at [redacted] who assisted in the smooth-sailing of the research implementation. Specifically, [redacted] for allowing the research to be aligned to his department and to [redacted] for the warmth and support she provided me throughout the process.

Aside from my thesis, I would like to make special thanks to Victoria Wilkes, Joanne Gliddon and James Bickley. As placement supervisors they provided me with an invaluable amount of support and encouragement. I've learnt life-long lessons from all three and could not have made it through the Doctorate without their endless faith in me.

The most important acknowledgement is to my parents, Nicki & David, the most intelligent people I know. They have supported me in every way possible, every day of my life. I hope one day I can make them as proud of me as I am of them. Thank you to my baby siblings who make me burst with pride. My second mother and only sister Cath, thank you for making sure I ever graduated in the first place. And Dan, who fuels the fire of sibling rivalry by recounting every mistake I've ever made and claiming to be the most successful. Thank you for maintaining my fear of failure and need to raise the bar.

Love always and forever to my partner Jack who has supported me in every way to follow my dreams. Thank you for your love, sense of humour and ability to

tolerate a hormonal student! And thanks to his parents and siblings, for keeping me loved, fed and watered.

An obvious thank you to all of my friends, who have been calling me Dr Fitz from the day I enrolled and gave me no option but to qualify! Big special thanks to Liz, Harriet, Sophie and Ashleigh, my forever friends who have provided the perfect balance of distraction and motivation. And finally to Chloe, Danielle, Nicola & Katie, we started this journey together and I am so proud of each and every one of you. You are amazing people and equally amazing Psychologists.

Abstract

This thesis explores the role of empathy in family violence, specifically child maltreatment (CM) and intimate partner violence (IPV). Chapter 1 provides an introduction to the thesis, introducing the phenomenon of family violence as a public health concern and detailing current knowledge about risk factors. With the thesis focusing specifically on empathy as a risk factor, Chapter 1 introduces the construct of empathy, its development and relevance to violence. Chapter 2 then explores the relationship between empathy and CM in a systematic literature review of 17 studies published between 1985 and 2014. Results found that maltreating parents demonstrate significantly lower empathic capacity and that this relationship is stronger for cognitive than affective empathy. Limitations of reviewed studies included small sample sizes, lack of UK data and recruitment of parents. Chapter 3 presents a critical analysis of the Interpersonal Reactivity Index (IRI; Davis, 1980) which is a commonly used multidimensional assessment of empathy. This critique demonstrates that the measure has good reliability, validity and a wide range of normative data available. Limitations of the measure include the validity of the Fantasy subscale and limitations of the IRI due to it being a questionnaire-based assessment. Chapter 4 presents a research report which explores the presence of empathy and emotional recognition skills in IPV (n=30), violent (n=20) and non-violent (n=20) offenders. Results found that IPV participants were more likely than NV offenders to interpret fearful faces as sad. Only the IRI personal distress scale (PD) showed a significant relationship with emotion recognition. As a whole group, higher PD scores were related to more errors identifying angry and high intensity faces. Analysing groups separately, IPV offenders PD scores were positively correlated with angry and fearful error scores, violent offenders PD scores were positively correlated with anger errors

scores and non-violent offenders PD scores were positively correlated with angry, high and low intensity error scores. The thesis conclusions are presented in Chapter 5 which identifies that empathy plays a role in family violence, although its influence in CM and IPV appears to be different. The chapter also highlights need for future research and to consider CM and IPV together, given high rates of co-occurrence.

Contents

	Page
List of Tables ……..	**7**
List of Figures ……	**8**
List of Appendices …….	**8**
List of Abbreviations …….....	**9**
Chapter 1: Introduction ……...	**11**
Chapter 2: A Systematic Literature Review Assessing Whether There Are Empathy Differences Between Maltreating and Non-Maltreating Parents …..	**27**
- Introduction ……	**29**
- Method ……...	**36**
- Results ……	**62**
- Discussion ……..	**81**
- Conclusions and Recommendations ……..	**90**
Chapter 3: Critique of a Psychometric Measure: The Interpersonal Reactivity Index (IRI) ……	**92**
- Introduction ……...	**93**
- Psychometric properties of the IRI ……	**97**
- Reliability ……..	**97**
- Validity ……..	**100**
- Normative Data …….	**108**
- Conclusions ……...	**112**
Chapter 4: A Comparison of Male Intimate Partner Violent Offenders and Non-Violent Offenders on Measures of Empathy …….	**114**
- Introduction ……	**116**
- Method ……...	**126**
- Results ……	**136**
- Discussion ……..	**132**
- Conclusion ……..	**149**
Chapter 5: Discussion ……..	**150**
References …….	**159**
Appendices ……	**180**

List of Tables

		Page
Table 1.	PICO Inclusion/Exclusion Criteria …….	**39**
Table 2.	Characteristics of Studies Included in the Review …….	**45**
Table 3.	Participant Selection and Empathy Measures Used in Studies ……..	**58**
Table 4.	Demographic Variables that Studies Considered ……...	**60**
Table 5.	Type of Empathy Measured in Studies ……..	**67**
Table 6.	Scales of the IRI Used in Studies ……...	**68**
Table 7.	Summary of Outcomes Related to Empathy ……..	**70**
Table 8.	Reliability Coefficients for IRI Scales ……	**99**
Table 9.	IRI Test-Retest Reliability Coefficients……..	**100**
Table 10.	Correlation Coefficients of the IRI and Other Measures of Empathy	**101**
Table 11.	IRI Normative Mean Data ……..	**110**
Table 12.	Descriptive Information of the Research Sample ……...	**128**
Table 13.	Inclusion/Exclusion Criteria for Each Participant Group ……...	**133**
Table 14.	MANCOVA output for IRI and DANVA2-AF Variables …….	**137**
Table 15.	Pearson's Partial Correlation between IRI and DANVA2-AF Output Variables for Whole Participant Group (N=70) …….	**140**

Table 16. Pearson's Partial Correlation between IRI and DANVA2-AF Output Variables for IPV Group (N=30) …… **141**

Table 17. Pearson's Partial Correlation between IRI and DANVA2-AF Output Variables for V Group (N=20) ……... **141**

Table 18. Pearson's Partial Correlation between IRI and DANVA2-AF Output Variables for NV Group (N=20) …… **142**

List of Figures

Page

Figure 1. Literature Review Study Selection Process ……..... **40**

List of Appendices

Page

Appendix 1. Risk and Protective Factors for Child Maltreatment, as Identified by the Child Welfare Information Gateway (2004) ……. **180**

Appendix 2. Details of Database Search Strategies for Literature Review ……. **181**

Appendix 3. Template for Contacting Researchers and List of Authors Contacted **184**

Appendix 4. Articles that were not Available Online and were Posted from Authors ……... **184**

Appendix 5. Articles Searched in Full That Did Not Meet the Inclusion/ Exclusion Criteria ……... **186**

Appendix 6. Quality Assessment Tool used in the Literature Review …… **191**

Appendix 7. Quality Assessment Results from the Literature Review ……... **192**

Appendix 8. Data Extraction Proforma for Literature Review ……... **195**

Appendix 9. References of the Studies Included in the Literature Review ……. **197**

Appendix 10 Participant Information Sheet ……. **199**

Appendix 11 Participant Consent Form ……... **201**

List of Abbreviations

AAPI:	Adolescent-Adult Parenting Inventory
CASP:	Critical Appraisal Skills Programme
CAPI:	Child Abuse Potential Inventory
CDQ:	Child Development Questionnaire
CM:	Child Maltreatment
CTS:	Conflict Tactics Scale
CWIG:	Child Welfare Information Gateway
EPHPP:	Effective Public Health Practice Project
EQS:	Empathy Quotient – Short
HES:	Hogan's Empathy Scale
IPV:	Intimate Partner violence
IRI:	Interpersonal Reactivity Index
PCL-R:	Psychopathy Checklist – Revised
POQ:	Parental Opinion Questionnaire
PPES:	Parent/Partner Empathy Scale
QMEE:	Questionnaire Measure of Emotional Empathy
RSVP:	The Risk for Sexual Violence Protocol
SAPROF:	Structures Assessment of Protective Factors for Violence Risk
STAXI:	State-Trait Anger Inventory

Chapter 1

Introduction

The term ‘family violence’ encompasses any violence (physical, emotional, psychological, sexual) or neglect that occurs in the context of a family environment, involving violence between any combination of parents, siblings, partners and children, including elder abuse (Browne & Herbert, 1999). The two most commonly cited forms of family violence are child maltreatment (CM) and intimate partner violence (IPV). These are the focus of this thesis.

The UK government recognises that CM is a multi-faceted concept encompassing physical, emotional and sexual abuse and/or neglect towards children (DfE, 2013). The World Health Organisation (WHO; Butchart, Putney, Furniss, & Kahane, 2006) defined CM as:

“all forms of physical and/or emotional ill-treatment, sexual abuse, neglect or negligent treatment or commercial or other exploitation, resulting in actual or potential harm to the child’s health, survival, development or dignity in the context of a relationship of responsibility, trust or power.”

What this definition fails to explicitly denote is that CM does not only refer to an individuals’ act of maltreatment, but a ‘failure to act to prevent harm’ (Department for Education, DfE, 2013). Thus CM also involves exposure to IPV, as research suggests this has an adverse effect on a child’s development (Edleson, 1999). This is now embedded into UK approaches and witnessing IPV is considered in parenting and violence risk assessments.

The UK government has recently updated its definition of IPV (Home Office, 2013), which is:

"any incident or pattern of incidents of controlling, coercive, threatening behaviour, violence or abuse between those aged 16 or over who are, or have been, intimate partners or family members regardless of gender or sexuality. The abuse can encompass, but is not limited to: psychological, physical, sexual, financial, emotional."

The prevalence and impact of family violence warrants it a significant public health concern (Krug, Dahlberg, Mercy, Zwi & Lozano, 2002). WHO (2015) found a lifetime prevalence of IPV victimisation to be 28.3% for women and 14.7% for men in England and Wales, with a direct annual cost to the UK of £15.7 billion (Walby, 2009). Severe CM is approximated to affect 24.5% of children and young people with lifetime prevalence's retrospectively reported by adults ranging between 25-42% (Radford et al., 2011). Professional bodies, such as the National Society for the Prevention of Cruelty to Children (NSPCC) and the National Institute for Clinical Excellence (NICE), have found it difficult to approximate the total cost of CM given its wide-ranging impact. However, considering some of the costs: child sexual abuse is approximated to cost the UK £3.2 billion, approximately £1.14 billion is spent annually on social services costs for "abuse and neglect" (Walby, 2004) and the weekly cost of placing one child in local authority care is £696 a week (Curtis, 2007). The collated statistics regarding family violence are likely to only represent the 'tip of the iceberg' and significantly underestimate the true impact due to confounding factors such as low rates of reporting (Gilbert et al., 2009), low rates of conviction (NSPCC, 2013) and difficulties estimating the long-term implications such as personal costs, the costs of counselling, increased need for mental health services and

involvement with the criminal justice system (WHO, 2013). For example, it has been suggested that the reality of CM is tenfold the official statistics (Lansdown, 2005).

The impact and costs of CM and IPV are calculated individually, rather than together as 'family violence'. This often portrays an assumption that they are two separate phenomena. However the two constructs are interlinked. The co-occurrence of CM and IPV is high, with research typically citing a 30 – 60% overlap (Edleston, 1999). A cumulative effect has been evidenced, with those experiencing both CM and IPV at greater exposure to the negative impact than those who experience one form of CM (Chiodo, Leschied, Whitehead, & Hurley, 2008). In addition to being at higher risk of physical CM (Hamby, Finkelhor, Turner & Ormrod, 2010), exposure to IPV is in itself harmful (Brandon & Lewis, 1996).

Furthermore research indicates that there is an 'intergenerational transmission' of family violence (Dixon, Hamilton-Giachritsis & Browne, 2005). Although prospectively a small number of those abused in childhood perpetrate CM as adults, a disproportionality high number of offenders and perpetrators of CM have been victims of CM (Barlett & Easterbrook, 2015; Williams, Papadopoulou & Booth, 2012). Additionally, witnessing IPV as a child increases the likelihood that an individual will perpetrate or be a victim of IPV (Russell, Springer & Greenfield, 2010), with this being an item included in risk assessments of IPV perpetration (Kropp, Hart, Webster & Eaves, 1995). Being a victim or perpetrator of IPV has also been evidenced as a risk factor for perpetrating CM (Slep & O'Leary, 2005), although this is something less ingrained into current practice.

In summary, family violence is evidenced to have a lifetime detriment to victims and a financial impact on the economy. As a societal problem it is important

for research to understand factors that increase the risk of perpetrating family violence in order to inform preventative policy and intervention strategies. Given the co-occurrence and intergeneration transmission of CM and IPV, it is important that policy makers and clinicians acknowledge these links and assess for perpetration and victimization of all forms of family violence regardless of the reason for contact. Research to support this has been conducted by Dixon, Browne, Hamilton-Giachritis & Ostapuik (2010) who found three distinct patterns of reciprocal family violence (Paternal/Maternal, Hierarchical and Reciprocal). Using a sample of 67 families with suspected CM and IPV, they found that the most common form of co-occuring family violence (43.3% of cases) involved the parent who experienced IPV perpetrating CM towards their child (Hierarchical). A similarly frequent pattern (occurring in 41.8% of cases) was of both parents perpetrating violence towards each other and their child (Reciprocal). Paternal/Maternal family violence, which involves one parent/partner perpetrating violence to both partner and child, occurred in 14.9% of cases.

Empathy

This thesis focuses specifically on the role of empathy in the perpetration and maintenance of family violence, as empathy is considered to be an important determinant of how humans relate to each other. Evidence suggests that empathy underpins pro-social behaviour and community cohesion and higher levels of empathy are related to helpful and caring behaviour (Batson & Ahmad, 2009). In line with this, aggression and anti-social behaviour have been associated with lower levels of empathy (Feshbach, 1964; Feshbach & Feshbach, 1969; Gibbs, 1987; Miller & Eisenberg 1988). Accordingly, empathy is hypothesised to be a modulator of family violence.

Origins of Empathy

In order to explore the role of empathy in family violence, it is necessary to understand where the concept of empathy originates and how it is defined in the literature. The etymology of empathy can be found in the Greek concepts ‘en pathos’ and ‘empatheia’ which respectively refer to having feeling and a passionate emotional state. In German philosophy, the concept of empathy was discussed in the late nineteenth century by those such as Vischer (1873) and Lipps (1897) using the term Einfühlung. They discussed empathy in the context of aesthetics to propose that as humans we empathise with our objective environment, providing affective qualities to what we visualise (note that they considered Einfühlung to be synonymous with sympathy). Einfühlung was translated into the English concept ‘empathy’ in 1909 by Edward Titchener.

The concept of empathy can also be found in evolutionary literature. Although not coining the concept empathy, in 1871 Darwin discussed altruistic behaviour and considered that helping others in distress was motivated by the observers desire to reduce their own negative affect (affective empathy). Darwin considered that empathy was favoured by natural selection as empathic individuals were likely to be of benefit to a community and result in a higher number of offspring. In more recent evolutionary literature, de Waal (2008) has expanded upon Darwin’s theory by linking empathy to the mirror neuron system. De Waal (2008) has suggested that the ability to understand and experience the perspective/affect of another may first be found in parent-child interactions long before evolution of the human species: a parent feeling compelled to nurture their child in response to their child’s cries and smiles. This empathic capacity evolved to apply outside of the realm of parenting, allowing us to empathise with others we encounter. de Waal identified that the fact adults retain

distress signals suggest that such signals maintained utility outside of the parent-child interaction.

De Waal proposed that the empathic experience is a fast hard-wired response that is difficult to control, giving the example of having to cover your eyes during a horror film as you have already identified with the character. De Waal identified however that processes exist that allow an individual/primate to detach from such empathic capacity, such as primates/humans killing each other in a brutal nature. For the purpose of the present thesis, it is important to understand these processes and how they link to family violence.

Definitions of Empathy

Throughout history, researchers and clinicians have debated which constructs empathy encompasses. The divergence is such that Batson (2009) dedicated a book chapter to describing eight different ways to describe empathy: (1) knowing the internal state of another; (2) matching the posture or neural response of another; (3) feeling how another feels; (4) understanding the situation of another; (5) imagining the thoughts and cognitions of another; (6) imagining how you would feel in the situation of another; (7) feeling distress at the suffering of another; (8) feeling for a person who is suffering. These separate constructs demonstrate the conflicts that exist as to whether empathy is:

1. conscious or subconscious
2. a process or an outcome
3. affective or cognitive
4. congruent with the feeling or with the welfare of another
5. a trait or a state

The first debate of conscious vs subconscious is one less focused upon in current literature, although it was debated by philosophers in the early stages of empathy conceptualisation. Empathy, first coined 'Einfühlung' by Lipp (1897), was considered an unconscious processes of inner-imitation. Although Lipp's theories were proposed to explain optical illusions, Titchener (1909) applied the concept to experimental psychology, ultimately forming the basis of empathically-derived processes such as emotional contagion (Allport, 1924). This concept of 'inner-imitation' has more recently regained attention and support with neuroimaging advancements implicating the mirror neuron system in empathy (Gallese, 2001; Yamada & Decety, 2009). Mirror neurons essentially fire when observing the behaviour of others, thus forming a neural imitation of another's actions. Linking mirror neurons to empathy suggests that empathy is underpinned by imagining how another feels.

In contrast to an unconscious conceptualisation of empathy, some suggest that empathy is underpinned by an egocentric drive. Hoffman (2000) described empathy in the context of child development, explaining that the second stage of empathy development, egocentric empathy, is characterised by children offering help in line with their own needs. Thus, although behaviour may be empathic, some studies suggest that the outcome of empathic behaviour is dependent upon the observer's own needs. Other researchers have also found that empathic behaviour involves self-orientated goals (Batson & Coke, 1981; Eisenberg, 1982).

The above examples introduce the second conflict of whether empathy is a process or an outcome (Davis, 1994). Again this disparity is relatively neglected in the literature. The example of an egocentric child involves both the process and outcome of empathy, the 'process' being how the child inferred ones behaviour and

the outcome being the egocentric behaviour that resulted. When describing empathy with relation to the therapeutic relationship, Rogers (1975) conceptualised empathy as a process of "temporarily living in his/her life, moving about in it delicately without making judgments, sensing meaning of which he/she is scarcely aware" (p.4). However, many theorists focus on emotional responses as a measure of empathy, these being examples of empathy as an outcome. Thus researchers who measure empathy behaviourally, for example with facial expressions or emotional reactions, operationalise empathy as an outcome.

The idea that empathy can be considered solely with relation to its outcome has been contended by Polaschek (2003) who proposed that behaviour is determined not only by ones affect, but is influenced by factors such as competing interests and personal gains (which again would relate to ego-centric empathy). Viewing empathy as a multi-stage process, the debate is whether empathy is the final stage (outcome), the stages before the outcome (the process) or both the process and outcome.

A more often debated question with regard to empathy is whether it is a cognitive or affective construct. Early descriptions of empathy (Lipps, 1897; Titchener, 1909) viewed empathy as a cognitive construct and early measures of empathy such as Hogan's Empathy Scale (HES; 1969) are underpinned by the theoretical stance that empathy is a cognitive process occurring "without actually experiencing that person's feelings" (Hogan, 1969, p.308). This construct is thus synonymous with those such as perspective-taking and theory of mind (Vollm et al., 2006). Others however consider empathy to be an affective state, with a further divide existing as to whether empathy is an emotion congruent with the affect of somebody else (Eisenberg & Strayer, 1987) or congruent with the welfare of somebody else (Batson & Coke, 1981).

The final debate with relation to empathy is whether empathy is a state or a trait. As a state, empathy is considered as a construct that can be induced (Batson et al., 1997) and is variable depending upon the context and who the person is interacting with. Trait empathy, often termed dispositional empathy, is conceptualised as a more stable personality characteristic.

Some theorists however do not consider empathy is as distinct as suggested in the above debates, proposing that processes interact and depend upon each other. Through this lens, cognitive empathy, such as perspective taking, is the first component of a wider empathy process. The early stages are able to predict latter stages, such as affective empathy outcomes (Strayer, 1987). In line with this inclusive construct, Cohen & Strayer (1996) define empathy as "the ability to understand and share in another's emotional state or context".

Despite the disparity between definitions and models of empathy, Barnett and Mann (2013) argue that most models suggest empathy is influenced by five sub-components: emotional contagion, perspective-taking, belief in the worth of others, situational factors and reaction to personal distress. They suggest the latter two factors negatively impact upon empathy by inhibiting the other three factors.

The current thesis adheres to a view of empathy that includes both cognitive and affective constructs. The thesis considers the experience of empathy (the process) to be distinct from the outcome and explores factors that may contribute to a disparity between process and outcome. For example, the thesis explores why an individual may be able to feel the emotions of others but not behave in an empathic way. The thesis also suggests that trait and state empathy are different constructs.

The Development of Empathy

Developmentally, the first stages of empathic capacity (such as perspective taking) can be seen in young children. Although specific ages at which such milestones are reached, it is widely accepted that a child's understanding of another originates from an egocentric perspective and later develops to understand that an individual may have a perspective different from their own. The latter skill is coined 'theory of mind' and Piaget (1965) suggested that this develops between the ages of 3 and 5. Evidence of empathic behaviours at an earlier age are argued to include new-borns who cry in response to hearing another child cry (Martin & Clark, 1982) and infants who become distressed when another child is distressed (Knafo et al., 2008) or demonstrate helpful behaviours when viewing someone in distress (Zahn-Waxler et al., 1992).

Research has demonstrated individual differences in the development of empathy from early ages. Research examining the heritability of empathy (using twin studies) found that between ages 2 and 3 years old, a third of empathic capacity could be linked to genetic influences (Knafo et al., 2008). Relevant to the current thesis is the finding that an infant's temperament and emotional regulation predicted their empathic capacity in adolescence. This suggests that high distress and poor emotional regulation is linked to poor empathic development.

Explanations for variations in emotional regulation can be found in attachment models. Crittenden's (2006) model of attachment (Dynamic Maturational Model) suggests that a child develops an ability to understand their own emotions and the emotions of others via modelling from their own parents. A child learns to manage their own emotions through the support of their parent, who guides them to self-

soothe during parenting. Disruptions to this process can occur in a number of ways. If a child learns that expressing their emotions is dangerous (e.g. if I cry I will be hurt, crying does not result in a respond), this child will develop a strategy that inhibits their affective experiences, rather surviving using cognitive information (Type A attachment). A child who receives inconsistent parenting is unable to predict the outcome of their behaviour and thus cognitions are less helpful for survival. Such children are guided by their affective experience and respond accordingly.

Although these strategies may have been adaptive at early stages of development, the utilisation of such strategies on a long-term basis are likely to result in emotion regulation difficulties. Linking this to de Waal's (2008) evolutionary theory of empathy, being overwhelmed with one's own emotions may be a factor that inhibits empathic capacity and increases perpetration of non-empathic behaviours such as violence. Thus would suggest that perpetrators of violence do not have an absence of empathy, rather obstacles exist that prevent perpetrators accessing their empathic capacity.

Empathy and Violence

As a factor that is known to mediate general violence, empathy is also suggested to be an influential factor linked to family violence. However, the way in which it is involved and how it influences family violence is less well understood. Thus intervention and policy makers are left to make assumptions from evidence relating to empathy and general violence in order to inform practice.

The most comprehensive review exploring the relationship between violence and empathy was conducted by Miller and Eisenberg (1988). They found that cognitive empathy significantly negatively correlated with physical aggression when

questionnaire measures of empathy were used with violent offenders. However when using task-based measures of empathy, picture or vignette assessment tools yielded a non-significant result, albeit a negative relationship. A more recent literature review by Joliffe and Farrington (2004) similarly found a negative correlation between empathy and offending behaviour (a stronger correlation for cognitive than affective empathy). A specific focus on violent offenders found a significant correlation only for cognitive empathy. Therefore, reviews of the literature suggest that there is a relationship between violent offending behaviour and cognitive empathy. The relationship between affective empathy and violence is less clear. The reviews also indicated that utilisation of different assessment tools results in different outcomes.

A finding of lower cognitive but not affective empathy has been replicated in some studies that explore correlates of violent offending (Lauterbach & Hosser, 2007; Seidal et al., 2013) and anger arousal (Day, Mohr, Howells, Gerance & Lim, 2012). Conversely, other studies have found affective empathy deficits in violent offenders (Beven, O'Brien-Malone & Hall, 2004; Joliffe & Farrington, 2007). Such disparity in research findings suggests that there are confounding factors in the data that have not been accounted for. For example, Jolliffe and Farrington (2004) found that when IQ and economic status were controlled for the relationship between cognitive empathy and offending remained but that of affective empathy and offending dissipated in offending populations. Authors of these findings however provide no theory as to why affective empathy is impacted by IQ and economic status rather than cognitive. Logically, it would be assumed that as IQ is a cognitive function, its effect would correlate with cognitive empathy. There is some suggestion that these findings may be related to the intellectual requirements for assessing affective empathy. A cohort of studies have demonstrated lower cognitive functioning in offenders compared with

community samples (Caddick & Webster, 1998). As suggested by Beven, O'Brien-Malone and Hall, (2004), items relating to affective empathy may require a greater level of insight and intellectual capability. However, it would be expected that as these participants had lower cognitive functioning, their cognitive empathy was also negatively impacted.

Alternatively, some research suggests that empathy deficits in perpetrators of violence are not related to general skill deficits per se but rather perpetrators have impairments in identifying specific emotions. Utilising facial emotion recognition tasks, a meta-analysis of 20 studies found a "robust link" between violence and the recognition of fearful faces (Marsh & Blair, 2008).

Recently, Marshall and Marshall (2011) proposed a three-stage model of empathy informed by research examining empathy in violent offenders. Primarily the emotion has to be recognised (stage 1), then the perspective of the person is taken (stage 2). Finally, both naming and understanding the emotion facilitates feeling the emotion of another person (stage 3). With regard to perpetrators of violence, Marshall and Marshall (2011) propose that having a hostile relationship with the subject or being overwhelmed by emotion inhibits achievement of stages 2 and 3.

Applying findings from violence literature to family violence would suggest that the relationship perpetrators have with family members and level of emotional arousal modulates empathic capacity. If parallel in findings between the general violence and family violence literature is hypothesised, it would be expected that perpetrators of family violence have lower levels of cognitive empathy, with less certainty regarding affective empathy. It may also be expected that deficits in

recognition of negative emotions is found and for there to be different findings between questionnaire and task-based assessments.

The assumptions drawn from the general violence literature can be seen in a number of practices related to family violence. For example, risk assessments of IPV assess a perpetrators attitudes towards assaulting their partner (Kropp, Hart, Webster & Eaves, 1995). Parenting assessments also assess a parent's ability to perspective take as well as assessing affective empathy ("emotional warmth") (NSPCC, 2014).

Aims of the Thesis

This thesis aims to explore the role of empathy in family violence, specifically exploring IPV and CM, in order to better understand the validity of assumptions made in practice, that lower levels of empathy are associated with higher risk of perpetrating family violence. In the next chapter, the thesis explores the role of empathy in IPV by conducting a literature review of all studies relating to empathy and parental CM. This review was concerned with understanding previous findings regarding the relationship between empathy and CM, whether empathy relates differently to male and female perpetrators and whether different measures of empathy provided different outcomes.

Chapter 3 critically analyses the psychometric properties of the Interpersonal Reactivity Index (IRI; Davis, 1980), which is one of the most commonly used assessments of empathy. The critical analysis considers the reliability and validity of the IRI as an assessment of empathy.

Chapter 4 explores the relationship between empathy and IPV, presenting findings from an empirical study exploring the differences in empathy between offenders who have perpetrated IPV, general violent and non-violent offences. This

Chapter aimed to explore whether facial emotion recognition and empathic deficits existed between IPV and other offender sub-types. However the study was also concerned with whether the empathic profiles of IPV offenders were significantly different from violent offenders, given that they are currently provided with different intervention pathways. The thesis conclusions are presented in Chapter 5, which considers how research relating to empathy and CM differs from the empathy and IPV research literature and how this can inform practice.

Chapter 2
A Systematic Literature Review Assessing Whether There Are Empathy Differences between Maltreating and Non-Maltreating Parents

Abstract

Child maltreatment (CM) affects approximately 24.5% of children and young people in the UK (Radford et al., 2011). A plethora of research has focused on risk factors for CM perpetration, however few studies have focused on the role of empathy. The current review aimed to explore whether there are empathic differences between maltreating and non-maltreating parents to better understand this relationship. 17 studies, published between 1981 and 2013, were deemed suitable for the review, as determined by inclusion/exclusion criteria and a quality assessment. 16 of the 17 studies found that maltreating parents had significantly lower levels of empathy than non-maltreating parents, using both standardised and non-standardised assessments. This effect was seen more-so for cognitive empathy than affective empathy. Three studies examined parental empathy towards their own child and found that maltreating parents were significantly less able to empathise with their child than non-maltreating parents. Interestingly, three studies that used task-based assessments found differing results regarding the empathic capacity of maltreating parents compared to non-maltreating parents. The review highlights limitations of the reviewed studies and important factors that future research should address.

Introduction

The global impact of child maltreatment (CM) is incalculable (NICE, 2009). In England there has been a steady increase in the number of 'looked after children'. At the end of March 2014 there were 68,840 looked after children with the two most common reasons being 'abuse or neglect' and 'family dysfunction'. Of all looked after children, 58% were subject to a care order (Department for Education, DfE, 2014). Research has demonstrated that victims of child abuse have an increased likelihood of education difficulties (Kendall-Tackett & Eckenrode, 1996), future mental health difficulties (Golding, 1999), criminal justice involvement (Jaffee, Caspi, Moffitt & Taylor, 2004) and becoming a perpetrator of CM (Pears & Capaldi, 2001). With such bleak outcomes it is essential that we develop a better understanding of factors that increase the risk of an individual perpetrating CM as well as factors that stop a CM perpetrator from continuing to cause children harm.

Empathy is a factor that has been linked to the development of secure parent-child attachments. In a deductive manner, lack of empathy is often considered to be a risk factor for perpetrating CM. In some forensic fields the role of empathy has been researched extensively. For example, models of empathy in sex offending have been developed (Wastell, Cairns & Haywood, 2009). However, CM literature is much more in its infancy, with the role of empathy less understood. Nonetheless, social work parenting assessments explore the extent to which a parent is able to empathise with their child and intervention for perpetrators of CM include strategies that increase a parent's ability to understand the perspective of their child. Thus present procedures include empathy in their strategies for managing CM, without a strong evidence base. The aim of the present literature review is to synthesise research that explores the relationship between empathy and CM in order increase understanding.

Child Maltreatment

CM is unfortunately widespread globally (WHO, 2014). Exploring CM in the UK, the NSPCC (Radford et al., 2011) conducted a survey with 4,036 children and young people, categorised in three age groups: under 11 years of age, 11 – 17 years of age, 18 – 24 years of age. Results found that 5.9%, 18.6% and 25.3% of children and young people respectively had experienced severe maltreatment, with 5%, 13.4% and 14.5% respectively experiencing severe maltreatment perpetrated by their parent or guardian. For all three age-groups, severe neglect was the most frequent form of maltreatment, severe physical violence the next most frequent, with contact sexual abuse being the least frequent form of severe maltreatment perpetrated by parents or guardians. The study found a lifetime prevalence of 24.5% of experiencing parent/guardian perpetrated maltreatment for those aged 18 – 24 years.

Understanding the context of these findings, it is concerning that only 1.5% of children are referred to child protection services for all abuse types (Gilbert et al., 2009). This suggests that there are a large proportion of CM cases undetected. Thus, in order to improve the way we protect children from maltreatment, we must develop a more comprehensive understanding of risk factors for perpetrating CM.

Risk factors for parent/guardian perpetrated child maltreatment

There is some level of understanding regarding risk factors for CM, with a number of theories developed to explain causes of CM (Azar, 1986; 2008; Milner, 1993; 2003). Although differing in their focus, most models agree that CM occurs as a consequence of a complex interaction between risk and protective factors, with there being no ‘single route’ to CM. A useful way to conceptualise risk factors for CM is via an ecological model (Belsky, 1980; Bronfenbrenner, 1977). The model considers

CM under four levels: individual, relationship, community and societal risk and protective factors. The model and example factors are depicted in Appendix 1.

An alternative theory, the cumulative model, focuses less on the nature of risk factors than the ecological model, and proposes that the more risk factors experienced in a parents' life, the higher the risk for CM perpetration (Appleyard, Egeland, van Dulmen, & Sroufe, 2005). As the name suggests, the cumulative model of CM proposes that maltreatment is due to an accumulation of risk factors.

The most commonly identified risk and protective factors for CM have been collated by the Child Welfare Information Gateway (CWIG, 2004) and considered with relation to the ecological model of CM (Appendix A). It is beyond the scope of this literature summary to explore all risk and protective factors, thus the most commonly found are discussed.

A commonly found 'relationship' risk factor is maternal depression (Conron, Beardslee, Koenen, Buka, & Gortmaker, 2009; Windham et al., 2004), which appears to impact negatively upon parenting style and attachment (Campbell et al., 2004). Hoffman, Crnic and Baker (2004) found that mothers with depression were more likely to have compromised parenting, which impacted negatively upon the emotional regulation of their child and intensified behavioural difficulties. This effect has been shown to reverse with the improvement of depression (Kahng, Oyserman, Bybee, & Mowbray, 2008). Thus, as a risk factor, maternal depression has a two-fold effect, both by impacting upon the way a mother interacts with her child (relationship risk factor) and in the way her child interacts with its environments (individual risk factor).

Parental substance abuse is a further 'relationship' risk factor for CM, with research suggesting parental substance abuse is apparent in a high proportion of

severe CM referrals (Forrester & Harwin, 2006) and in a quarter to a third of all referred cases (Cleaver, Unell & Aldgate, 2011). Forrester and Harwin's (2006) study examined all social service referrals in four London boroughs over one-year (290 families; 534 children). As well as finding that 100 (34.5%) of the families were affected by parental substance misuse, they found that other relationship risk factors were: parental involvement with social services as a child, previous criminal convictions and living in a violent household. The study also indicated that living in temporary accommodation was a 'community' risk factor for CM.

Violence in the household, as identified by Forrester and Harwin (2006), is a more recently identified risk factor for CM. Casaneueva, Martin and Runyan (2009) found that children whose mothers had experienced IPV in the last twelve-months were twice as likely to be reported to child protection services for maltreatment compared with mothers who had not experienced IPV. Brandon et al. (2012) found that 63% of the 139 serious cases (involving the death or serious harm of a child) occurring between 2009 and 2011 involved IPV. This evidence indicates the importance of considering CM from a broader ecological perspective and in relation not only to the child's relationships, but the interactions and relationships a child is exposed to in its environment.

A further risk factor identified in Forrester and Harwin's (2006) study is parental experience of maltreatment as a child. Dixon, Hamilton-Giachritsis and Browne (2005) followed 4351 families during the first 13-months post-birth and found that 6.7% of parents who had experienced maltreatment as a child, compared with 0.4% of parents who had not, were found to have maltreated their child in this period. 62% of this effect was accounted for by the following risk factors: young parenting, history of maternal mental illness, living with a violent adult and poor

parenting styles. Thus parental experience of maltreatment as a child, often termed the intergenerational transmission of CM, demonstrates the complexity and multi-faceted nature of CM and interaction of risk factors.

It is important however to acknowledge that risk factors are not determinate and the complex interaction of CM risk involves the influence of protective factors. An important finding from Dixon et al.'s (2005) study is that although parents who had experienced maltreatment as a child were more likely to abuse their child, 93.3% of maltreated parents did not, indicating the importance of protective factors in the risk factor interaction. Commonly cited protective factors include the 'resilience' of a child (such as such as above-average intelligence, positive peer-relationships, calm temperament and internal locus of control) (CWIG, 2004), financial solvency, social support and two-parent families (Dixon, Browne & Hamilton-Giachritsis, 2009). A noteworthy finding from this study was that the absence of protective factors was more important than the presence of risk factors.

A limitation of the CM literature is that a 'higher order' model has not been developed to explore the pathway to CM perpetration. Rather models discussed above present a list of independent correlates. Although understanding risk factors independently is an important step, in order to be able to identify and support parents at risk of CM perpetration, it is important to know why risk and protective factors play a role. For example, what is it about substance misuse that increases risk of CM perpetration? Is it a primary risk factor (e.g. once a parent starts taking substances they become a risk for reasons such as availability to their child is limited and their behaviour is dysregulated) or is it a secondary risk factor (e.g. a parent already at risk of perpetrating CM is likely to misuse substances as a coping mechanism).

There are higher-order models available in the CM literature, for example looking at why maltreated children are at risk of later health complication (ACE model, Felitti et al., 1998). In its most basic form, the model suggests that adverse childhood experiences lead to social, emotional and cognitive impairment, which leads to adoption of health–risk behaviour and later health complications. The CM perpetration literature would similarly benefit from looking at the processes that occur in the pathway towards CM perpetration and desistance. For example, it is interesting that the most commonly cited risk and protective factors all relate to the extent to which a parent is able to prioritise their child's needs over their own and how well a parent is able to regulate their own emotions. Mental health difficulties, experience of violence and substance misuse are all obstacles to a parent's ability to regulate their own emotions and respond to their child's needs (Cleaver, Unell & Aldgate, 2011).

Understanding these risk factors from the ACE model perspective, being a victim of child maltreatment is an 'adverse childhood experience' (stage 1), that may lead to emotional regulation and mental health difficulties (stage 2: social, emotional and cognitive impairment) which then leads to substance misuse (stage 3: health-risk behaviours) and later violence towards family. This is merely a simplistic proposal of how a higher-order model of child maltreatment perpetration may look. In reality, the interaction of factors is complex but certainly is something that must be explored and depicted.

Empathy and Child Maltreatment

Although there is not a clear understanding of how empathy relates to CM, one can formulate hypotheses from other areas of research. For example, literature regarding attachment (Crittenden, 2006) indicates that attunement with a child is a key determinant of the development of a secure parent-child attachment. In order to

be attuned to their child a parent requires an understanding of how their child feels, what their child is experiencing and the appropriate way in which to respond. All of these processes are key stages for the development of cognitive and affective empathy (Marshall & Marshall, 2011) and thus it may be hypothesised that empathic capacity impacts upon the development of attachment, which impacts upon the likelihood of perpetrating CM. Certainly this is the assumption adopted in current practice.

Research also indicates that the relationship with an individual is a key determinant of empathic capacity. Having a negative relationship with somebody is likely to present an obstacle to experiencing affective empathy (Marshall & Marshall, 2011). To date, there is less of an understanding regarding how this applies to CM. Do parents who maltreat their children have general empathic deficits, or a deficit specifically in their ability to empathise with their own child?

With evident gaps in our knowledge regarding the role of empathy in the perpetration and perpetuation of CM, the present literature review aimed to understand the relationship between empathy and CM.

Current Review

With the aim of clarifying whether prior literature reviews have examined the relationship between parental empathy and CM perpetration, a preliminary search was conducted on 25th March 2015. The databases searched were: Cochrane Database of Systematic Reviews, PsycArticles, PsychINFO, Medline and The Campbell Library of Systematic Reviews. No relevant systematic reviews were found that looked at empathy in maltreating parents, highlighting the importance of the current review.

One literature review examined the relationship between empathy and aggression (Miller & Eisenberg, 1988). Considering relevance to the current chapter,

this review included four papers that examined the relationship between CM, empathy and aggression and briefly explored this relationship.

Method

Search Strategy

Prior to conducting searches for the review, searches of the databases were conducted to identify the extent of available data and inform potential search parameters. The preliminary search revealed that many studies examining empathy were conducted in the 1980's, thus informing the decision not to include a parameter of date to the search. If required by the limits of databases, the earliest date was chosen.

A search of electronic databases was undertaken on 18th April 2015. The databases searched were: OVID PsychINFO (1806 – April Week 3 2015), PubMed, Web of Science, Cochrane Library, Science Direct, Medline, ProQuest Applied Social Science Index & Abstracts (1987 – current), ProQuest Nursing & Allied Health Source , ProQuest Social Services Abstract (1979 – current) & ProQuest Sociological Abstracts (1952 - current). There were no limits applied to the searches at the stage of electronic search, to enable the search to be more encompassing. Google was also searched, although this did not reveal any papers that had not already been identified by the searches of databases.

The initial scoping exercise informed the development of search terms utilised in the search strategy (see below) by examining key words that were used for relevant articles. Furthermore, PubMed provides a list of all the words searched for when using wildcards which further informed the word variations of the search. When available, terms were mapped to subject headings in order to account for coding

differences between databases. Wild cards were used to broaden the search and the term 'perspective taking' was included, as this is a commonly investigated empathy construct. The search terms were as below:

(empath*) OR (perspective taking)

AND

(child* abus*) OR (emotion* abus*) OR (physical* abus*) OR (sex* abus*) OR (child* neglect*) OR (abus*) OR (neglect*)

AND

(parent*) OR (father*) OR (dad*) OR (mother*) OR (mom*) OR (Mum*) OR (carer*) OR (caregiver*)

A list of the terms included for each database search is provided in Appendix 2.

The search of electronic databases yielded 10595 articles. The titles and abstracts of these articles were read by the researcher in order to remove those that were not relevant to the review question. This process removed 10453 articles. Of the remaining 142 articles, 93 duplicates were removed. The full texts of the remaining articles were read to search for relevant articles that had not been identified via the electronic search of databases. This search identified 8 relevant articles, one of which could not be accessed. Where articles could not be accessed online, authors were contacted for full text copies (Appendix 3). Authors were also contacted to query whether they were aware of any relevant unpublished studies. Those who replied did not recommend any articles that had not been identified in the search.

The 57 full-text articles were assessed with relation to the PICO inclusion/exclusion criteria (Table 1). These criteria were based upon earlier literature scoping exercises and the aims of the research question. 39 articles were removed as

they did not meet the inclusion/exclusion criteria. Subsequent to quality assessment (discussed further below) 1 article was removed, leaving 17 articles remaining for inclusion. A detailed overview of this search process is provided in Figure 1. Appendices D and E respectively detail the articles that could not be accessed and those excluded when applying the PICO criteria.

Table 1

PICO Inclusion / Exclusion Criteria

	Inclusion	Exclusion
Population	• Parents or caregivers who have maltreated their children (or at high-risk of child maltreatment) • Males and females • All nationalities • All ethnicities	• 'antisocial' or 'aggressive' parents or caregivers who have not maltreated their own children.
Intervention / Exposure	N/A	N/A
Comparator	• Parents or caregivers who have not maltreated their children (or at low-risk of child maltreatment)	
Outcome	• Level of empathy, utilising an empathy measure that had standardized norms, has high validity and reliability.	• Level of empathy measured utilising an empathy measure that does not have data available for its validity or reliability, and does not have standardized norms
Study Design	• Cohort • Case study • Case series • Case control	• Narratives, Reviews, Commentaries, Editorials • Unpublished theses/dissertations
Other Factors	• Language of publication: English	• Published in a language other than English

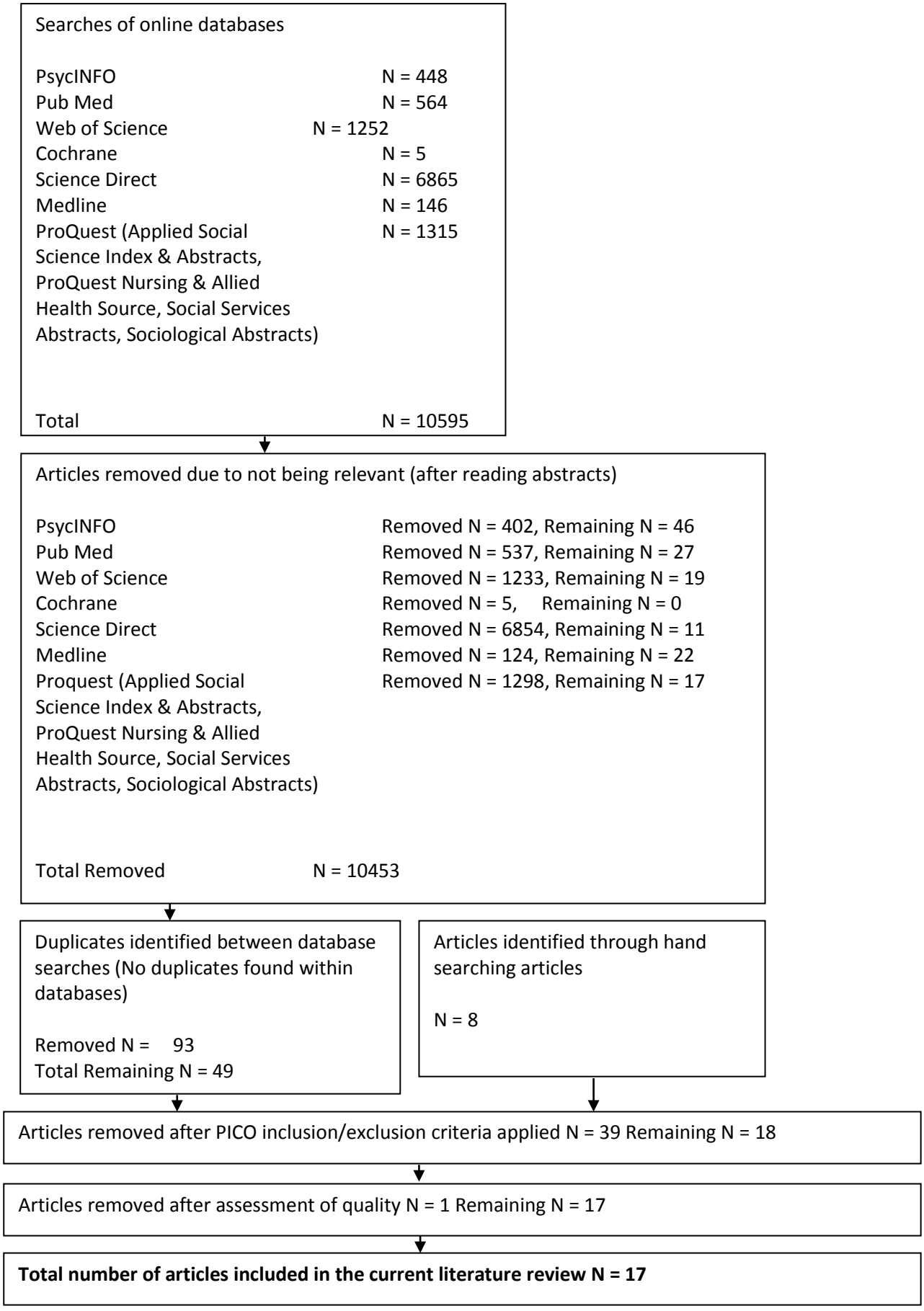


Figure 1. Study Selection Process

Quality Assessment

Studies deemed suitable via the inclusion/exclusion criteria were each assessed for their methodological quality. The assessment tool was informed by the Critical Appraisal Skills Programme (CASP, 2004) and the Quality Assessment Tool for Quantative Studies (Effective Public Health Practice Project; EPHPP, 1998). It was modified by the researcher to make it relevant to the current literature review. Although both tools include similar questions, the scoring methodology differs. CASP (2004) scores each question on a three-point scale of 'yes (2)' 'partially (1)' and 'no (0)', and an overall score is given which can be converted into a percentage. EPHPP (1998) similarly uses a classification of 'strong' 'moderate' or 'weak', although numerical equivalents are not given. This score has eight sub-sections assessing different aspects of a study, thus the tool produces eight sub-section classifications which determine a global rating of 'strong', 'moderate' or 'weak'.

It is commonly found in quality assessments that different tools yield different results (Juni, Witschi, Bolsch & Egger,1999). For the current review of papers the EPHPP (1998) adapted tool did not effectively capture the quality of papers, due to the global score being derived from the sub-section classifications. Rating papers in such a way, it was found that although some questions of a subsection were marked as 'weak' the overall subsection was rated as 'moderate', thus diluting the study flaws. Furthermore, there were an unequal amount of questions in each subsection meaning that the weightings of each question determining the subsection score differed. It was thus decided that the quality of papers would be measured utilising the modified CASP tool (Appendix 6).

The 18 assessed papers were either case-control or case-series studies, requiring the development of two separate tools. The tools were similar, with the case-series tool omitting questions relating to control groups.

Quality assessment items considered the following factors:

- Maltreating Parents: Maltreating participants were classified as being representative if the recruitment process reduced participant bias. These items focused on whether the sample was an opportunistic sample and whether sample sizes were adequate. The sample size cut-offs were based on a power analysis of the differences on the perspective-taking scale of the IRI found by Beven, O'Brien-Malone and Hall (2004), which suggested a total participant group of 126 would be needed with a target and control group. This would give 80% power to detect a difference of eight points on the Interpersonal Reactivity Index (IRI, Davis, 1980) perspective-taking scale, at the conventional 5% level of significance. This was intended to be an estimate of participant sizes needed. For case-control studies, it was also considered whether the control group was an appropriate comparator for the target population.
- Empathy: This item measured whether a valid and reliable measure was used to measure empathy.
- Confounding factors: Case control studies may be considered as more vulnerable to bias, due to the absence of random-assignment to groups (Sackett, 1979). This can be controlled for by matching-participants. This question focuses on whether the study explicitly matched participants pre-

assessment, or whether participant demographic variables were accounted for in analysis.

- Analysis: This item assessed the quality and appropriateness of analysis, and whether the analysis was adequate. In particular this item focuses on whether the analysis was limited to correlational analysis, or whether it extended to bivariate or multivariate analysis.

The case-control and case-series assessments contained a different number of items, thus scores were converted into percentages for the purpose of quality comparison. There is little information available as to what is considered an acceptable cut-off for inclusion. However, using information collated by Kmet, Lee and Cook (2004) a liberal cut-off of 60% was applied. A low cut-off was applied to make the review more encompassing and 60% was chosen over 50% to ensure that at least one item had a score of 2. Quality assessment results are provided in Appendix 7.

The inter-rater reliability of the quality assessment tool was assessed by providing 10 randomly selected papers (56% of total number of papers) to be assessed by an external researcher. There was agreement in ratings for all of the papers, with a maximum disparity of two points between raters. This score difference did not change the papers excluded from the review. The question that caused disparity was whether control groups were appropriate. The external rater scored this with relation to whether maltreatment status was adequately defined and there were enough participants, and the current author placing less emphasis on these factors as they were assessed in a separate question.

Using a cut-off of 60%, one study was excluded (Robyn & Fremouw, 1996). This study yielded a quality assessment score of 50%. The study was a case-control design with only nine participants in each group (maltreating and non-maltreating parents). Authors described that participants were 'controlled as closely as possible' however there is no detail given of this matching process or the demographic differences between participants (or groups), except for there being significant difference between groups of age and marital status. There is very little detail given in any of the sections of the article. The study only utilised t-tests to measure the group differences, although these were not reported descriptively.

Data Extraction

The data extraction proforma (Appendix 8) was developed by the researcher to facilitate data synthesis and capture information relevant to the research question. As well as article reference, this process extracted information relating to sample demographics (parenting status, sample size, age of sample size), how parents were recruited, how maltreating status was determined, the type of maltreatment assessed, how empathy was measured, the level of analysis used, study findings and the strengths and limitations of the study.

Results

Descriptive Data Analysis

Subsequent to the study selection process, 17 articles were included in the review (Appendix 9). Tables 2, 3, and 4 condense study characteristics, which will be discussed descriptively throughout the review.

Table 2
Characteristics of studies included in the review

Reference	Study Design	Quality Assessment Score	Participants	Comparators	How was 'child abuse status' determined	Measure of empathy and outcome relating to empathy and child maltreatment	Strengths and Weaknesses
de Paul, Perez-Albeniz, Guibert, Asla & Ormaechea (2008) Spain/ University of Basque Country	Case Control	79.17%	Neglectful mothers N=37 Mean age: 33.83 Recruited from nine child protection services in Spain.	1. Mothers at high risk for child physical abuse N=22 Mean age: 40.58 Recruited from five public schools of the Basque Country. 2. Non-maltreating mothers N=37 Recruited from 17 public schools in low-income areas of	Neglectful mothers were categorised based on information from child protection services. 'High risk' mothers scored above 32 on the abuse scale of the Child Abuse Potential Inventory. Non-maltreating mothers were "well known not to be neglectful or physically abusive".	The IRI was used as a measure of empathy A MANOVA was conducted with one between-subjects factor (parenting classification) with the three dependent variables from the IRI factors (personal distress and perspective taking). A significant main effect was found, Wilks's λ = .813, $F(6, 180) = 3.28$, $p = 0.004$. One way ANOVAs were conducted for each IRI factor. There was a significant difference between groups for personal distress, $F(2, 92) = 6.43$, $p = 0.002$, and perspective taking, $F(2,92) = 3.25$, $p = 0.04$ 'High risk' mothers had a higher score on personal distress and lower scores on the perspective taking than	There was no justification for comparing mothers who were proven to be neglectful with mothers at high-risk of physical maltreatment. Non-maltreating mothers did not complete the CAPI. The study would have benefitted from all participants completing the CAPI, to ensure the accuracy of group classification.

				the Basque country.		neglectful and non-neglectful mothers. No significant difference between neglectful mothers and comparison mothers	
Francis & Wolfe (2008) Canada/ McMaster University	Case Control	83.33%	Physically abusive fathers N = 23 Mean age: 38.44 Referred from Child Protection Services Child sexual abusers were excluded	Non-abusive fathers N = 25 Mean age: 39.20 Recruited from the same community as abusive fathers.	Physically abusive fathers had a record with Child Protective Services as being physically abusive to a biological or step-child under the age of 12 within the past three years. Non-abusive fathers were chosen from the same community. Scores on the CAPI, Childhood Trauma Questionnaire-Short Form and the CAGE	The IRI and IFEEL measures were used to assess empathy. IRI: Abusive fathers has lower empathic concern ($F(1, 45) = 5.27$, $p < .05$) and perspective taking ($F(1,45) = 5.84$, $p < .05$) scores. Performance on the IFEEL task demonstrated that abusive fathers gave more responses falling in the Anger, Fear and Other categories and fewer falling in the interest category.	Described the recruitment process in detail, including difficulties with recruitment and drop-outs.

					alcohol dependence screening questionnaire were also used to screen parents pre-study.		
Letourneau (1981) Washington / National Catholic School of Social Service	Case Control	87.5%	32 mothers identified as physically abusive by Child Protection Services Mean age: 24.6	30 non-abusive mothers Mean age: 23.9 Recruited from a day care centre and public health programme.	Physically abusive mothers were categorised accordingly due to their involvement with Child Protection Services. No detail is given regarding how non-abusive mothers were determined to be non-abusive.	The HES and EQ were used to measure empathy. There was a significant difference in empathy scores for abusive and non-abusive mothers ($t = -6.77$, $p < .0001$) on HES. There was a significant difference between abusive and non-abusive mothers ($t = -2.74$, $p < 0.004$) on the EQS.	There is no information given regarding how non-abusive mothers were classified as not abusive.
Leon, Rodrigo, Quinones, Hernandez, Lage & Padron (2014)	Case Control	66.67%	Neglectful mothers attending a preventative parenting programme due to having low	Non-neglectful mothers attending a preventative parenting programme due to having low socioeconomic	Neglectful mothers identified by the Maltreatment Classification System and had a history of maltreatment	The IRI was used a measure of empathy. Neglectful mothers scored significantly lower on the IRI empathic concern subscale ($t(27) = 2.35$, $p<0.05$).	The sample were not representative of the general parenting population.

Spain/ University of La Laguna & Basque Center on Cognition, Brain and Language			socioeconomic status N = 28 Mean age: 33.6	status N=30 Mean age: 36.6	Non-neglectful mothers were chosen as they did not have a history of maltreatment with social services.		
McElroy & Rodriguez (2008) University of North Carolina	Case Series	88.89%	Mothers of 5-12 year old children who had externalizing behaviour problems. N = 73 Mean age: 40	N/A	The study utilised the Parenting scale, CAPI and parent-child CTS scores as a measure of child abuse potential.	The IRI was used as a measure of empathy. There was a significant negative correlation between the perspective taking subscale of the IRI and score on the CAPI ($r(73) = -.48$, $p < .01$) The CAPI abuse scale scores in the whole sample were significantly higher than the reported normative mean of 91 ($p < 0.001$).	The analysis was correlational and no record of how many parents score high- and low- risk for maltreatment
Mennen & Trickett (2011) University of Southern California	Case Control	94.44%	Mothers who had abused their children N = 83 Mean age: 35.9	1. Non-maltreating mothers N = 100 Mean age: 32 2. Foster mothers	Maltreating mother status was determined utilising records from Child Protection Services.	The AAPI 'lack of empathy' scale was used to measure empathy. There was a significant relationship between caregiver relationship and lack of empathy on the AAPI. Using an ANCOVA, comparison mothers scored significantly higher on the lack of empathy subscale than foster	There is no description of how (or whether) non-maltreating mothers, foster mothers and female relative were determined to be non-

			Recruited from Child Protection Services.	N = 100 Mean age: 49 3. Female relatives N = 52 Mean age: 44		mothers and maltreating mothers. (F (3, 282) = 3.51, $p < 0.05$). Statistics regarding post-hoc analyses were not provided, except for denoting the relationship. There was a significant correlation between lack of empathy and belief in corporal punishment score (r = 0.30, $p < 0.01$)	maltreating (or low-risk of maltreating).
Milner, Halsey & Fultz (1995) Northern Illinois University	Case Control	83.33%	Parents at high risk of child maltreatment. N=10 Mean age: 33 Recruited from a social service agency, public school, a medical clinical and a	Parents at low risk of child maltreatment. N=10 Mean age: 34 Matched on race, age, educational level, marital status and number of children. Recruited from a	High risk parents were classified as scoring above 166 on the CAPI. (Mean CAPI score of 254.3) Low risk parents were classified as scoring below 166 on the CAPI. (Mean CAPI score of 74.2).	The IRI and a novel task whereby mothers watched videos of infants were used to measure empathy IRI: High-risk mothers showed significantly higher levels of personal distress compared to low-risk mothers, F (1,18) = 12.90, $p < .005$. No significant difference ($p > .05$) for perspective taking or empathic concern between high-risk and low-risk mothers were found. Video task: No significant difference between conditions. However there was a significant increase in empathy between baseline and crying infant	Participants were matched exactly on gender, ethnic background and marital status.

			public aid office.	social service agency, public school, a medical clinical and a public aid office.		for low-risk mothers, F(3,27) = 11.83, $p < 0.0005$. This was not significant for high-risk mothers, F(3,27) = 2.91, p = 0.53	
Perez-Albeniz & de Paul (2003) Spain/ University of Basque Country	Case Control	91.76%	High-risk parents for child maltreatment 11 Male 25 Female N = 36 Mean age: 40.3 Recruited from seven public schools	Low-risk parents for child maltreatment N = 38 16 Male 22 Female Mean age: 39.3 Recruited from seven public schools	High risk parents scored above 32 on the Abuse Scale of the CAPI. Low risk parents scored below 6 on the Abuse Scale of the CAPI	The IRI, HES and QMEE were used to measure empathy. A MANOVA revealed a significant difference between high- and low-risk parents on all measures of empathy. (Wilks $\lambda = .283$, $F(12,59) = 12.430$, $p < .001$). There was a significant difference between high- and low- risk parents on the HES total score, $F(1,10) = 40.82$, $p < .001$, and for the QMEE total score, $F(1,71) = 5.25$, $p < .05$. High-risk parents showed significantly lower scores on the IRI 'empathic concern' dimension and significantly higher scores on the 'personal distress' dimension compared with low-risk parents.	Utilised multiple tests of empathy. Relatively good sample size, with relation to other studies. Did not split the analysis into male and female.
Perez-Albeniz &	Case Control	89 %	High risk parents.	Low risk parents	High risk parents scored above 32	The IRI and PPES were used to measure empathy.	There were no significant

de Paul (2004) Spain/ University of Basque Country			N=19 Mean age: 37.4 9 fathers 10 mothers Four public schools	N = 26 Mean age: 37.2 12 fathers 14 mothers Four public schools	on the Abuse Scale of the CAPI. Low risk parents scored below 6 on the Abuse Scale of the CAPI	A significant main effect for risk status was found, Wilks λ = .336; F(7,35) = 9.89 $p < 0.001$. IRI: High-risk parents showed a significantly higher score on the personal distress dimension F(1, 41) = 39.05 $p < .001$ and lower score on perspective taking dimension f(1,41) = 11.51 p = .002 compared with low-risk parents. There was no significant difference on empathic concern dimension. PPES: High-risk parents showed a significantly lower total score, F (,410 = 21.70 $p < .001$ and lower score on empathy with their partner, F (1,41) = 23.22 $p < .001$ and child, f (1,41) = 20.15 $p < .001$ compared with low-risk parents. There was no effect of gender on either the IRI or PPES.	differences between both groups on age, number of children, gender, marital status and educational level. Low response rate (of 1, 514 questionnaires sent out, 331 were returned).
Rodriguez (2013) University of North	Case Series	70%	Mother-child dyads. Children aged between 4 – 9 years.	Those in the bottom 5th of the distribution of the CAPI	High risk parents were those who scored above the 166 on the CAPI.	The IRI, plotkin child vignettes (PCV) and MATCh task were used to measure empathy. Participants completed the AAPI-2 but the scores relating to the empathy scale were	No information given regarding how participants were recruited.

Carolina			N=20 Recruited from two separate parenting studies. Mean total age of group: 36.83	N=26 Recruited from two separate parenting studies Mean total age of group: 36.83	Low risk parents were those who scored in the bottom fifth of the CAPI score distribution, in the overall sample of 135 mothers.	not reported. There was not a significant relationship between CAPI scores and self-reported empathy. There was not a significant difference between high risk and low risk mothers on the IRI empathic concern ($t(48) = 0.11$, $p > 0.5$) or perspective taking ($t(48) = 0.68$, $p > 0.05$) On the MATCh task, there was a significant correlation between total score and CAPI score ($t(48) = 1.99$, $p < .05$). Thus high-risk mothers were less able to read their child's emotion responses than low-risk mothers. Furthermore, a significant correlation between MATCh score and IRI empathic concern ($r(46) = .33$, $p < .01$) and perspective-taking ($r(46) = .25$, $p < .01$)	No information given regarding the demographics of those classified as high- and low-risk of maltreatment.
Rodriguez, Cook & Jedrziewski (2012) University	Case Series	75%	Parents N = 26 19 mothers 7 fathers	N/A	Parents completed the CAPI and the APPI	PCV and the EQS were used to measure empathy. The AAPI-2 was administered with participants but the empathy scale results were not reported.	Did not report the relationship between AAPI scores and CAPI score.

of North Carolina			Average age: 30.85 Recruited from after school programmes			There was a significant correlation between the EQS score and CAPI abuse score (r (26) = 0.61, $p < 0.01$), the EQS score and AAPI-2 total score (r (26) = -0.49, $p < 0.01$) and the EQS and PCV punishment score (r (26) = 0.50, $p < 0.01$). There was no significant relationship between CAPI score and eye-tracking scores for attribution (r (26) = .00, $p > .05$) and empathy vignettes (r(26) = .26, $p > 0.05$). There was a significant correlation between AAPI total score and the total time taken to read the empathy vignettes (r(26) = .35, $p < .05$).	Only used correlational analysis. Did not report how many parents were at high- and low-risk of child maltreatment.
Rodriguez & Richardson (2007) University of North Carolina	Case Series	75%	Parents N = 115 86 mothers 29 fathers Mean age: 37.62 Recruited from primary	N/A	The CAPI and Parent-Child conflict tactics scale was administered with parents. However there is no information in the articles detailing how many parents reached the cut-off's for being	The IRI was used as a measure of empathy. There was a significant negative correlation between score on the CAPI Abuse Scale and the IRI perspective taking scale (r(115) = -.25, $p < .01$) and a positive correlation between the CAPI abuse scale and the parenting stress index (r(115) = .70, $p < .001$).	There was no information given regarding how many participants fell into high- and low- risk of maltreatment With regards to empathy, only correlational

			and secondary schools.		high-risk of maltreatment.	IRI was not included in the final hierarchical multiple regression predicting CAPI abuse Scale score.	information available.
Rodriguez & Tucker (2014) University of North Carolina University of Alabama	Case Series	81.25%	Mother-child dyads N = 95 Mean age: 37.89 Recruited from a day care and after school centres	N/A	Mothers completed the CAPI and AAPI-2	The IRI and EQS were used to measure empathy. The AAPI-2 was used in the study, but results from the empathy scale were not reported. Utilising a regression model, found that lower empathy and more negative child attributions accounted for some of the variance in increased risk of abuse. Those with high distress but low social support had the highest abuse risk scores (M=1.22), followed by those with low distress and low social support (M = -0.07), followed by high distress and support (M=-0.27)	The study uses the IRI, EQ and AAPI as measures of empathy, however only reports one global empathy score. There is no breakdown providing with regards to results on separate subscales. Participants were recruited from a range of community services, increasing representation of the sample
Rosenstein (1995)	Case Series	85%	Caretakers referred to child	The study split parents into four groups:	All participants had been referred to child protection	The AAPI empathy scale was used to measure empathy.	The study utilised participants from

Corpus Christi			protection services. N = 29 20 Females 9 Males 26 parents 2 grandparents 1 unrelated carer Mean age: 32	1. first referral and no abuse found 2. First referral and abuse found 3. More than one referral but no abuse found 4. More than one referral and abuse found.	services for child maltreatment. Two participants were removed as the maltreatment was considered 'minor'. The author decided whether the abuse was 'valid' based on the definition of child physical abuse in the Texas Family Code (1993).	There was not a significant empathy difference between abusive and non-abusive parents. There was a significant correlation between parental levels of empathy and whether the parent values corporal punishment (r (29) = -.753, $p < .0001$). For 75% of parents, there was a negative correlation between empathy levels and parent-child stress.	another study to conduct ANOVAS. Split participants into four groups; 29 participants is a small sample for this split.
Thompson, Jones, Litrownik, English, Kotch, Lewis & Dubowitz (2014) Calica Center for Innovation in Children	Case Series	75%	Mother-child dyads recruited from the LONGSCAN longitudinal sample. N=412 No mean age reported	N/A	Maltreatment status was determined by CPS records and a questionnaire provided to children	The AAPI empathy scale was used as a measure of empathy Significant correlation between mothers empathy (measured using the AAPI) and emotional maltreatment (r=0.18, $p < 0.05$). Correlation between mother's empathy and physical abuse was non-significant.	Controlled for demographic variables Empathy was not the main focus of the study.

and Family Services, Juvenile Protective Association			Recruited from a number of sites for children at risk of maltreatment				
Wiehe (1986) University of Kentucky	Case Control	66.67%	Abusive mothers N=32 Mean age: 25	Mothers who accessed a community centre who were not known to be abusive. N = 32 Mean age: 26	Abusive mothers were chosen randomly from a child protection case load. Non-maltreating mothers "were never known to the community centre staff as having been reported for child abuse"	The HES was used as a measure of empathy A one-way ANOVA was used to analyse outcomes from the HES identified that abusive mothers had significantly lower empathy scores than non-abusive mothers $F(1, 62) = 10.62$, $p < 0.01$. When marital status, income and education were controlled for, education showed a significant interaction $F(2,57) = 6.89$, $p < 0.01$.	Abuse status was assumed rather than measured Sample were from a community centre and child protection case load, thus limited in ability to apply findings to other populations
Wiehe (2003) University of Kentucky	Case Control	70.83%	Physically and emotionally abusive parents N = 52	Non abusive foster parents N = 101 16 males 85 females	Abusive parents were selected from child protective services, and were being investigated for child abuse	The IRI was used as a measure of empathy Abusive parents showed significantly lower scores on perspective taking ($t(153) = -3.66$, $p < .05$) and empathic concern ($t(153) = -3.41$, p	The control group were assumed to be non-abusive.

	10 males 42 females Average age: 34	Average age: 42	Foster parents were assumed to be non-abusive.	< .05) and higher scores on personal distress (t(153) = 3.75, *p* <.05)

Note. CAPI = Child Abuse Potential Inventory; IRI = Interpersonal Reactivity Index; IFEEL = IFEEL task based assessment of empathy; HES = Hogan's Empathy Scale; EQ = Empathy quotient; CTS = Conflict Tactics Scale; AAPI = Adult-Adolescent Parenting Inventory; QMEE = The Questionnaire Measure of Emotional Empathy ; PPES = Parent/Partner Empathy Scale; MATCh = Matching Affect to Child Task; PCV = Plotkin Child Vignettes

Table 3

Participant selection and empathy measures used in studies

	Recruitment		Parents		Defining Maltreatment			Measure of Empathy							
	Schools/ nursery/ day care centres	Child Protection Services	Mothers	Fathers	Used parents proven have maltreated	Only used CAPI to define maltreatment status	Information not given of how controls were defined	IRI	AAPI	HES	PPES	EQS	QMEE	PCV	Other
de Paul et al., (2008)	X	X	X		X		X	X							
Francis & Wolfe (2008)		X		X	X			X							IFEEL pictures
Letourneau (1981)	X	X	X		X		X			X		X			
Leon et al., (2014)		X	X		X			X							
McElroy & Rodriguez (2008)			X			X		X							
Mennen & Trickett (2011)		X	X		X		X		X						
Milner et al., (1995)	X	X	X			X		X							Videotape of Infants
Perez-Albeniz & de Paul (2003)	X		X	X		X		X		X			X		
Perez-Albeniz & de Paul (2004)	X		X	X		X		X			X				
Rodriguez (2013)	X		X			X		X	X					X	MATCh
Rodriguez et al., (2012)	X		X	X		X			X			X		X	Eye tracking
Rodriguez & Richardson (2007)	X		X	X		X		X							
Rodriguez & Tucker (2014)	X		X			X		X	X			X		X	
Rosenstein (1995)		X	X	X	X				X						

Thompson et al., (2014)			X			X			X			
Wiehe (1987)	X		X			X			NO			X
Wiehe (2003)			X			X			NO		X	

Note. CAPI = Child Abuse Potential Inventory; IRI = Interpersonal Reactivity Index; IFEEL = IFEEL task based assessment of empathy; HES = Hogan's Empathy Scale; EQ = Empathy quotient; CTS = Conflict Tactics Scale; AAPI = Adult-Adolescent Parenting Inventory; QMEE = The Questionnaire Measure of Emotional Empathy ; PPES = Parent/Partner Empathy Scale; MATCh = Matching Affect to Child Task; PCV = Plotkin Child Vignettes

Table 4
Demographic variables that studies considered

	Marital status	Education	Parent age	# of children	Children's age	Nationality / race / ethnicity	Employed	Household income	Offence History	Alcohol abuse	Social Class	Family Structure	Maltreatment as child	Total number of variables considered
de Paul et al., (2008)	X	X	X	X										4
Francis & Wolfe (2008)	X	X	X	X		X	X	X	X	X			X	10
Letourneau (1981)	X	X	X	X		X		X			X	X		8
Leon et al., (2014)	X	X	X	X	X		X							6
McElroy & Rodriguez (2008)		X	X	X		X								4
Mennen & Trickett (2011)		X	X			X								3
Milner, Halsey & Fultz (1995)	X	X	X	X	X	X								6
Perez-Albeniz & de Paul (2003)	X	X	X	X										4
Perez-Albeniz & de Paul (2004)	X	X	X	X										4
Rodriguez (2013)			X		X	X	X	X						5
Rodriguez et al., (2012)		X	X	X		X								4
Rodriguez & Richardson (2007)	X	X	X	X				X						5

Rodriguez & Tucker (2014)	X	X	X	X	X	X	6
Rosenstein (1995)			X	X	X		3
Thompson et al., (2014)	X				X	X	3
Wiehe (1986)	X	X	X		X	X	5
Wiehe (2003)		X			X		2

Results

The 17 reviewed studies can be considered in three categories: studies looking solely at empathy (de Paul et al., 2008; Milner et al., 1995; Perez-Albinez & de Paul, 2003, Rodriguez et al, 2012; Rosenstein, 1995), studies looking at empathy alongside another construct (Letourneau, 1981; Perez-Albinez & de Paul, 2004; Rodriguez, 2013; Wiehe, 1986, 2003) and studies examining broader concepts which included a measure of empathy (Francis & Wolfe, 2008; Leon et al., 2014; McElroy & Rodriguez, 2008; Mennen & Trickett, 2011; Rodriguez & Richardson, 2007; Rodriguez & Tucker, 2014, Thompson et al., 2004). All of the studies tested participants once, none conducted follow-up assessments. Seven of the studies used a case-series design and ten used a case-control design.

Quality of Studies

The quality assessments revealed disparity between the 17 studies regarding quality. As described in Table 1, six studies scored between 85-94% (Letourneau, 1981; McElroy & Rodriguez, 2008; Mennen & Trickett, 2011; Perez-Albinez & de Paul, 2003; 2004; Rosenstein, 1995), seven studies scored between 75-84% (de Paul, et al., 2008; Francis & Wolfe, 2008; Milner et al., 1995; Rodriguez et al., 2012; Rodriguez & Richardson, 2007; Rodriguez & Tucker, 2014), and four studies scored between 65-74% (Leon et al., 2014; Rodriguez, 2013; Wiehe 1986; 2003). The key weaknesses of studies were a) small sample sizes, b) participants being recruited from specific services (such as Child Protection Services) and therefore not representative of the general parent population and c) studies not accounting for confounding factors between groups. A further weakness in some studies was that control parent groups were not assessed for maltreatment risk, relying upon researcher assumption. Furthermore, some of the studies utilised correlational analysis and did not give detail

as to how many parents scored high- and low- risk for CM. These studies failed to give enough information as to whether there were significant empathy differences between parents at high- and low- risk of CM.

Study Sites

Nine studies were conducted at individual institutions across America and Canada, with the remaining eight studies conducted by the University of Basque Country (De Paul et al., 2008; Perez-Albinez & de Paul, 2003; Perez-Albinez & de Paul, 2004) and the University of North Carolina (McElroy & Rodriguez, 2008; Rodriguez, 2013; Rodriguez & Richardson, 2007; Rodriguez et al., 2012; Rodriguez & Tucker, 2014). Thus 47% of the studies in the current review were conducted by two institutions. All eight of these studies relate their findings to the social information processing model (Milner, 1993; 2003), suggesting homogeneity of perspective. There are also concerns as Rodriguez (2013) stated that their participant sample came from two other studies, although these studies were not named. As four (23.5%) of the studies in the review were undertaken by Rodriguez, it is a possibility that data from some participants has been represented twice in the review.

Participants

Seven studies, utilising case-series design, did not give information regarding how many participants were maltreating and non-maltreating. Of the ten studies that did, the average sample size for maltreating parents was 32 (SD = 19.82). The smallest sample included ten maltreating participants (Milner et al., 1995) and the largest maltreating sample was 83 (Mennen & Trickett, 2011). The average age for maltreating parents was 34.5 (SD = 4.2) and 35.1 (SD = 4.6) for non-maltreating parents. Most notable is that only four studies included an adequate number of

participants (Mennen & Trickett, 2011; Rodriguez, 2013; Thompson et al., 2014; Wiehe, 2003), as defined by the quality assessment.

The majority of studies, apart from Francis and Wolfe (2008), included mothers in their sample, with six studies also including fathers (Perez-Albinez & de Paul, 2003; 2004; Rodriguez & Richardson, 2007; Rodriguez et al., 2012; Rosenstein, 1995; Wiehe, 2003). Francis and Wolfe (2008) looked exclusively at fathers in their study. Fathers represented 21% of the overall sample in the six studies that included both mothers and fathers. Only one of these studies split analysis between male and females (Perez-Albinez & de Paul, 2004), finding that there was not a significant difference between the empathy of mothers and fathers. Other carer types in the studies were foster mothers (Mennen & Trickett, 2011), female relatives (Mennen & Trickett, 2011), grandparents and unrelated carers (Rosenstein, 1995).

Participant recruitment

With relation to participant recruitment, 10 studies recruited parents through Child Protection Services (De Paul et al., 2008; Francis & Wolfe, 2008; Leon et al., 2014; Letourneau, 1981; Mennen & Trickett, 2011; Milner et al., 1995; Rosenstein, 1995; Thompson et al., 2014; Wiehe, 1987; 2003) and 10 studies recruited from schools, day care centres and nurseries (De Paul et al., 2008; Letourneau, 1981; Milner et al., 1995; Perez-Albinez & de Paul, 2003; 2004; Rodriguez, 2013; Rodriguez et al., 2012; Rodriguez & Richardson, 2007; Rodriguez & Tucker, 2014; Wiehe, 1987). Out of 17 reviewed studies, only 9 (53%) utilised parents who were confirmed to have perpetrated CM (De Paul et al., 2008; Francis & Wolfe, 2008; Letourneau, 1981; Leon, et al., 2014; Mennen & Trickett, 2011; Rosenstein, 1995; Thompson et al., 2014; Wiehe, 2003). The remaining studies utilised the Child Abuse

Potential Inventory (CAPI; Milner, 1986) to classify parents as high- and low- risk of CM, although the consistency of using it and how it was used varied between studies.

Demographic variables

All of the studies considered at least some demographic variables in data collection, however only one study matched participants (Milner et al., 1995). A summary of demographic variables considered is provided in Table 3. The most commonly considered demographic variables were parent education, parent age, number of children and ethnicity, with 9 studies only considering between 2-4 variables (De Paul et al., 2008; McElroy & Rodriguez, 2008; Mennen & Trickett, 2011; Perez-Albinez & de Paul, 2004; Rodriguez et al., 2012; Rodriguez & Tucker, 2014; Rosenstein, 1995; Thompson et al., 2014; Wiehe, 2003). Although 14 studies recorded parents' education level, none of the studies considered more explicit tests of intellectual functioning, to prevent bias in participant groups. Furthermore, only one study (Wiehe, 2003) considered the impact of personality styles on empathy, but focused only on narcissism.

As already mentioned, one study matched participants pre-assessment (Milner et al., 1995). Of those with a case-control design, three found significant demographic differences between groups (de Paul et al., 2008; Mennen & Trickett, 2011; Wiehe, 2003), only one controlling for this in analysis (Mennen & Trickett, 2011). There were studies included in the review whereby the comparator group was not justified. This was true particularly for de Paul et al. (2008), who included mothers proven to be neglectful and mothers and high- and low- risk for physical abuse based on the CAPI (Milner, 1986). There was no justification in the article as to why mothers proven to be neglectful were compared with those at risk for physical abuse. Furthermore, five studies did not assess control participants for maltreatment status, assuming they

would be non-maltreating (de Paul et al., 2008; Letourneau, 1981; Mennen & Trickett, 2011; Wiehe, 1987; 2003).

Measuring empathy

As previously noted, the concept of empathy is complex, which necessitates studies to be explicit about the aspect of empathy they are examining and utilising a tool that accurately measures this construct. However, only two of the reviewed studies explicitly outlined the aspects of empathy the study was concerned with (de Paul et al., 2008; Perez-Albinez & de Paul, 2003), both looking at dispositional empathy which they describe to be the propensity to use the capacity to empathise. None of the 17 articles justified their choice of empathy assessment.

Of the 17 studies, eight measured general empathy exclusively (de Paul et al., 2008; Letourneau, 1981; Leon et al., 2014; McElroy & Rodriguez, 2008; Perez-Albeniz & de Paul, 2003; Rodriguez & Richardson, 2007; Wiehe, 1987; 2003) with an additional six including measures of general empathy alongside other empathy measures (Francis & Wolfe, 2008; Milner et al., 1995; Perez-Albeniz & de Paul, 2004; Rodriguez, 2013; Rodriguez, et al., 2012; Rodriguez & Tucker, 2014). Eight studies measured empathy towards general children (Francis & Wolfe, 2008; Mennen & Trickett, 2011; Milner, et al., 1995; Rodriguez, 2013; Rodriguez et al., 2012; Rodriguez & Tucker, 2014; Rosenstein, 1995; Thompson et al., 2014) and three assessed mother's empathy towards their own children (Francis & Wolfe, 2008; Perez-Albeniz & de Paul, 2004; Rodriguez, 2013). These results are displayed in Table 5.

Table 5
Type of empathy measured in studies

	General Empathy	Empathy towards children	Empathy towards own child
de Paul et al., (2008)	X		
Francis & Wolfe (2008)	X	X	X
Letourneau (1981)	X		
Leon et al., (2014)	X		
McElroy & Rodriguez (2008)	X		
Mennen & Trickett (2011)		X	
Milner, Halsey & Fultz (1995)	X	X	
Perez-Albeniz & de Paul (2003)	X		
Perez-Albeniz & de Paul (2004)	X		X
Rodriguez (2013)	X	X	X
Rodriguez et al., (2012)	X	X	
Rodriguez & Richardson (2007)	X		
Rodriguez & Tucker (2014)	X	X	
Rosenstein (1995)		X	
Thompson et al., (2014)		X	
Wiehe (1987)	X		
Wiehe (2003)	X		

Studies that used the IRI

The most commonly used psychometric assessment of empathy in the 17 studies was the IRI (Davis, 1980, 1983) self-report questionnaire, used in 11 studies (de Paul et al., 2008; Francis & Wolfe, 2008; Leon et al., 2014; McElroy & Rodriguez, 2008; Milner et al., 1995; Perez-Albinez & de Paul, 2003; 2004; Rodriguez, 2013; Rodriguez & Richardson, 2007; Rodriguez & Tucker, 2014; Wiehe, 2003). The IRI has four subscales, measuring perspective-taking, identification with fictional characters (fantasy), personal distress and empathic concern. Thus the perspective taking subscale may be considered a measure of cognitive empathy and the empathic concern scale a measure of affective empathy. However not all studies used all of the IRI scales (Table 5). Four studies utilised all of the IRI scales (de Paul et al., 2008; Leon et al., 2014; Perez-Albeniz & de Paul, 2003; 2004). All of the studies used the perspective taking scale and only two studies did not use the empathic concern scale (McElroy & Rodriguez, 2008; Rodriguez & Richardson, 2007). Six studies used the

personal distress scale (de Paul et al., 2008; Leon et al., 2014; Milner et al., 1995; Perez-Albeniz & de Paul, 2003; 2004; Wiehe, 2003) and only the four studies that used all of the scales used the fantasy scale.

Table 6
Scales of the IRI used in studies

	Perspective Taking	Fantasy	Empathic Concern	Personal Distress
de Paul et al., (2008)	X		X	X
Francis, K & Wolfe (2008)	X		X	
Leon et al., (2014)	X	X	X	X
McElroy & Rodriguez (2008)	X			
Milner, Halsey & Fultz (1995)	X		X	X
Perez-Albeniz & de Paul (2003)	X	X	X	X
Perez-Albeniz & de Paul (2004)	X	X	X	X
Rodriguez & Richardson (2007)	X			
Rodriguez (2013)	X		X	
Rodriguez & Tucker (2014)*	X		X	
Wiehe (2003)	X	X	X	X

Note: * Rodriguez & Tucker (2014) combined the perspective taking and empathic concern scores to make a global score and did not report the scores individually.

Other measures of empathy

Another commonly used measure was the Adult-Adolescent Parenting Inventory (AAPI; Bavolek, 1984; 2001), which was used in six studies (Mennen & Trickett, 2011; Rodriguez, 2013; Rodriguez et al., 2012; Rodriguez & Tucker, 2014, Rosenstein, 1995; Thompson et al., 2014). The AAPI is a self-report measure with four subscales assessing: expectations of the child, empathy for the child, value of physical punishment and parent-child role reversal. These measures are very different: the IRI assesses general empathy whilst the AAPI focuses on empathy towards children. However, only three of the six studies utilised the AAPI as a measure of empathy (Mennen & Trickett, 2011; Rosenstein, 1995; Thompson et al., 2014). The other three studies utilised the AAPI as part of their assessment of CM.

Other self-report measures of empathy used in the studies were the Hogan Empathy Scale (HES; Hogan, 1969) (Letourneau, 1981; Perez-Albinez & de Paul, 2003; Wiehe, 1987), The Questionnaire Measure of Emotional Empathy (QMEE; Mehrabian & Epstein, 1972) (Perez-Albinez & de Paul, 2003), Parent/Partner Empathy Scale (PPES; Feshbach & Caskey, 1985) (Perez-Albinez & de Paul, 2003), The Empathy Quotient-Short (EQ-S; Wakabayashi et al., 2006) (Letourneau, 1981; Rodriguez et al., 2012; Rodriguez & Tucker, 2014) and IFEEL pictures (Emde, Osofsky & Butterfield, 1994) (Francis & Wolfe, 2008). Novel behavioural assessments of empathy were used in three studies (Milner et al., 1995; Rodriguez, 2013; Rodriguez et al., 2012; Rodriguez & Tucker, 2014) as outlined in Table 3.

Outcomes related to empathy

A summary of study outcomes is provided in Table 7. Of the 17 included studies, three (18%) did not find a significant relationship between parenting maltreatment status and empathy (Milner et al., 1995; Rodriguez, 2013; Rosenstein, 1995). Milner et al., (1995) only found a significant difference on the IRI personal distress scale between high and low risk mothers. Non-significant results were found for the perspective taking and empathic concern scale, with Cohen's d effect sizes of 0.34 and 0.23 respectively. Similarly, Rosenstein (1995) had a small sample (n=29) which was split into four groups for analysis. Thus design limitations may have impacted upon their ability to find reliable results. Calculating effect sizes from these studies, Rosenstein (1995) obtained a Cohen's d score of 0.62, suggesting that although results did not reach significance, an effect of empathy was observed. It is also worth noting that although Rodriguez (2013) did not find a significant difference between IRI measures of empathy, the novel MATCh (Matching Affective to Child) task did find a significant effect (described below).

Table 7
Summary of outcomes related to empathy

	IRI PT	IRI FS	IRI EC	IRI PD	HES	EQS	AAPI	OTHER
de Paul et al., (2008)	Maltreating mothers had significantly lower scores		X	Maltreating mothers had significantly higher scores				
Francis & Wolfe (2008)	Maltreating fathers had significantly lower scores		Maltreating parents had significantly lower scores					IFEEL: Abusive fathers gave more responses falling in the Anger, Fear and Other categories and fewer falling in the interest category.
Letourneau (1981)					Maltreating mothers had significantly lower score	Maltreating mothers had significantly lower score		
Leon et al., (2014)	X	X	Maltreating mother had significantly lower scores	X				
McElroy & Rodriguez (2008)	Maltreating mothers significantly more likely							

	to have lower PT score					
Mennen & Trickett (2011)					Comparison mothers scored significantly higher on lack of empathy subscale (better empathic capacity) than foster mothers and maltreating mothers.	
Milner et al. (1995)	X	X	High risk mothers showed significantly higher levels of PD			Video task: No significant difference between conditions. However there was a significant increase in empathy between baseline and crying infant for low-risk mothers. This was not found for high risk mothers

Perez-Albeniz & de Paul (2003)	X	X	High risk showed significantly lower score	High risk showed significantly higher scores	High risk significantly lower score		QMEE: High risk significantly lower score
Perez-Albeniz & de Paul (2004)	High risk had significantly lower score	X	X	High risk had significantly higher score			PPES: High Risk had significantly lower score
Rodriguez (2013)	X	X	X	X			MATCh task: High risk mothers were significantly less able to read their child's emotion compared to low risk mothers
Rodriguez et al., (2012)						Lower There was a significant correlation between the EQS score and CAPI abuse score, AAPI-2 total score, and PCV punishment score.	PCV and eye tracking: There was no significant relationship between CAPI score and eye-tracking scores for attribution and empathy vignettes

Rodriguez & Richardson (2007)	High risk had significantly lower PT score					
Rodriguez & Tucker (2014)	The study used the IRI PT and EC scales and the EQS but only reported a global score. Utilising a regression model, found that lower empathy and more negative child attributions accounted for some of the variance in increased risk of abuse.					
Rosenstein (1995)						No significant difference in empathy between abusive and non-abusive parents.
Thompson et al. (2014)						Correlation between empathy and emotional maltreatment
Wiehe (1987)					Maltreating mothers had significantly lower scores	
Wiehe (2003)	Maltreating parents had	X	Maltreating mothers had	Maltreating mothers had		

	significantly lower PT scores	significantly lower EC scores	significantly higher PD scores

Note: X = no significant difference; IRI = Interpersonal Reactivity Index; PT = Perspective Taking Subscale; FS = Fantasy Subscale; EC = Empathic Concern Subscale; PD = Personal Distress subscale; HES = Hogan's Empathy Subscale; EQS = Empathy Quotient-Short; AAPI = Adult-Adolescent Parenting Inventory; IFEEL = IFEEL emotion recognition task; QMEE = The Questionnaire Measure of Emotional Empathy; PPES = Parent/Partner Empathy Scale; MATCh = Matching Affect to Child Task; PCV = Plotkin Child Vignettes; HES = Hogan's Empathy Scale; CAPI = Child Abuse Potential Inventory

Outcomes from the IRI

Of the eleven studies that utilised the IRI, 6 out of 10 that reported perspective taking scores found that maltreating parents scored significantly lower than controls (de Paul et al., 2008; Francis & Wolfe, 2008; Letourneau, 1981; McElroy & Rodriguez, 2008; Perez-Albinez & de Paul, 2004; Rodriguez & Richardson, 2007; Wiehe, 2003). 4 of the 8 studies that reported empathic concern scores found maltreating parents scored significantly lower than non-maltreating (Francis & Wolfe, 2008; Leon, et al., 2014; Perez-Albinez & de Paul, 2003; Wiehe, 2003) and 5 of 7 studies that reported results on the personal distress scale found that maltreating parents scored significantly higher than non-maltreating parents (De Paul et al., 2008; Milner et al., 1995; Perez-Albinez & de Paul, 2003; 2004; Wiehe, 2003).

Outcomes from other empathy questionnaires

Of the six studies that included the AAPI, three utilised the tool as a measure of CM potential rather than to indicate empathy, thus the subscale scores were not reported (Rodriguez, 2013; Rodriguez et al., 2012; Rodriguez & Tucker, 2013). Rosenstein (1995) did not find a significant relationship between the subscale scores of the AAPI and parenting maltreatment status, however found that the empathy subscale was most predictive of CM. Mennen and Trickett (2011) found that non-maltreating mothers scored higher on the empathy subscale than maltreating mothers and foster mothers. Thompson et al. (2014) found a significantly weak correlation between AAPI empathy and emotional maltreatment. Of the other measures administered, studies that utilised the HES, QMEE and PPES found that maltreating parents had significantly lower levels of empathy than non-maltreating parents.

Studies utilising multiple measures of empathy demonstrated a lack of convergence between measures, which is noteworthy considering that they are intended to measure the same construct. This is demonstrated by Perez-Albinez and de Paul (2003) who reported the correlation between the IRI four subscales and the QMEE and HES for the 440 parents who

responded to their questionnaires. The QMEE had a small correlation with the perspective taking subscale ($r(440) = .26$, $p < .01$) and moderate correlation with the empathic concern subscale ($r(440) = .54$, $p < .01$). The HES had a small correlation on both the perspective-taking ($r(440) = .38$, $p < .01$) and empathic concern ($r(440) = .28$, $p < .01$) subscales. The small correlations between empathy measures emphasises that different measures assess different empathy constructs. Thus the measures used impacts upon the internal validity of studies.

Outcomes from behavioural measures of empathy

Three studies utilised novel tasks to assess behavioural measures of empathy. Milner et al. (1995) administered the ERQ whilst mothers watched videotapes of infants in four conditions (baseline, smiling, quiet and crying). Although they did not find a significant difference between parents on individual conditions, they found a significant increase in empathy between baseline and crying conditions for low-risk mothers, but not for high-risk mothers. Furthermore, high-risk mothers showed significant increased distress in the crying condition, whereas low-risk mothers did not. This finding suggests that empathy is a state and a process affected by cognitions.

Rodriguez et al. (2012) utilised eye-tracking whilst parents read vignettes which were designed to evoke empathy and to measure attributions. There was no significant difference found between high- and low- risk parents, although there was a significant correlation between AAPI total score and the total time taken to read the empathy vignettes (which authors imply as a processing difficulty). The effect of time taken suggests an effect of cognitive empathy. The study may have benefited from including a measure of cognitive ability (such as IQ) to ascertain whether the effect found was of empathy or of processing speed.

Rodriguez (2013) utilised the MATCh task as a behavioural measure of empathy. Parents were required to watch videos of their children listening to stories and determine the emotion that their child reported. They also completed Plotkin Child Vignettes (PCV; Plotkin, 1983), which are vignettes detailing hypothetical situations of the respondent's child misbehaving. Respondents have to identify the extent to which they considered the child to be misbehaving to annoy the parent and the punishment that would be given. There was a significant correlation found between score on the CAPI and AAPI empathy scale and greater disparity between mothers' and children's emotional rating. The MATCh score was significantly correlated with PCV scores of punishment and annoyance and IRI scores of empathic concern and perspective-taking. This study did not find significant differences on the IRI scales between maltreating and non-maltreating parents.

Outcomes relating to empathy towards own child

The study by Rodriguez (2013) was one of three studies in the review that measured parental empathy towards their own children. Francis and Wolfe (2008) found that maltreating fathers had significantly lower perspective taking and empathic concern scores towards their own child than non-maltreating fathers. Perez-Albinez and de Paul (2004) measured parental empathy towards their child using the PPES and found a significant difference between high and low risk parents. Although Rodriguez (2013) did not find significant differences on the IRI between maltreating and non-maltreating parents (using the standard IRI), he found a significant negative correlation between parents CAPI abuse risk score and parents ability to read their own child's emotion. Thus, all three studies found that abusive parents were less able to empathise with their own child.

Other Dependant Variables

8 of the 17 reviewed studies measured empathy alongside other constructs. Two studies assessed family violence and neglect utilising the parent-child Conflict Tactics Scale (PC-

CTS; Straus, Hamby, Finkelhor, Moore & Runyan, 1998) (McElroy & Rodriguez, 2008; Rodriguez & Richardson, 2007), two studies measured parental anger using the State-Trait Anger Inventory (STAXI; Spielberger, 1988) (Francis & Wolfe, 2008; Rodriguez & Richardson, 2007), four measured parental stress using the Parenting Stress Index (PSI; Abidin. 1990) (Francis & Wolfe, 2008; Rodriguez & Richardson, 2007; Rosenstein, 1995), one measured parental stress using the Perceived Stress Scale (PSS; Cohen et al., 1983) (Rodriguez & Tucker, 2014) and one used the Schedule of Recent Life Experiences (Holmes & Rahe, 1967) (Letourneau, 1981).

Another investigated construct was parental expectations of the child, which was measured in two studies (McElroy & Rodriguez, 2008; Rodriguez & Richardson, 2007) utilising the Parental Opinion Questionnaire (POQ; Azar, Robinson, Hekimian & Twenyman, 1984) and The Child Development Questionnaire (CDQ; Mash, 1980) respectively. Parent attributions were measured in three studies, utilising the Parent Attribution Test (Bugental, 1998) (Rodriguez & Richardson, 2007) and Plotkin Child Vignettes (Plotkin, 1983) (Rodriguez, 2013; Rodriguez et al., 2012; Rodriguez & Tucker, 2014). Other assessed factors were mental health (Francis & Wolfe, 2008; Rodriguez & Tucker, 2014), disciplinary styles (McElroy & Rodriguez, 2008), locus of control (McElroy & Rodriguez, 2008; Wiehe, 1987), frustration tolerance (McElroy & Rodriguez, 2008), family environment (Mennen & Trickett, 2011), affective attachment (Rodriguez & Richardson, 2007), narcissism (Wiehe, 2003), social support and loneliness (Rodriguez & Tucker, 2014).

Regarding the relationship between the CTS and empathy, Rodriguez and Richardson (2007) found a significant correlation between the IRI perspective-taking subscale and the parent-child CTS physical assault scale ($r(115) = -.28$, $p < .01$). McElroy and Rodriguez (2008) found a similar correlation between the two subscales ($r(73) = -.26$, $p < .05$). This

suggests that there may be a moderate relationship between physical abuse and cognitive empathy.

Two of the reviewed studies examined anger as a CM risk factor. Rodriguez and Richardson (2007) found a correlation between the STAXI and both the IRI perspective-taking subscale ($r(115) = -.54$, $p < .001$) and CAPI abuse scale ($r(115) = .49$, $p < .001$). Francis and Wolfe (2008) found that abusive fathers scored significantly higher on state anger ($F(1, 45) = 5.59$, $p < .05$), anger expression out ($F(1,45) = 8.34$, $p < .01$) and anger expression index ($F(1,45) = 4.06$, $p < .001$), and that maltreating fathers showed significantly more clinically elevated scores on all of the four subscales. Unfortunately for the aims of the current review, authors did not report the relationship between empathy measures and the STAXI.

With relation to parenting stress, Francis and Wolfe (2008) found that abusive fathers reported significantly greater stress on all of the PSI subscales and the PSI total score ($F(1,37) = 14.07$, $p < .0001$). They did not report the relationship between PSI and empathy scores. Rodriguez and Richardson (2007) found a significant relationship between PSI total score and IRI perspective taking scores ($r(115) = -.28$, $p < .01$) and CAPI abuse scale ($r(115) = .25$, $p < .01$). Similarly, Rosenstein (1995) found a significant relationship between PSI and empathy scores ($r(29) = -.48$, $p < 0.015$). Considering that all maltreating parents scored higher on the IRI 'personal distress' scale, the results suggest that there may be an interaction between parenting stress, cognitive empathy and CM. Rodriguez and Tucker (2014) did not report the relationship between empathy and parenting stress, although they found a significant correlation between parent distress and abuse risk ($r(= 0.62$, $p < 0.001$). They found that parents with the highest CAPI abuse risk scores had high levels of personal distress and lowest levels of social support, linking the role of social support in this interaction.

Another factor relating to cognitive processes is parental attributions. Rodriguez et al. (2012) and Rodriguez (2013) did not report parental attribution scores with relation to empathy however Rodriguez and Richardson (2007) found a significant relationship between IRI perspective taking and parental attributions ($r(115) = .31$, $p < .001$).

Discussion

The current review aimed to systematically review existing literature to understand the relationship between empathy and CM. A total of 17 studies were included in the review, as a result of meeting the inclusion/exclusion criteria and scoring above 60% on the quality assessment. Combining results from both standardised and non-standardised empathy measures, 16 studies found that maltreating parents, compared to non-maltreating parents, showed significantly lower levels of empathy to some degree. Thus with relation to the review question, it appears that there are empathy differences between maltreating and non-maltreating parents. It is important to recognise however that the quality of reviewed studies varied. Of concern for the validity of findings is that five studies (30%) failed to assess whether control participants had a history of maltreatment and rather assumed that they had not.

The review suggests that maltreating parents specifically show cognitive empathy deficits, with 6 out of 10 (60%) studies finding significant differences on the perspective taking subscale of the IRI. Findings related to affective empathy were less consistent. Similar findings from the current review were found by Beven, O'Brien-Malone and Hall (2004) who did not find a significant difference on the empathic concern scale, but did on the other three IRI scales. Given that only four of eight studies that included the empathic concern subscale found that maltreating parents scored significantly lower than controls, it may be that maltreating parents do not have significantly different empathic concern but have deficits in other more cognitive aspects of empathy, such as perspective-taking. This hypothesis is more consistent with the findings of the current review and Joliffe and Farrington's (2004) meta-analysis of offenders.

The above findings indicate that although most studies found a significant difference in empathy on some measures, no scale or assessment resulted in a significant outcome

difference in all studies. This highlights the disparity across studies as well as the complexity of CM. The most robust finding was that 5 out of 7 (71%) studies found maltreating parents showed significantly higher PD scores than non-maltreating parents. The PD scale assessed how parents respond in emergency situations, higher scores indicating higher distress. Thus the findings suggest that maltreating parents have poorer emotional regulation than non-maltreating parents. Considering how this links to empathy, evidence has shown that experience of intense emotions negatively impacts upon information processing (Easterbrook, 1959). Thus it may be hypothesised that the higher emotional distress of maltreating parents inhibits their cognitive processing which would explain the current finding of cognitive empathy deficits in maltreating parents. Neurological research has found that the prefrontal cortex is involved in reducing amygdala activation (distress) when experience adversity or witnessing others in distressed states. This facilitates the observer to access more cognitive brain structures by accessing right temporo-parietal junction, precuneus and posterior cingulate brain regions, which are involved in distinguishing personal distress from empathy. (Decety & Jackson, 2006). In this respect, rather than having a lack of empathy, parent's emotional dysregulation overrides/inhibits their empathic experience. Thus if emotional regulation is considered a confound in the assessment of empathy processing, it would also provide an explanation for why outcomes are inconsistent within and between studies.

Given that emotion regulation appeared to be a factor that significantly differed between maltreating and non-maltreating parents, it is unfortunate that only one study measured personality traits. In addition to using the IRI (and finding significant differences between maltreating parents on the EC, PT and FS scales), Wiehe (2003) measured narcissism and found that maltreating parents scored significantly higher on the Hypersensitivity Narcissism Scale and on 3/6 scales of the Narcissistic Personality Inventory.

Combining these outcomes maltreating parents had poorer impulse control, lower self-confidence and higher narcissistic traits than controls.

Referring to CM perpetration literature, evidence suggests that CM perpetrators have a high prevalence of CM victimisation and emotion regulation deficits. Both of these factors are commonly found in other Cluster B personality disorders (Borderline and Antisocial Personality Disorder) (American Psychiatric Association, 2013). I have already proposed that emotion dysregulation may inhibit empathic capacity thus personality styles defined by emotional dysregulation are likely to have an increased risk of empathic deficits. Accordingly it would conceptually make sense for other cluster B personality disorders to link to empathy deficits and CM. Thus rather than researching one personality disorder in isolation, it would be more informative to investigate wider personality patterns to ensure that correlation between personality styles does not confound findings.

A stark finding from the review is the scarcity of research investigating the effect of empathy in CM perpetration in the UK. What is available internationally is saturated by two research institutes in Spain and America who adhere to the same model of CM. The lack of interest in the current topic is evidenced in that many reviewed studies did not explicitly aim to measure empathy, but did so as part of a broader construct such as parenting attitudes or cognitive differences. Thus many studies did not report their findings with relation to empathy and CM. Furthermore, given that neglect is the most common form of CM, it is interesting that only two studies included neglectful parents (de Paul et al., 2008; Thompson et al., 2014).

The current review identified that the disparity in empathy definition conflates how it is operationalised and measured, as previously highlighted in Jolliffe and Farrington's (2004) review. For example, consider IRI empathic concern items, such as 'When I see someone being taken advantage of, I feel kind of protective towards them'. These items suggest that

the ‘empathic concern’ scale adheres to Batson’s (1991) definition of empathy being an affective state coherent with another’s welfare. None of the studies utilising the IRI made it explicit whether authors intended to assess this more sympathetic aspect of empathy.

None of the reviewed studies provided justification for any of their chosen empathy measures and only two explicitly stated the type of empathy they were examining. This is problematic considering the current review highlights a lack of consistency between measures and that different measures assess different aspects of empathy. In summary, the identified disparity in defining empathy as a construct makes the synthesis of available literature problematic. Results have been published under the umbrella term of empathy, but refer to different constructs.

The extent to which findings from the current literature can be applied to the wider parent population may be questioned, as many participants were recruited from child protection services. Studies utilising child protection referrals were limited by this process. Considering the finding that child protection services receive referrals for 1.5% of the total child population (Gilbert et al., 2009) and that the NSPCC found a CM prevalence in the UK of 24.5% (Radford et al., 2011), there is likely to be participant bias in studies recruiting from child protection services. Studies that attempted to capture a more representative population by giving questionnaires to school populations were limited by low response rates.

Most studies (apart from Thompson et al., 2014) were retrospective, measuring parents at one time on a number of measures, thus the designs of all reviewed studies were limited. Less than half of the reviewed papers included parents who had been proven to have maltreated their children, hampering the applicability of findings to the maltreating population. Four studies conflated maltreatment status by assuming their control population was non-maltreating, without any measure administered. For example, Wiehe (2003) included foster parents as a control group, justifying this decision by emphasising that foster

parents had empathy training and were less likely to be abusive. However, research has indicated that CM is also sometimes perpetrated by foster parents (Biehal & Parry, 2010) thus the applicability of their findings is confounded.

The utilisation of the CAPI as a measure of maltreatment is likely to have distorted findings, considering the CAPI can often result in false negative classifications (Milner, 1989). The CAPI further limits applicability of findings as it was designed to indicate potential for physical abuse rather than neglect or emotional or physical abuse (Milner, 1986). Thus with the relatively small literature cohort there is a lack of research examining empathy of neglectful parents. Furthermore, the literature search found no studies examining empathy of parents who sexually abuse their children, an unexpected finding considering the maturity of empathy research and intervention in the sexual offender field (Ward & Durrant, 2013). Thus with relation to the review applicability, the findings may only be applied to physically abusive parents.

Applicability of the findings may be further limited, considering studies in the review were over-represented by females. The lack of male representation in study populations is in part due to research focusing on mothers, but those that included both male and female carers reported greater difficulty recruiting males than females. However, because there are conflicting findings as to whether there are empathy differences between males and females (Eisenberg & Lennon, 1983; Klein & Hodges, 2001), it is important to understand this effect in maltreating parent populations and understand the impact it has upon children, as well as its ability to indicate maltreatment risk. Perez-Albeniz and de Paul (2004) were the only study to report gender differences (as this was the aim of their research question). They did not find an effect of gender on empathy measures despite a cohort of literature suggesting that females have a greater empathic capacity than males (Davis, 1983; De Corte et al.,

2007). Perez-Albinez and de Paul's results however must be interpreted with caution considering the small sample sizes in the study.

Furthermore, research suggests that empathic responding differs depending upon the person with whom you are empathising (Cialdini, Brown, Lewis, Luce, & Neuberg, 1997). None of the studies made it clear whether they were assessing a parent's empathy towards their own child, children or to people in general. This has to be assumed from the assessments used. For example, the AAPI questions relate to 'children', whereas the IRI examines general empathy towards hypothetical people. The more global assessment of empathy (IRI) may not adequately capture a maltreating parents' state empathy towards their child. This is more apparent considering the aforementioned sexual offender literature that suggests empathy is dependent upon the person and situation (Fernandez, Marshall, Lightbody & O'Sullivan, 1999; Webster & Beech, 2000).

Thus CM literature is limited in its ability to determine whether maltreating parents show differing levels of empathy towards their own child or towards other people/children. The three studies that examined this in the current review (Milner et al., 1995; Rodriguez, 2013; Rodriguez et al., 2012) suggested that maltreating parents had empathic deficits towards their children. However as none of the designs directly compared empathy towards own child with empathy towards others (adults or children), deducing conclusions based on these studies is limited. Understanding the differing empathic capacity of maltreating parents towards different people (adults, family members, children, own children) may be a useful consideration in the development of risk assessment tools that include empathy as a risk factor.

Measures of empathy utilised in the literature contain subscales that assess measures other than empathy, such as the AAPI. It would be outdated to consider CM through a single-causal lens of parental empathy, given the extensive research identifying risk factors in all

domains of the ecological model (CWIG, 2004). Thus it is important to consider other factors that interact with low parental empathy and CM. All studies administering the IRI found that maltreating parents showed significantly higher levels of personal distress. Added to this, other measures found that maltreating parents had significantly higher levels of anger and stress and negative attributions towards children. Thus, considering the ecological and cumulative models of empathy, the reviewed research suggests that variables such as anger, stress and negative attributions hamper the actualisation of empathy. Understanding this interaction will be important for future research in the field.

Further Limitations of Reviewed Studies

With empathy being such a broad interactive construct, it is of concern that the current review is largely informed by the outcomes of 'paper and pencil' questionnaires, rather than behavioural assessments. As highlighted in Rodriguez's (2013) study that utilised questionnaires and behavioural assessment, parents' perception of their empathy (as indicated by questionnaire responses) may be different to their utilisation of empathy. Disparity in results from questionnaires and behavioural assessments was also found in Miller and Eisenberg's (1988) review.

The analysis used in studies also contributed to the difficulty in understanding the relationship between empathy and parental perpetrated CM. Parametric and non-parametric tests were used between studies, making the synthesis of data problematic. Some authors only conducted correlations between variables, without reporting how many parents scored above and below the CAPI cut-off. Thus readers are only informed that there is a relationship between CAPI and empathy score, rather than specifically knowing whether parents at high-risk have significantly lower levels of empathy than parents at low-risk (Rodriguez et al., 2012).

Studies that did not find empathy deficits have not seemed concerned by this finding. However their studies may be suggesting that maltreating parents have the capacity to empathise, but not with their own children (akin to sexual offender research) (Fernandez, Marshall, Lightbody, & O'Sullivan, 1999; Webster & Beech, 2000). Alternatively it suggests that maltreating parents are able to understand the emotions of their child, they are able to feel the pain of their child, but this negative painful feeling does not deter maltreating behaviour. The latter finding would suggest a maltreating parent cohort with sadistic traits, which would necessitate a change of focus for parenting interventions, which currently focus on parent education to increase cognitive empathy (Sanders et al., 2004). Both hypotheses would necessitate further research to understand how and why parents without empathy deficits maltreat their children.

Thus considering maltreating parents from a personality perspective, it is surprising that only one study in the current review considered personality as a factor that interacted with empathy. Some personality traits are characterised by empathy deficits, thus it would be relevant to the CM literature to understand whether maltreating parents with different personality traits have different empathy deficits.

Another factor the reviewed studies have failed to account for is IQ. Many studies recorded parental education level, however did not further assess the cognitive capabilities of parents. This is problematic as there may have been cognitive functioning differences between maltreating and non-maltreating parents that could have accounted for findings. Indeed, Joliffe and Farrington (2004) found that the significant relationship between empathy deficits and offending disappeared once participants were matched for IQ and socio-economic status. Other studies however have not found an effect of IQ on empathy (Davis, 1983; Mayer & Geher, 1996). Thus future research will benefit from including IQ, to ascertain how this interacts with empathy and CM.

Strengths and Limitations of the Review

The purpose of a systematic review is to apply a robust search strategy for extracting information, to prevent a bias when answering a question and generating new hypotheses (Sayers, 2007). In an attempt to remove bias, multiple databases were searched, broad search terms were included, papers hand searched for their applicability, and both PICO and quality assessments were applied. Nonetheless, there is potential that the current review contains bias.

Primarily, unpublished theses were not considered. Some researchers recognise the limitations of including dissertation theses as retrieval of these documents can be problematic and often contain methodological confounds (Vickers & Smith, 2000). However others believe the inclusion of dissertation theses is crucial for an eclectic unbiased review (Egger, Dickerson, & Smith, 2007).

Key authors in the field were contacted to retrieve articles that were not available online or via institution libraries and to question whether they knew of any unpublished documents that would be of use to the literature review. Authors were able to provide requested documents. Authors however did not know of any other studies that would be useful in the field. On the one hand, this provides confidence that the review encompasses the important articles in the field. One may however consider there to be bias in the selection of papers, as they have all been subject to peer review and to the criteria for publication. This is more pertinent in understanding that nine of the seventeen studies (53%) were published in the journal 'Child Abuse and Neglect'. Song et al., (2009) identified that studies finding positive results are more likely to be published, thus creating a bias in available data.

The review may be considered biased as it was conducted by one researcher. Although it is generally accepted that one person is able to conduct a literature review, it is subject to the opinion of one researcher and others may have placed different emphasis on different

factors. In an attempt to overcome this, inter-rater reliability was assessed when completing quality assessments.

The current review marginally diverted from the research question in discussing other factors that were measured and the outcomes of these. Considering CM from an ecological model, it is useful to understand how empathy interacts with other risk/protective factors.

Conclusions and Recommendations

Empathy is a concept that has been researched since the late 18th century. Its relationship with CM has been investigated in the last forty years. Despite this forty year span there has been little progress or advancement of knowledge with regard to the relationship between empathy and CM, with studies conducted in the last five years utilising similar paradigms and having the same aims as preliminary studies. There appears to have been little attention paid to synthesising results from available research.

Despite limitations, this review suggests that maltreating parents have significantly lower levels of empathy than non-maltreating parents, both general empathy and empathy towards their own children. Specifically, research suggests that deficits are related to cognitive more so than affective empathy.

Whilst integrating and clarifying findings from previous research, the current review has identified where future research should focus. There is primarily a need for UK based studies, ideally prospective longitudinal cohort studies, that include a large sample and a range of questionnaire and behavioural assessments of both empathy and CM. Undertaking this research will be costly in relation to recruiting and follow-up of participants, however it will give a more concise understanding of how empathy, along with other risk and protective factors, interact with CM risk.

Studies, whether prospective or retrospective, will benefit from larger sample sizes to ascertain a more robust effect. Furthermore, studies need to more clearly operationalise empathy and consider to whom they are interested in parents showing empathy towards (e.g. to the child, to children or to adults). Researchers need to be aware of the empathy assessment used and whether it is a valid measure of the empathy construct they intend to measure.

There is a need for a clearer understanding of whether parents who perpetrate different types of maltreatment (sexual, physical, emotional, neglect or abuse) show different types of empathy, as well as whether there is a difference between maltreating mothers and fathers. Such information, as it has done with sexual offender treatment programmes, will inform risk assessments and intervention.

Chapter 3

Critique of a Psychological Assessment: The Interpersonal Reactivity Index

Introduction

Chapter 2 explored the role of empathy in child maltreatment, finding support for the notion that lower empathy is linked to child maltreatment perpetration. The chapter identified that the Interpersonal Reactivity Index (IRI; Davis, 1980) was one of the most commonly used assessments of empathy in the empathy and child maltreatment literature and benefitted from measuring both cognitive and affective empathy. Accordingly, this chapter explores the psychometric properties of the IRI.

As explored in the introduction chapter of this thesis, empathy is a multi-faceted construct. It is considered a pivotal element in a number of human behaviours, not only in aggression, but in the success of job roles (Paterson, Reniers & Vollm, 2009), effectiveness of treatment (Odgers, 2014) and development of relationships (Britton & Fuendeling, 2005). It is important to understand the role empathy plays in human behaviour in order to be able to develop interventions to prevent harmful behaviours and increase effectiveness of positive behaviours.

In order to effectively measure the influence of empathy in human behaviour a psychometric must be able to assess and distinguish between the different facets of empathy. Standardised measures of empathy differ in their ability to achieve this goal, as discussed further in this chapter. It was considered important to analyse the psychometric properties of the IRI in this chapter due to its assessment of multiple empathy domains and its popularity within the literature as a measure of empathy. Additionally, as the IRI was used in the study of empathy in offenders presented in Chapter 4, it was important to understand the validity and reliability of the measure in order to effectively interpret outcomes and limitations.

Development of the IRI

The IRI was developed in a climate whereby there were a number of already available measures of empathy used to examine individual empathic differences in a number of different settings. An early scale designed for use with children, the Affective Situations Test for Empathy (Feshbach & Roe, 1968), required respondents to report how they felt when viewing pictures of other children in varying emotional states. Although intended to be a measure of affective empathy, critics argued that the assessment required cognitive empathy as the respondent primarily had to interpret the stimulus emotion before evoking their own response. Similarly, the Emotional Empathy Scale developed by Mehrabian and Epstein (1972) was intended to be an assessment of affective empathy but was confounded by inclusion of cognitive questions. An alternative measure of empathy was developed by Hogan (1986), named the Hogan Empathy Scale, which provides questions to asses both cognitive and affective empathy. However, as with the aforementioned measures, the outcomes from this assessment provide a single empathy score. Thus prior to the development of the IRI, the three most commonly used assessments of empathy provided a global perspective of empathy.

Davis (1980) however recognised that cognitive and affective empathy were distinct constructs and that assessments of empathy were confounded by integrating cognitive and affective outcomes into a global score. He stated in his research paper that the true impact of empathy on individual behaviour could only be understood by examining empathic constructs individually. Accordingly, the IRI was developed as a multidimensional assessment of empathy.

Administration of IRI

As an assessment of cognitive and affect empathy, the IRI has four scales, each with seven items:

- Perspective-taking (PT): the ability to understand the perspective of another person.
  E.g. "I try to look at everybody's side of a disagreement before I make a decision".

- Empathic concern (EC): The ability to experience compassion for another person experiencing negativity.
  E.g. "I often have tender, concerned feelings for people less fortunate than me".

- Personal distress (PD): The extent to which an individual experiences distress when witnessing someone else's misfortune.
  E.g. "In emergency situations, I feel apprehensive and ill-as-ease".

- Fantasy (FS): the tendency to identify with fictional characters
  E.g. "I daydream and fantasize, with some regularity, about things that might happen to me".

Affective empathy is assessed using the IRI EC and PD scales, whilst cognitive empathy is assessed using the PT scale. Debates exist regarding whether the FS scale assesses cognitive or affective empathy and as explored later in this report, whether the FS scale assesses empathy at all.

The IRI takes approximately 10 minutes to administer. When completing the IRI respondents are required rate each item on a 5-point Likert Scale whether they consider the

items describe them “not well” to “very well”. In this respect, the IRI is considered an ordinal scale. Four separate subscale scores are derived, with no ‘total score’. The scale was designed to be scored from 0-4 (Davis, 1980). However, Konrath, O’Brien and Hsing (2011) reviewed all papers that utilised the IRI as a research tool and found that only 18 used the 0-4 scale whilst 84 used a scale of 1-5.

Application of the IRI

Since its development, the IRI has been utilised as an assessment of empathy in a variety of different contexts, for both clinical and research purposes. A large cohort of literature has utilised the IRI to assess empathy in medical and nursing clinician’s through-out training and clinical practice (Bellini & Shea, 2005; Hojat, Mangione, Kane & Gonnella, 2005; Yarnold, Bryant, Nightingale & Martin, 1996). Other areas of research have utilised the IRI to examine deficits in patients with dementia (Oliver, Mitchell, Dziobek, MacKinley, Coleman, Rankin & Finger, 2015), autistic spectrum disorders (Rogers, Dziobek, Hassenstab & Wolf, 2007), schizophrenia (Haker, Schimansky, Jann & Rössler, 2012) and other disorders. The tool has also been used to assess age and developmental differences in empathy (Konrath, O’Brien & Hsing, 2011) as well as a way to distinguish characteristics of groups of people, such as health volunteers (Paterson, Reniers & Vollm, 2009).

With relation to forensic practice, the IRI has been used in multiple contexts. Some examples relating to research include examining the characteristics of parents who abuse their children (Milner, Halsey & Fultz, 1995), differentiating between different types of offenders (Lindsey, Carlozzi & Eells, 2001) and understanding the relationship between empathy and offending (Joliffe & Farrington, 2004). The IRI has also been used as a pre- and post-treatment assessment measure in prison offending programmes (Eisenberg & Fabes, 1990; Kim, Choi, Rhee, Kim, Joung & Kim, 2012).

Despite the IRI being such a prominent measure of empathy in current practice, its psychometric properties have not yet been thoroughly amalgamated and analysed. The present report aims to critically analyse the utility of the IRI as a measure of empathy.

Psychometric Properties of the IRI

In order to assess the quality of a psychometric assessment, Kline (1986) suggested five criteria need to be met. Two important aspects of a psychometric measure that he identified are **reliability** and **validity**, which are the ability for the assessment to produce an outcome that is accurate and replicable and is an accurate measurement of the construct. He considered that the assessment tool had to provide outcomes that were of an **ordinal or ratio scale**. Doing so provides meaningful differences between the ratings thus allowing parametric analyses to be used. This is an important component for the development of the criteria: **applicable normative data**. Normative data allows an assessment to have standards and therefore provides a basis for future respondents and samples to be compared. Another important component of a psychometric assessment is **discrimination**. This is the ability for the assessment to discriminate between participants based upon the construct measured and for scores to be spread over the outcome range.

Reliability

Internal Reliability

Internal reliability is the extent to which each item within the assessment consistently measures the same construct. The most commonly used assessment of internal reliability is Cronbach's Alpha which provides a measure of reliability from 0 to 1, with 1 indicating high internal reliability. Although the cut-off for reliability is arbitrary, a score of 0.7 or higher is generally considered to indicate good internal reliability (Nunnaly, 1978).

A number of papers with large samples have reported the internal reliabilities of each of the IRI scales. Table 8 provides a summary of the findings of four papers spanning the last four decades. In the paper describing the development of the IRI scale, Davis (1980) reported the internal reliabilities of the IRI scales in two separate research projects: the development of the scale (n = 427) and re-testing the scale (n = 1, 161). All of the scales produced alpha coefficients between 0.71 – 0.79 indicating a high level of internal reliability. Similar findings of high internal reliability were found in a longitudinal cohort study (O'Brien, Konrath, Gruhn & Hagen, 2012) that utilised the EC and PD scale of the IRI with 75,263 participants when each participant was aged 18 and 21. Alpha coefficients between 0.75 and 0.83 were reported.

In contrast however, lower alpha coefficients were reported in a large sample of 289 prisoners (Ireland, 1999), which found that when negatively worded items were removed from the EC and PD scales, internal reliabilities increased. Similarly, Lauterbach and Hosser (2007) found that removing negatively worded items increased the alpha coefficients for the FS and EC scales. In a study with 88 high security prisoners, Beven, O'Brien-Malone and Hall (2004) used Corrected Item-Total Correlations to measure internal reliabilities of each item and found that all of the negatively worded items had scores lower that 0.30, indicating that the items did not contribute towards the measurement of empathy (DeVellis, 1991) and negatively impacted upon the internal consistency of the IRI. Thus although research has demonstrated the positive internal reliability of the IRI, this data suggests that internal reliabilities with prisoners may be improved by removing negatively worded questions. Authors suggested that these questions require longer processing time and higher level of cognitive functioning. However, there have been a number of studies using the IRI that have demonstrated no effect of IQ scores (Davis, 1983; Mayer & Geher, 1996). Thus it does not appear that the impact of cognitive functioning on IRI scoring is as linear as suggested.

It is important also to consider the statistical/methodological processes that occur when removing negatively worded items. Each scale has 2 negatively worded items, thus the adapted scale would have only 5 questions per scale, removing 28% of the questions from the full questionnaire. This would thus reduce the bandwidth of possible outcomes resulting in an artificial rather than genuine increase in reliability.

Table 8.
Reliability Coefficients (Cronbach's Alpha) for IRI scales

		Fantasy	Perspective Taking	Empathic Concern	Personal Distress
Davis (1980)	Male Community (N=579)	0.78	0.75	0.72	0.78
	Female Community (N=582)	0.75	0.78	0.70	0.78
Ireland (1999)	Male and Female prisoners (N=289)	0.64	0.70	0.43/0.60*	0.52/0.60*
Lauterbach & Hosser (2007)	Male Prisoners (N=839)	0.66/0.74*	0.77/0.77*	0.77/0.81*	0.63/0.63*
O'Brien, Konrath, Gruhn & Hagen (2012)	Male and Female community (N=75,263)			0.75 - 0.83	0.82

Note: * = when reverse items were removed;

Test-Retest Reliability

Test-retest reliability is a measure of external reliability. It is the extent to which a measure is able to produce the same outcome at different assessment intervals (in the absence of intervention). The test-retest reliability of the IRI has not been extensively researched, as

many applications of the measure have been to examine changes over time following intervention or training, with expected change.

In the development of the IRI, Davis (1980) reported test-retest reliabilities in a sample of 109 undergraduates (56 males; 53 females). Participants completed the IRI on two separate occasions 60-75 days apart. Correlations between the two testing periods ranged from 0.61 to 0.81 (Table 9), indicating a good level of test-retest reliability. Test-retest reliabilities were also obtained by Davis and Franzoi (1991) in a sample of 205 secondary school students. Students completed the IRI at 1-year intervals for four years. The results of this study (summarised in Table 9) demonstrated moderate to high test-retest reliabilities for all of the scales. Authors found higher correlations between IRI scores at year 2 and 3 which indicated increased empathic stability with age.

Table 9.
IRI Test-retest reliability coefficients

		Fantasy	Perspective Taking	Empathic Concern	Personal Distress
Davis (1980)	Adult Male (N=56)	0.79	0.61	0.72	0.68
	Adult Female (N=53)	0.81	0.62	0.70	0.76
Davis & Franzoi (1991)	Adolescent Male (N=103)	0.62 - 0.70	0.58 - 0.65	0.48 - 0.50	0.58 - 0.62
	Adolescent Female (N=102)	0.62 - 0.70	0.63 - 0.72	0.35 - 0.59	0.59 - 0.66

Validity

Face Validity

Face validity is the extent to which items on a measure appear to relate to the construct. Face validity alone is not a robust enough assessment of validity, however it is a useful consideration. The IRI scale items were developed from a theoretical understanding of

cognitive and affective empathy, in addition to questions used in established measures of empathy. Thus the questions of each scale appear to have face validity.

Concurrent Validity

Concurrent validity of the IRI is assessed by the extent to which the measure correlates with other measures of empathy. Perfect correlations are not desirable, as the measure would have no added value. Davis (1983) correlated outcomes from the IRI with two empathy measures: the Emotional Empathy Scale (EES; Mehrabian & Epstein, 1972) and the Hogan Empathy Scale (HES; Hogan, 1969). As described in the introduction the EES is a measure of affective empathy whilst the HES is a measure of cognitive empathy. Participants were 677 male and 667 female psychology students. In a more recent study, Lawrence, Shaw, Baker, Baron-Cohen and David (2004) correlated the IRI against their newly developed measure: the Empathy Quotient (EQ), with a sample of 28 male and female volunteers. The EQ was designed to be a measure of cognitive and affective empathy.

Results from these studies are provided in Table 10 and indicate that the IRI has good concurrent validity. The HES was moderately positively correlated with the IRI PT scale and moderately negatively correlated with the IRI PD scale. The EES showed moderate to strong positive correlations with the IRI FS and EC subscales and the EQ showed moderate correlations with the IRI PT and EC subscales.

Table 10.
Correlation coefficients between the IRI and other measures of empathy

		Fantasy	Perspective Taking	Empathic Concern	Personal Distress
HES	Male (N=677)	0.15	0.42	0.11	-0.25
	Female (N=667)	0.15	0.37	0.25	-0.40
EES	Male (N=677)	0.48	0.22	0.63	0.36
	Female (N=667)	0.56	0.17	0.56	0.12
EQ	Male and Female (N=28)	0.42	0.49	0.42	0.16

Content Validity

Content validity is the extent to which a psychometric measures all aspects of the construct. Thus in reference to the IRI, content validity refers to the extent to which it measures all the features of empathy. As already described, Davis's (1980) purpose for developing the IRI was in order to create a multidimensional measure of empathy that assessed both cognitive and affective empathy. Functional analysis and correlation with other measures of empathy indicates that the IRI is able to assess both of these constructs well. These are described in detail later in the chapter.

However, a construct that the IRI is unable to measure is dispositional empathy. This is the utilisation of empathy 'in the moment'. There are a number of research papers that purport to the IRI assessing dispositional empathy, however the self-report reflective nature of the questionnaire refutes the validity of this claim. Rather the structure of the assessment provides an assessment of trait empathy (Geer, Estupinan & Manguno-Mire, 2000). Some study designs have overcome this difficulty by using the IRI alongside task-based assessments. Additionally, the IRI does not require an individual to rate their empathic capacity towards a specific person and thus it is unable to be used to test for theories that suggest empathy differentiates depending upon a subject's relationship with the person (Marshall & Marshall, 2011). In order to overcome this, some studies have adapted questions of the IRI to include specific people (e.g. your child, your partner) (Hawk, Keijseres, Branje, Van der Graaff, de Wied & Meeus, 2013).

Factor Analysis

One method of assessing the reliability and validity of the IRI is to analyse the extent to which the items of each scale contribute to the scale outcome and whether there are other factors that contribute to the outcome, by computing a factor analysis. Davis (1980) used

Jöreskog's (1969) method of oblique rotation to factor analyse the initial IRI questions and then again to re-assess the reliability of the items. This is a 'confirmatory maximum likelihood factor analysis' which allows researchers to pre-specify factor loadings whilst free parameters and alternative factors can be explored utilising a method of maximum likelihood. This was an appropriate factor analysis choice for the development of the IRI, as Davis (1980) had developed 50 questions designed to assess aspects of both cognitive and affective empathy.

From the 50 questions which Davis (1980) factor analysed, four factors were demonstrated (FS, EC, PT, PD) in both male (n=201) and female (n=251) samples. Davis also reported that smaller factors emerged, however it was found that these were uninterpretable and thus focused on the four strongest factors, in line with procedures outlined in Comrey (1978). He then developed a 45-item questionnaire utilising the questions that contributed to the factors and new questions were developed in line with the four factors. This questionnaire was given to 427 psychology students (221 males; 206 females). Using Jöreskog's factor analysis with oblique rotation, Davis predefined the four factors into the analysis and found that most items loaded onto one item and supported the model.

The final version of the IRI was developed by including items that loaded most heavily onto the factors and excluding items that loaded heavily onto more than one item. This was provided to 1,161 students (579 male; 582 female) utilising the same factor analysis procedure as described above. Results demonstrated that for both males and females the items loaded most heavily onto the factor they were defined to.

Thus the development of the IRI was guided by factor analysis and findings provided support for the presence of the four scales. Subsequent research has also provided support for

the presence of the four factors in samples of children (Litvack-Miller, McDougall & Romney, 1997) and adults (Carey, Fox & Spraggins, 1988; Cliffordson, 2002).

However, there is also a cohort of data that contests the construct validity of the IRI and proposes alternative models. For example, Alterman, McDermott, Caccioa and Rutherford (2003) criticised Davis (1980) for using confirmatory factor analysis due to having a bias for the pre-specified four factors. Furthermore, they recognised that 6 of the items for men and 7 for women had loadings below the suggested cut-off of 0.40, compromising the test validity.

In response to their critique, Alterman, McDermott, Caccioa and Rutherford (2003) examined the factorial validity of the IRI in 241 patients undergoing a methadone treatment programme using confirmatory factor analysis, loading Davis's original four factors and then using exploratory factor analysis to explore alternative models. Utilising this methodology, Alterman et al. (2003) found that a three-factor model consisting of 18 items best represented empathy. The first factor, described as 'empathy' consisted of nine items from the EC and PT subscales, the second scale consisted of five PD items and the third scale consisted of four FS items. It was found that the PD scale did not correlate with other scales and thus Alterman et al. suggested that it did not represent empathy and rather reflected traits linked to narcissism. Although these scales had high internal consistencies (all above 0.60), combining the EC and PT subscales undermines the purpose of the IRI, which was to separately analyse cognitive and affective empathy constructs.

The IRI factors have additionally been challenged, due to the correlations found between the four scales. Authors have suggested that the inter-correlation of the IRI scales indicates that there is an underlying component explaining findings. For example, Cliffordson (2002) examined the construct validity of the IRI in a sample of 127 undergraduates. This study confirmed the fit of the four factor model, however authors contested the linearity of

the model (that the four factors were independent of each other). The study found that the EC subscale accounted for the measure outcomes and overlapped with outcomes from the PT and FS scale. As with findings from Alterman, McDermott, Caccioa and Rutherford (2003), the PD scale did not relate to the model. Thus, Cliffordson (2002) proposed a uni-dimensional model indicating that EC is the main component of the measure and influences outcomes on the other measures.

This model counters theories of empathy which suggest that perspective taking is a primary stage of empathy as one must understand how another feels in order to respond. It also counters Davis's (1983) theory that perspective taking is the principle component that links empathy with social functioning. Cliffordson (2002) explained that the EC scale was concerned with both cognitive and affective aspects of empathy and thus measured an individual's ability to interpret and respond to another's reaction. Again however, this model counters Davis's (1980) aim of analysing cognitive and affective empathy separately. Whilst the guiding theories of empathy suggest that cognitive and empathy are distinct but related constructs, it is important that they are able to be analysed separately.

One theme however that appears to emerge throughout the literature is the question regarding the validity of the FS scale as a measure of empathy. Many researchers have excluded the scale from their design (e.g. O'Brien, Konrath, Gruhn & Hagen, 2012). In addition to concern raised in the other studies, Baron-Cohen and Wheelwright (2004) suggested that the FS scale assesses a construct more related to imagination and the PD scale assessed emotional control, rather than affective experience. Furthermore, Corte, Buysse, Verhofstadt, Roeyers, Ponnet and Davis (2007) identified that the seven items of the FS scale highly correlate and assess the same concept, rather than differing concepts that define the same construct.

The FS scale was designed in order to measure empathy for fictional characters, under the premise that empathy towards fictional characters was a distinct construct to empathy towards people. However, Nomura and Akai (2012) examined this concept in 95 undergraduates (53 males; 42 females) utilising the IRI and a 'fictional IRI', made up of modified IRI questions to reflect empathy towards fictional characters. Results from this study indicated that there were no significant differences between empathy for fictional characters and real people, thus authors concluded that the concept underlying the FS was unsupported.

Predictive Validity

Predictive validity is the extent to which a measure is able to predict future outcomes. As previously described, empathy is considered to be a modulator of social processes and thus researchers have been interested in the role of empathy in predicting a number of related constructs, such as criminal behaviour, recidivism, helping behaviour and personality constructs.

With relation to forensic populations, the IRI has been utilised as a tool to predict reoffending. Lauterbach and Hosser (2007) utilised logistic regression to examine the ability of IRI scores to predict violence recidivism (within 24 months of release) in 577 offenders. Whilst controlling for IQ, age, sentence length and socio-economic status, lower PT scores predicted a higher likelihood of reoffending, alongside younger age and longer sentence. In a more recent study, Bock and Hosser (2014) found that the PT scale again predicted violent recidivism as did the FS scale and 'global score' (removing negatively scored items). Thus these studies suggest that the cognitive components of the IRI are able to contribute to prediction of violence recidivism.

The IRI has also been used to predict perceptual accuracy. For example, Larson, Fair, Good and Baldwin (2010) were interested in examining the relationship between IRI scores and processing of errors on the Stroop task (Stroop, 1935) in a sample of 30 undergraduates. The Stroop task presents names of colours in coloured fonts. It is found consistently that longer response times occur when the word and colour do not match. They found that higher IRI scores were related to increased error-related negativity amplitude and that higher PD scores were related to post-error slowing. However, when negative affect was controlled for, EC score also predicted error-related negativity amplitude. Although the relationship between these variables was complex, the results provide evidence to suggest that empathy modulates error processing.

Similarly, Haas, Anderson and Filkowski (2015) required 16 adult participants to complete the IRI and subsequently complete emotion attribution tasks whilst undergoing functional MRI. These results demonstrated that participants with high PD scores were significantly more likely to make quicker emotion attribution decisions and had increased anterior insula activity. The anterior insula cortex is considered to be a critical neurological component of emotional awareness (Gu, Hof, Friston & Fan, 2013) and although the direction of causality cannot be deduced from the research, findings link the PD scale with emotional processing and suggest increased emotion processing facilitates emotional attribution. Additionally, those with higher PT scores had significantly delayed emotion attribution reaction times and increased prefrontal cortex and premotor activity. The prefrontal cortex is considered to be the most evolved brain region and involved in cognitive processing, consequential thinking and decision making. In this respect, the study provides evidence that the PT scale correlates with neurological cognitive structures.

The IRI has also been used to predict altruistic pro-social behaviour. For example, Oswold (2003) found that PT scores were able to predict the likelihood that participants

would volunteer time to support others, with higher scores predicting a higher likelihood. Additionally, the EC and PT subscales have been demonstrated to positively correlate with pro-social tendencies and altruistic behaviour (Davis, Mitchell, Hall, Lothert, Snapp & Meyer, 1999; Paterson, Reniers & Vollm, 2009; Wilhelm & Bekkers, 2010). Regarding personality characteristics, the EC scale has been shown to correlate with less egotistic traits and higher emotional reactivity (Davis, 1983), whilst the PT subscale has been shown to correlate with high self-esteem and good social functioning (Davis, 1983). Thus, the IRI appears to have good predictive validity for a number of outcomes hypothesised to relate to empathic capacity.

Normative Data

As a widely used assessment of empathy, the IRI benefits from having normative data for a number of different samples, including community, forensic and clinical populations, such as patients with schizophrenia and Asperger's syndrome. Outcomes for these samples are summarised in Table 11. The IRI has also been adapted and normative data available for a number of different cultural populations, including Spanish (Carrasco, Delgado, Barbero, Holgado, & del Barrio, 2011), Chilean (Fernández, Dufey & Kramp, 2011), Chinese (Siu & Shek, 2005) and Dutch (De Corte, Buysse, Verhofstadt, Roeyers, Ponnet & Davis, 2007).

In his inaugural research paper, Davis (1983) provided normative data for male and female adults, finding that females had higher scores on all of the scales. However only the FS scale showed significant difference between male and female scores $F(1,1176) = 96.28$, $p < .001$. This finding has been replicated in a number of studies (e.g., De Corte, Buysse, Verhofstadt, Roeyers, Ponnet & Davis, 2007). Comparison of forensic outcomes (outlined in Table 11) however indicates disparity. Normative data from a high-security prison sample obtained by Bevan, O-Brien-Malone and Hall (2004) suggested that offenders have lower

scores on the FS, PT and EC subscale and higher scores on PD, compared with Davis's (1983) community sample. However, Ireland's (1999) study with offenders suggests that offenders have similar or higher scores on all of the IRI subscales compared with Davis's (1983) community sample. Thus, with regard to normative data for forensic populations, it appears that a more comprehensive review is needed in order to ascertain reliable normative data.

Comparing normative data is also limited by the scoring procedure used in studies. As mentioned in the introduction, some studies score using a scale of 0-4 and others use a scale 1-5. Often papers do not state which method they use making comparison of normative outcomes difficult. Additionally, the standardised scoring procedure for the IRI is cumulative, so that the scores for each item are added to produce an overall score. Using the 1-5 scale, each scale has 7 items with a maximum item score of 5 and a total scale score of 35. There are a number of research papers however that report notable low mean scores, as demonstrated in Table 11 with the outcomes from Hawk, Keusers, Branje, Van der Graaf, de Wied and Meeus's (2013) study of adolescents and their mothers. Other studies that report similar low mean scores include O'Brien, Konrath, Gruhn and Hagen's (2012) longitudinal study with 75,263 community participants and Konrath, O'Brien and Hsing (2011) longitudinal study with 13,737 college students. Unfortunately none of these research papers explain their scoring methodology, which limits their utility as normative data. Given the large samples in these studies, being able to utilise them for the purpose of normative data would be an advantage.

Table 11.
IRI Normative Mean Data

Authors	Sample	FS	PT	EC	PD
Davis (1983)	Adult Male (N=579)	15.73	16.78	19.04	9.46
	Adult Female(N=582)	18.75	17.96	21.67	12.28
Beven, O-Brien-Malone & Hall (2004)	Offenders (N=88)	9.28	12.99	12.83	10.14
Ireland, (1999)	Male adult prisoner (N=140)	17.6	23.7	24.0	17.1
	Male young offender (N=74)	18.7	19.8	22.3	17.6
	Female Adult prisoner (N=50)	18.4	24.6	25.8	19.2
	Female Young offender (N=20)	20.3	22.9	25.7	19.4
Hawk, Keusers, Branje, Van der Graaf, de Wied & Meeus (2013)	Early adolescent males (N=148)	2.0	1.99	2.30	1.82
	Early adolescent females (N=121)	2.40	2.16	2.68	2.18
	Late adolescent males (N=123)	2.07	2.20	2.28	1.33
	Late adolescent females (N=121)	2.67	2.52	2.76	1.90
	Mothers (N=501)	2.18	2.60	2.85	1.57
Haker, Schimansky, Jann & Rössler (2012)	Adult patients with Schizophrenia (N=21)	15.3	16.8	19.3	15.7
Rogers, Dziobek, Hassenstab, Wild & Convit (2007)	Adults with Asperger's syndrome (N=21)	11.4	10.5	16.9	15.8

Note. FS = Fantasy Scale; PT = Perspective Taking Scale; EC = Empathic Concern Scale; PD = Personal Distress Scale

The Scale Structure

There are some flaws inherently related to the IRI being a likert scale and the scoring procedure adhered to. It is unclear why Davis (1980) has developed a theory of empathy that considers affective and cognitive empathy, but developed a scale that includes an additional two concepts (personal distress and fantasy). Having four scales derived from 28 questions, thus seven questions per scale, it is questionable the extent to which the limited number of questions are able to effectively assess each scale construct.

An issue that relates to a general criticism of Likert scales is the way in which the IRI is scored. The IRI is an ordinal scale as the labels for each number are conceptual and thus intervals between them cannot be assumed equal. In this respect, the numbers of the scale are categorical labels rather than representative of quantity. Summing the label numbers to

generate an ‘empathy score’ implies that there are meaningful differences between scores. There is debate within the literature regarding whether Likert scale output should be summed (as with the IRI) or whether it should be averaged, both arguments identifying the flaws of converting ordinal data into interval data. (Carifio & Perla, 2008; Norman, 2010; Sullivan & Artino, 2013). The argument for averaging is that it provides a smaller bandwidth of responses and provides a more representative account of the individual’s responses. Considering the IRI, two participants may obtain a scale score of 15. Participant 1 answering ‘0’ to four questions and answering ‘5’ to three questions. Participant 2 answering ‘2’ to six questions and ‘3’ to one question. Using the ‘adding’ methodology, both participants are deemed as having the same empathic capacity, whereas averaging would portray the difference in their responses.

An additional issue relates to the reliability of data obtained from Likert scales. Research has found that often respondents are drawn to make ‘extreme responses’, thus choosing scores at either end of the scale. This is particularly true when labels are only applied to the end of a scale (in the case of the IRI applying labels to only 1 and 5). An effect known as ‘acquiescence response style’ has also been found, with respondents showing a tendency to agree with items irrespective of the content (Moors, Kieruj & Vermunt, 2014). The impact of such formatting issues is demonstrated in a piece of research by Ogden and Lo (2012) who gave participants a Likert scale and completed an interview with them. Results from the study found that the information from the Likert scale contradicted the information participants provided in interview suggesting that Likert scales did not accurately portray the subjective experiences of participants. Despite there being a plethora of research supporting the reliability and validity of the IRI as a measure of empathy, it is important to be aware of these structural flaws and their potential impact on the IRI.

Conclusions

This review has critically analysed the psychometric properties of the IRI, in line with the psychometric standards outlined by Kline (1986). As a questionnaire based assessment, the IRI benefits from being easy and quick to administer. The IRI has been used extensively in research in a variety of clinical and forensic fields which, as a benefit of having an ordinal scale, has resulted in the availability of normative data for a wide variety of populations. The validity of the IRI cross-culturally has also contributed to a plethora of data exploring empathic capacity. One limitation however is that researchers have scored the IRI in differing ways, which increases the complexity of comparing normative data.

The IRI is based upon a multidimensional concept of empathy and thus encompasses an assessment of both cognitive and affective empathy and was the first assessment to explicitly do so. The IRI's construct and predictive validity has been evidenced by its ability to predict and correlate with other constructs known to relate to empathy, such as altruism and pro-social behaviour. Extensive research has confirmed the presence of the four scales proposed by Davis (1980) and its ability to differentiate between cognitive and affective empathy constructs.

There have been concerns raised regarding the hierarchy of the measure. At present the four scales are distinct and measure cognitive and affective empathy separately. However, models of empathy suggest that empathy is a progressive construct, requiring an understanding of an emotion before being able to respond to the emotion. In this respect, findings that the EC scale is the most related to empathy is expected, given that it is one of the latter stages of empathy. Additionally, research appears to indicate that the FS scale is not a valid measure of empathy as being able to relate to fictional characters does not correlate with empathic capacity.

There have also been questions raised regarding the validity of the reverse items, with preliminary research suggesting that removing the reverse item questions increases the internal reliability of the IRI. Authors proposed that this effect was due to the increased processing time required and thus a similar effect should be seen for all samples with low IQ's rather than just offenders, which has not been shown consistently. Thus, this effect is something that would be useful to explore in future research.

Although a multidimensional measure of empathy, there are limitations regarding the use of the IRI, including its ability to assess person-specific empathy and empathy 'in vivo'. Researchers have overcome this difficulty by incorporating task-based assessments or adapting the IRI to relate to a specific population. The adaptability of the IRI is one of its strengths. Despite the limitations of the IRI, it provides an easy to use valid and reliable holistic assessment of empathy. In this respect, it meets all of the criteria outlined by Kline (1986) and thus can be considered a good quality assessment of empathy.

Chapter 4

A Comparison of Male Intimate Partner Violent, Violent and Non-Violent Offenders on Measures of Empathy

Abstract

The empirical study presented in this chapter explored empathic differences between male prisoners who had perpetrated intimate partner violence (IPV), violence (V) and non-violent offences (NV). Specifically it tested differences between self-reported empathy and facial emotion recognition between groups and explored the relationship between self-reported empathy and facial emotion recognition. 70 male offenders volunteered to complete the study: 30 IPV, 20 V and 20 NV. Two measures of empathy were included in the study: The Interpersonal Reactivity Index (IRI) questionnaire and The Diagnostic Analysis of Nonverbal Accuracy 2 – Adult Facial Expressions (DANVA2-AF) facial emotion recognition task. Results showed that IPV offenders had significantly lower empathic concern scores and were more likely to interpret fearful faces as sad than NV offenders. Measuring the correlation between IRI and DANVA2-AF scores, only personal distress (PD) showed a significant relationship. As a whole group, higher PD scores were related to more errors identifying angry and high intensity faces. Analysing groups separately, IPV offenders PD scores were positively correlated with angry and fearful error scores, V offenders PD scores were positively correlated with anger errors scores and NV offenders PD scores were positively correlated with angry, high and low intensity error scores. Findings suggest that there was preliminary evidence for differences in affective empathy between IPV and NV offenders and little evidence of differences between groups relating to cognitive empathy. The findings provide evidence against the supposition that IPV and V offenders have differing empathic profiles.

Introduction

Previous chapters have explored the role of empathy in family violence by systematically reviewing literature relating to empathy and child maltreatment and then by analysing the psychometric properties of the Interpersonal Reactivity Index (IRI; Davis, 1980), which is a commonly used multidimensional assessment of empathy. This chapter will continue to explore the role of empathy in family violence by detailing an empirical study that explored the role of empathy in incarcerated perpetrators of intimate partner violence (IPV) and compared these outcomes with incarcerated perpetrators of violent and non-violent crimes.

Within literature and forensic practice IPV is considered a special type of violent crime, distinct from general violent behaviour. Accordingly, theories of IPV have been developed independent from general theories of crime and violence (Felson, 2002; Gottfredson & Hirschi, 2007) which has influenced the distinction made for IPV perpetrators in practice. For example, within Her Majesty's Prison Service IPV offenders have a different treatment pathway to general violent offenders, with IPV treatment programmes such as the Healthy Relationships Programme (HRP; Bullock, Sarre, Tarling & Wilkinson, 2010). This is costly to the prison service, as HRP is not delivered at all prison sites thus IPV prisoners have to be moved prison for the purpose of treatment. This cost is justified if the intervention increases programme responsivity and reduces re-offending. However, there is little research that explores differences between IPV and violent offenders. With an awareness that IPV intervention programmes have highest drop-out rates (Holtzworth-Munroe & Meehan, 2004) and offenders who don't complete treatment are the most likely to re-offend (Hanson & Bussière, 1998), it is important that research focuses on increasing an understanding of treatment needs of IPV offenders and how this relates to the needs of violent offenders in general.

Recent research by Moffit, Krueger, Caspi and Fagan (2000) aimed to explore the similarities and differences between general violent and IPV offenders. They explored data from a sample of 800 young adults who had been part of longitudinal research from birth. The study looked at personality data that had been collected aged 18 and self-report and official offending information aged 21. Data were analysed using modelling latent constructs which identified that whilst the IPV and violent offences were distinct, they were "moderately related". Most IPV offenders had committed violence outside of their relationships, which is in line with previous research that indicates generally violent IPV perpetrators make up the largest majority of IPV offenders (Holtzworth-Munroe & Stuart, 1994). They also found that poor emotional regulation and hostile attribution bias, a trait they coined 'negative emotionality', was strongly linked to both IPV and violence. This would suggest IPV shares many similarities with violent offending, in line with the perspective of a general theory of violence (Felson, 2002). Hence, further exploring of this overlap is necessary to inform IPV treatment needs.

Empathy

The present study was interested in understanding whether empathic capacity, another criminogenic factor, differed between offenders. Empathic capacity is conceptualised as the ability to understand (cognitive empathy) and/or to experience (affective empathy) the emotions of another person (Davis, 1980). Therefore theorists suggest that in order to engage in violent behaviour an individual must either have empathic deficits (Eisenberg & Miller, 1987) or be able to dissociate from their empathic experiences (Abel et al., 1989). In line with this, empathy is a risk (Douglas & Reeves, 2010) and protective factor (de Vogel, de Ruiter, Bouman & de Vries Robbé, 2009) for violence and offender rehabilitation (Mulloy, Smiley & Mawson, 1999).

It is important to recognise however that empathy is not a trait considered unique to violent offending. One of the main principles of restorative justice literature and practice, which focuses on both violent and non-violent offending, is that by supporting a perpetrator to understand the impact of their behaviour on their victims they will increase their perspective taking and thus reduce likelihood of reoffending (Wallis, 2014). Zehr (2002) links the restorative justice process to the neural correlates of empathy, suggesting that by seeing and hearing the experiences of victims, offenders are able to identify and relate to victims which reduces their cognitive biases and increases responsibility taking.

The field of IPV similarly engages in the notion that perpetrators must have empathic deficits. Accordingly, intervention includes skills to improve empathic capacity with an aim to reduce recidivism (Bullock, Sarre, Tarling & Wilkinson, 2010). There has been relatively little research carried out to investigate empathy deficits in IPV perpetrators and such conclusions are therefore merely assumptions. This causes a number of difficulties. Primarily, there is a lack of clarification as to whether IPV perpetrators have empathy deficits. Secondly, if empathy deficits are present, it is unclear whether they are risk factors for IPV or whether increasing empathic capacity reduces recidivism. Thirdly, there is a lack of empirical testing regarding the overlap between violent and IPV behaviour and the risk/protective factors that make the behaviours distinct.

The absence of an empirically validated understanding of how empathy interacts with other individual and environmental factors to mobilise IPV is problematic for the reduction of IPV recidivism. In the absence of this understanding intervention for IPV is being designed and delivered without a robust theoretical foundation and therefore violating the 'need' principle of rehabilitation (Andrews & Bonta, 2003).

Empathy and General Violent Behaviour

Some understanding of how empathy relates to IPV has been deduced from research examining the relationship between empathy and general violence, which was described in Chapter 1. Traditionally empathy has been assessed using questionnaire-based methods. Limitations of using questionnaires has included participants having time to consider responses, thus not assessing 'bottom-up' dispositional empathy (Jenkins & Dillman, 1995). More recently, researchers interested in the role of empathy have used task-based assessment tools to overcome limitations of questionnaires, such as facial emotion recognition tasks. Findings suggest that self-reported affective empathy correlates positively with facial emotion recognition (Gery, Miljkovitch, Berthoz & Soussignan, 2009) with less understanding regarding the relationship between cognitive empathy and emotion recognition.

Relating this to offending populations, Robinson et al., (2012) found that offenders (both violent and non-violent) showed significant impairment in recognition of negative emotions (particularly anger, sadness, fear and disgust) compared with a community sample. However there was no significant difference in emotion recognition between violent and not violent offenders. Seidal et al. (2013) found that violent offenders demonstrated altered skin conductance responses for recognising faces of fear and disgust compared with a matched community sample. Furthermore, there was a significant negative correlation between number of violent assaults and accuracy in perspective taking when viewing angry scenarios.

Considering the amount of disparity across studies when utilising alternative methods of emotion recognition and empathy assessment, such as facial recognition and skin conductance, it is important to understand which components of empathy are being assessed, as the type of processing required for the task is likely to impact on outcomes. Emotions that

are only presented briefly to participants are more likely to use bottom-up processing and success on these tasks has been linked to increased emotional empathy (Martin et al., 1996). For example, Besille and Yuille (2010) found that when showing facial emotions at 50ms, the Interpersonal Reactivity Index (IRI; Davis, 1980) empathic concern subscale was significantly correlated with emotion recognition and the Empathy Quotient (Baron-Cohen & Wheelwright, 2004) was not. The reverse effect was seen at 2000ms. This study suggests that facial recognition is linked to empathy, but different empathic processes are linked to different stages of facial recognition.

Empathy and IPV

There are a few studies that have sought to explore emotional processing and empathic capacity in IPV offenders. One group of researchers have utilised video vignette paradigms to explore dispositional empathy in IPV perpetrators. The first of these studies (Schweinle, Ickes & Bernstein, 2002) recruited 86 married men from the community. Outcomes from Conflict Tactics Scale (Straus, Hamby, McCoy & Sugarman, 1996) responding indicated that 11 (12.8%) of these men had perpetrated mild to moderate physical violence towards their wives, although none had done so in the last year. Additionally, 72 (83.7%) reported verbal aggression only and three (3.5%) reported no aggression. Participants completed a series of questionnaires and were then presented with three videos of three separate women discussing their relationship difficulties. Each video was divided into 30 15-second clips (totalling 90 video-clips). After viewing each 15-second video-clip, the men were required to write down the thought or feeling that they believed the woman in the video was experiencing and to decide whether the thought or feeling was critical/rejecting, not critical/not rejection or ambiguous.

Analysis of the results found a weak positive (albeit non-significant) correlation between level of aggression perpetrated towards partners and propensity to misidentify affect as critical/rejecting or ambiguous. This over identification of critical/rejecting and ambiguous emotions was also found to significantly positively correlate with level of aggression, when relationship instability was controlled for. Authors suggested that the results provided evidence for a general empathy deficit towards women in IPV men, with IPV males misinterpreting negativity in non-threatening situations. However, as participants were not shown videos of men, it is difficult to determine whether IPV males had an empathy deficit towards females specifically or towards both males and females. Additionally, given the small sample of IPV perpetrators and low frequency and moderate intensity of their violence, the validity of these findings is questionable.

Subsequent to the publication of these findings, the researchers concurrently published two further studies adapting the original paradigm. Schweinle and Ickes (2007) recruited 80 married men and used one of the three video examples originally used to develop 30 shorter clips of 15-seconds. Additionally, participants were covertly recorded whilst watching the clips and then (after being informed that they had been recorded) were required to watch their own reactions and label their emotions. Results from the experiment found a significant positive moderate correlation between IPV and over-identification of critical/rejecting and ambiguous emotions when participants' own perception of marital criticism was controlled for. This over-identification bias was significantly negatively correlated with the accuracy of affect identification. Additionally, they found a positive relationship between level of concern for the female and ability to correctly identify her emotion. These results are consistent with their previous findings (Schweinle et al., 2002) and suggest that IPV perpetrators have a general empathy deficit with a propensity to misinterpret negativity in non-negative stimuli. It is difficult however to interpret these findings as researchers did not

provide information about how many participants reported perpetrating IPV or the level and frequency of IPV perpetrated.

Using a different design, Clements, Holtzworth-Munroe, Schweinle and Ickes (2007) examined 71 community couples. Couples were divided into three groups: those experiencing IPV, those not experiencing IPV but showing distress, those not experiencing IPV and not showing distress. Men were shown 10-minutes of two of the videos used in previous experiments and were required to write down what they believed the women were thinking and feeling. After watching these videos, couples were then videotaped discussing a pre-agreed relationship difficulty. Following this, partners viewed their recording and independently rated their own thoughts and affect and then their partners thought and affect. Analysis revealed that men in violent couples were significantly worse at identifying their partner's emotion than men in non-violent relationships who were not distressed. There was no significant difference between the three groups regarding male responses towards female strangers. There were no significant differences between the three groups regarding female participants' recognition of their partner's emotions. This finding of partner-specific empathy deficits in male IPV perpetrators conflicts with the researcher's previous above-described findings of a general empathy deficit in IPV males (Schweinle & Ickes, 2007; Sweinle, Ickes & Bernstein, 2001). This would provide support for Marshall and Marshall's (2011) theory of empathy in violent offenders, explaining that the relationship an offender has with a subject modulates empathic capacity.

The three studies described are limited by the lack of clarity regarding the nature of IPV amongst their participants as well as some design flaws such as the absence of counter balancing. Some research however has used alternative paradigms to provide further insight into the relationship between IPV and empathy. Romero-Martínez, Lila, Sariñana-González, González-Bono and Moya-Albiol (2013) examined empathy as part of a study which aimed

to explore hormone dysfunction in IPV perpetrators when exposed to stress. Their study suggested that IPV perpetrators had higher levels of testosterone at baseline and differing testosterone and cortisol responding to controls in response to stress. As measured by the IRI, IPV perpetrators had significantly lower perspective taking scores and significantly higher personal distress scores than controls. There was not a significant difference between empathic concern and fantasy scores.

Additionally, Covell, Huss and Langhinrichsen-Rohling (2007) administered the IRI and the Revised Conflict Tactics Scale 2 (CTS2; Straus, Hamby, McCoy & Sugarman, 1996) with 104 IPV perpetrators who were undergoing community intervention. The study focused on the psychological aggression, physical assault and coercion scales on the CTS2 and accumulated these scores to make a 'total violence' score. Similar to findings from the general violence literature, IPV perpetrators who most frequently engaged in psychological aggression and total violence had lower perspective taking scores. Additionally, those who reported higher personal distress had higher total violence and physical assault scores. There were no significant relationships found between the fantasy or empathic concern subscales.

Furthermore, Covell et al. (2007) found that IPV offenders who used sexual violence had a differing empathic capacity pattern to IPV offenders who were physically violent only. They also identified five distinct IPV profiles based upon CTS2 outcomes. Those who scored highly on physical assault items had low perspective taking scores, high personal distress scores but high empathic concern scores. Those reporting high levels of psychological aggression had low fantasy and perspective taking scores but high personal distress scores. Those reporting high levels of overall violence could be subdivided into two groups. The first group reported higher perspective taking and personal distress scores and low fantasy scores. The second group had low perspective taking, fantasy and personal distress scores. Males who were sexually abusive to their partners had high personal distress and perspective-taking

scores. This research suggests that empathic deficits may be different amongst IPV perpetrators, depending upon the type of violence that they perpetrate.

Similarly, Babcock, Green and Webb (2008) explored empathic deficits amongst a community sample of 110 males. Both male participants and their partners completed the CTS2 and the males completed further questionnaires to assess their personality patterns. Participants also completed the Picture of Facial Affect series (Ekman & Friesen, 1976) which required them to identify the emotion expressed in 60 facial images. Out of the sample, 32 (29%) reported that they had never perpetrated any form of violence against their partners. Analysing their results based upon Holtzworth-Munroe and Stuart's (1994) three typologies of IPV perpetrators, Babock et al. (2008) found that emotion recognition deficits were not present in all IPV perpetrators. Interestingly, there were no significant differences found between IPV perpetrators and non-violent males. Those perpetrators identified as having borderline traits performed well on the emotion recognition task. Perpetrators identified as violent in and outside of their relationships obtained significantly more total errors on the emotion recognition task than the other groups. These males were likely to misidentify happy faces as sad or surprised and surprised faces as fear. They were also more likely to misidentify neutral and anger as disgust. IPV perpetrators who were only violent in relationships showed no empathic deficits.

The findings from this study would be consistent with theories suggesting that some perpetrators have a hostile attributional bias (Byrne & Arias, 1997), thus their violence is enacted in response to a misidentified negative provocation. This is also consistent with the findings of emotion specific deficits in violent offenders (Robinson et al., 2012) and IPV offenders (Schweinle & Ickes, 2007; Schweinle, Ickes & Bernstein, 2002).

It is also important to acknowledge however that offenders rarely specialise in one offence type. In 2012 85% of convicted offenders had a previous offence (MoJ, 2012). The latest figures from the MoJ (2014) suggest that 102,600 adult offenders had at least 15 prior convictions and 40% of them had long criminal records (on average 33 previous convictions). To provide an example of criminal diversity, of the 1,222 people convicted of rape in 2014, 85.3% had a previous conviction. Of the previous convictions, 4% were for rape, 11% for other sexual offences, 24% for other violent offences and 61% for 'other' offences (MoJ, 2015). Obtaining these statistics for IPV offenders is difficult as IPV is not recorded as an offence type, rather as a violent offence. It is important however when analysing IPV research to consider that the sample is likely to contain a variety of additional offence behaviours and offender profiles.

Aims of the Current Study

Reviewing outcomes from the limited available research reveals that all studies to date have been conducted with community based male perpetrators. There appears to be an indication that IPV perpetrators have empathy deficits and hostile attribution bias. However, it is unknown whether such deficits exist in clinical samples. There is evidence to suggest that using questionnaire and facial emotion recognition tasks to measure empathy provides a more comprehensive assessment amongst offenders. However to date, there has not been a study that has used questionnaire and facial emotion recognition tasks to investigate whether IPV perpetrators have an empathic profile that is unique from generally violent and non-violent offenders. The current study aims to extend the work completed by Babock et al. (2008) by examining self-reported empathy and facial emotion recognition in three participant groups of incarcerated males: IPV (IPV), violent (V) and non-violent (NV) perpetrators.

Specifically, three research questions were investigated:

1. Are there significant differences in self-reported empathy between offenders who have perpetrated intimate partner violence (IPV), offenders who have perpetrated violence (V) and offenders who have never perpetrated violence (NV)?
2. Are there significant differences in facial emotion recognition between offenders who have perpetrated intimate partner violence (IPV), offenders who have perpetrated violence (V) and offenders who have never perpetrated violence (IPV)?
3. Is there a relationship between self-reported empathy and facial emotion recognition?

Method

Ethical Approval

The current study adheres to the ethical guidelines stipulated by the University of Birmingham and the British Psychological Society. Ethical approval was received from both the University of Birmingham (ERN_1-0088) and the University of Coventry and the study was logged with the national Integrative Research Application System (#66225/163353/8/845). Additionally, local approval was received from the prison establishment where data collection took place.

Participants

Participants were recruited during a five week period between March and April 2014. In order to inform the number of participants needed for the study, a priori power analysis was conducted. Assuming an effect size of 0.15 with four response variables and three groups ($\alpha = 0.05$, power = 0.80), a total sample of 57, with 19 in each group, was required.

In total, 70 convicted male offenders participated in the study: 30 IPV perpetrators (IPV), 20 violence perpetrators (V), 20 non-violent offenders (NV). All of the participants

were serving a custodial sentence at the same prison, which had both closed (Category B and C) and open (Category D) conditions. Different times and days were used in order to recruit a representative sample and to prevent prison employment and leave impacting upon participation. The average population of the prison was [redacted] in the recruitment period (MOJ, 2014) and therefore the sample represented 7% of the population. Table 12 provides descriptive information regarding the participant demographics.

The age of participants ranged between 22 and 66 years (M = 35, SD = 9.6), which is reflective of the overall prison population (Berman & Dar, 2013). There was not a significant difference between the ages of IPV (M = 34.04, SD = 8.1) and V (M=30.2, SD = 7.2) offenders or IPV and NV (M=39.5, SD = 41) offenders. However there was a significant difference between the ages of V and NV participants ($F(2,67) = 5.374$, $p < 0.05$).

Regarding ethnicity, 64.29% of the sample identified themselves as White-British, 17.14% as Asian and 15.71% as Afro-Caribbean. The latest statistics suggest that a quarter of the UK prison population are from an ethnic minority group (Berman & Dar, 2013), therefore the sample has a slight under-representation of white-British prisoners. This may reflect that the region has a high level of multiculturalism (58% of residents White-British) compared with the national average (87% of UK residents White-British) (Office for National Statistics, 2011). There was not a significant difference between the ethnicity of prisoners in each group x^2 (1, N = 70) = 8.92, P = 0.35.

Table 12.
Descriptive information of the research sample

	IPV (N = 30)		Violent (N = 20)		Non-Violent (N = 20)	
	N	%	N	%	N	%
Ethnicity						
White - British	20	66.7	15	75	10	50
Afro - Caribbean	5	16.7	3	15	3	20
Asian	3	10	2	10	7	30
Other	2	6.6	0	0	0	0
Index Offence						
Murder/manslaughter	1	3.1	1	4.8	0	0
Violence (including kidnap, wounding, GBH, aggravated burglary)	9	28.1	5	23.8	0	0
Robbery/armed robbery	1	3.1	2	9.5	0	0
Possession of weapon	2	6.3	1	4.8	0	0
Arson	1	3.1	1	4.8	0	0
Theft or burglary not from person	7	21.9	3	14.3	3	15
Criminal Damage	0	0	1	4.8	0	0
Fraud/money laundering/cheating public revenue	2	6.3	3	14.3	13	65
Perverting course of justice	0	0	0	0	1	5
Handling stolen goods	1	3.1	0	0	0	0
Breaching restraining order	3	9.4	0	0	0	0
Supply, import or possessions of drugs	5	15.6	4	19	3	15
Previous Offences						
Violence (including kidnap, wounding, GBH, aggravated burglary)	12	36.4	7	25	0	0
Robbery/armed robbery	1	3	0	0	0	0
Possession of weapon	1	3	2	7.1	0	0
Escape from custody	0	0	1	3.6	0	0
Theft or burglary not from person	1	3	6	21.4	0	0
Criminal Damage	2	6.1	3	10.7	0	0
Fraud/money laundering/cheating public revenue	1	3	0	0	2	10
Driving offence	4	12.1	1	3.6	2	10
Supply, import or possessions of drugs	2	6.1	3	10.7	0	0
No Previous Convictions	9	27.3	5	17.9	16	80

Measures

The Revised Conflict Tactics Scale (CTS2; Straus, Hamby, McCoy & Sugarman, 1996)

The CTS2 is a measure of aggression used in relationships. The full CTS2 consists of five subscales to assess conflict tactics used in relationships: negotiation, psychological aggression, physical assault, sexual coercion and injury. Each item describes a behaviour and respondents are required to state whether they have engaged in the behaviour. Within each subscale, both 'minor' and 'severe' acts of the behaviour are included.

In the current study, the CTS2 was used as a screening tool in order to assist in assigning participants into the correct group. Therefore, only the sexual coercion and physical assault scales were used. These have been demonstrated to have internal consistencies of 0.86 and 0.87 respectively (Straus, Hamby, Boney-McCoy & Sugarman, 1996). In the current study, internal consistencies of 0.83 and 0.81 were found respectively. Additionally, one item from the 'psychological aggression' scale was used: 'threatened to hit or throw something at a partner'.

Accordingly, the questionnaire given to participants consisted of 20 items. The questionnaire was adapted so that participants were required to respond whether they had engaged in each behaviour towards: a stranger, an acquaintance, an intimate partner and another family member. This was in order to screen for violent behaviour in addition to IPV.

The Balanced Inventory of Desirable Responding (BIDR; Paulhaus, 1998)

The BIDR is a 40-item questionnaire which measures social desirability. Each item presents a statement and respondents are required to rate how true the statement is on a 7-point scale from 'not true' to 'very true'. There are two subscales:

- Self-Deceptive Enhancement (SDE): Tendency to provide honest but positively over-emphasised description of self
- Impression Management (IM): The intentional tendency to provide a deceptive self-description.

Previous research has shown that offenders often engage in higher self-deception than controls (Day, Mohr, Howells, Gerace & Lim, 2012), therefore the BIDR was included to asses this. Dichotomous scoring procedure was used which involves only taking account for the 'extreme' scores. This is a commonly used method and produces internal consistencies of 0.65 – 0.8 (Paulhaus, 1994). In the present study, the SDE had an internal consistency of 0.72 and the IM had an internal consistency of 0.76.

The Interpersonal Reactivity Index (IRI; Davis, 1980)

As seen in Chapter 3, the IRI is a 28-item questionnaire designed to assess multiple aspects of empathy. Each item presents a statement and respondents are required to rate how well the statement describes them on a 5-point scale from 'not well' to 'very well'. The IRI is comprised of four subscales with seven items each. The subscales are:

- Perspective-taking (PT): the ability to understand the perspective of another person
- Empathic concern (EC): the ability to experience compassion for another person experiencing negativity
- Personal distress (PD): the extent to which an individual experiences distress when witnessing someone else's misfortune
- Fantasy (FS): the tendency to identify with fictional characters

The IRI was used as a self-report assessment of empathic capacity. The current study was interested in exploring both cognitive and affective empathy, thus the IRI was considered the most appropriate measure. The validity of the IRI as a measure of empathy has been

demonstrated, with internal consistencies between 0.70 – 0.78 produced (Christopher, Owens & Stecker, 1993; Davis, 1994). The internal reliabilities in the present study were: 0.78 for PD, 0.72 for EC, 0.76 for PD, and 0.81 for FS. Therefore the IRI was considered an effective measure to assess self-reported empathic capacity.

Chapter 3 discussed findings that suggested removing negatively worded items from the IRI scales increased its reliability. The methodological issues of this procedure were discussed in Chapter 3 and accordingly the decision was made to use the full IRI questionnaire. Analysis of outcomes did explore whether removing the negatively worded items impacted upon findings and scale reliability. This was not the case and thus the findings discussed in this chapter relate the full IRI questionnaire.

The Diagnostic Analysis of Nonverbal Accuracy 2 – Adult Facial Expressions (DANVA2-AF; Nowicki, 2010)

The DANVA2-AF consists of 24 adult faces expressing happy, sad, angry and fearful emotions (six of each emotion). Each emotion consists of three high-intensity and three low-intensity emotional expressions. Each emotion is presented individually at 2000ms and participants are required to choose whether the emotion expressed is happy, sad, angry or fearful.

Previous research has found that higher error scores on the DANVA2-AF to be correlated with higher frequency of conflict (Verbeek, 1996), higher levels of depression (Radloff, 1977) and poor emotional control (Mcintyre, Danforth & Schneider, 1997, as cited in Nowicki, 2010). The DANVA2-AF has been found to have internal consistencies of 0.78 to 0.90 and test-retest reliabilities of 0.81 – 0.91 (Baum, Logan, Walker, Tomlinson & Schiffman, 1996; McIntyre, Danforth & Schneider, 1997; Spell, 1996). In the present study, internal consistencies were: 0.83 for happy, 0.76 for sad, 0.71 for angry, 0.71 for fearful, 0.72

for high intensity and 0.71 for low intensity. Given the associations with conflict and affective functioning and strong construct validity, the DANVA2-AF was considered an appropriate measure of facial affect recognition for the current investigation.

Procedure

Offenders were recruited within the prison opportunistically. The researcher recruited prisoners from one wing and three workshop locations within the closed prison and a category D prison. Suitable participants were identified by searching through prisoner files and computer systems. Those deemed suitable based upon file information (see Table 13 for inclusion criteria) were sent a copy of the information sheet via internal prison postal system, which they could keep. The sheet contained details of the researcher who they could contact if they wanted to take part in the study. When reviewing files, most prisoners fit into one of the three offender groups in the study. The only exclusion criteria was if offenders had a sexual offence. This is because research has suggested that offenders who sexually assault their partners have a different empathic profile to offenders who are otherwise violent in relationships (Covell et al., 2007). Given the relatively small sample size, we did not want this to confound the findings. One prisoner was excluded post-participation based upon this criteria. Offenders and prison staff were not told that the study was specifically exploring IPV, but rather just violence, in order to safeguard participants.

If offenders agreed to take part they were taken to a room containing the research materials. Offenders first read the participant information sheet (Appendix 10) and completed the consent form (Appendix 11). The questionnaires and computer task were randomly allocated to prevent order effects. All of the tasks were completed in the presence of the researcher in order to ensure that all of the items were completed and to answer any questions.

When completing the DANVA2-AF tasks, participants sat in front of the laptop and the researcher sat to the side. The researcher controlled the mouse of the laptop in order to ensure that all of the items were answered and prevent the ability of the prisoner to use the computer causing distraction.

Although prisoners were provisionally categorised prior to participation, definitive allocation to offender groups was made following their participation. Criteria for each group is found in Table 13. A number of steps were taken in order to correctly allocate prisoners to groups. When completing pre-participation file reviews, information stored on the prisons electronic information system (PNOMIS) and risk assessments (OASys) were analysed. Post-participation, offenders CTS2 results were analysed to assess for incidents of violence and IPV that prisoner's had not been convicted for and file reviews were re-reviewed to ascertain whether there was undisclosed incidents of violence and IPV.

Table 13.
Inclusion/exclusion criteria for each participant group

	IPV	V	NV
Inclusion	Convicted of a violent offence against partner Self-report of violence against partner in CTS2 responding Prison documentation confirms that prisoner has perpetrated IPV Offender has not committed a sexual offence	Convicted of a violent offence (not against partner) No self-reported violence against partner in CTS2 responding Prison documentation confirms that prisoner has perpetrated violence (not against partner) Offender has not committed a sexual offence	Convicted of non-violent offence No self-report of any violence and aggression Prison documentation shows no history of violent offending Offender has not committed a sexual offence

Analysis

Research Question 1: Are there significant differences in self-reported empathy between offenders who have perpetrated intimate partner violence (IPV), offenders who have perpetrated violence (V) and offenders who have never perpetrated violence (NV)?

Kolmogorov-Smirnov was used to test for normal distribution and Levene's Test was used to assess whether the assumption of homogeneity was met. Box's M Test was used to measure whether the correlation between covariates were significantly different. A one-way multivariate analysis of covariance (MANCOVA) was used to explore whether there were significant differences in empathy between IPV, V and NV offenders. The offender category (IPV, V, NV) was the independent variable and the four subscales of the IRI were the dependent variables. The total IRI score was not used in analysis as the four scores do not positively correlate thus the 'total score' does not provide a reliable measure of empathy (D'Orazio, 2004). The two subscales of the BIDR were inputted as covariates in order to control for impression management and self-deception enhancement. The procedure provided by Saunders (1991) was followed, conducting both a MANOVA and a MANCOVA to observe and compare the effect of the covariates. MANCOVA's Wilks' Lambda was used as there were three groups. Bonferroni was used to compute a post-hoc analysis.

Research Question 2: Are there significant differences in facial emotion recognition between offenders who have perpetrated intimate partner violence (IPV), offenders who have perpetrated violence (V) and offenders who have never perpetrated violence (IPV)?

Kolmogorov-Smirnov was used to test for normal distribution and Levene's Test was used to assess whether the assumption of homogeneity was met. Box's M Test was used to measure whether the correlation between covariates were significantly different. An

ANCOVA was computed to compare total DANVA2-AF scores between the groups, whilst controlling for desirable responding. This was analysed separately as the score correlates with the DANVA2-AF emotion and intensity variables.

MANCOVA's explored whether there were significant differences in emotion and intensity between IPV, NV and V offenders, as measured using the DANVA2-AF. Two separate MANCOVA's were run. Offender category (IPV, V, NV) was the independent variable and the DANVA2-AF scores were the dependent variables (DANVA2-AF emotion and DANVA2-AF intensity). A third MANCOVA was computed to explore whether there were any significant interactions between groups regarding the error rate of emotion identification (for example, interpreting a happy face as a sad face). The two subscales of the BIDR were inputted as covariates in order to control for impression management and self-deception enhancement. Again, the procedure provided by Saunders (1991) was followed, conducting both a MANOVA and a MANCOVA to observe and compare the effect of the covariates. MANCOVA's Wilks' Lambda was used as there were three groups.

Research Question 3: Is there a relationship between self-reported empathy and facial emotion recognition?

This research question was explored by looking at the correlation between IRI scores and DANVA2-AF scores for the whole participant group and individual participant groups. Pearson's Partial Correlation was used in order to control for desirable responding covariates.

Results

Research Question 1: Are there significant differences in self-reported empathy between offenders who have perpetrated intimate partner violence (IPV), offenders who have perpetrated violence (V) and offenders who have never perpetrated violence (NV)?

All variables were normally distributed, the assumption of homogeneity was met and the correlation between the covariates did not differ significantly. MANCOVA results are presented as for all analyses there was no difference between MANOVA and MANCOVA.

Analysis revealed there was a non-significant multivariate outcome (Wilks Lambda = 0.839, $F(8, 124) = 1.422$, $p = 0.194$) indicating that there was no significant difference between three offender groups on IRI output. Post-hoc analysis however revealed that that IPV offenders (M=16.64, SE = 0.95) had significantly lower EC scores than NV offenders (M=20.55, SE = 1.19) when desirable responding was controlled for. The effect size of this interaction however was small. Outcomes are provided in Table 14.

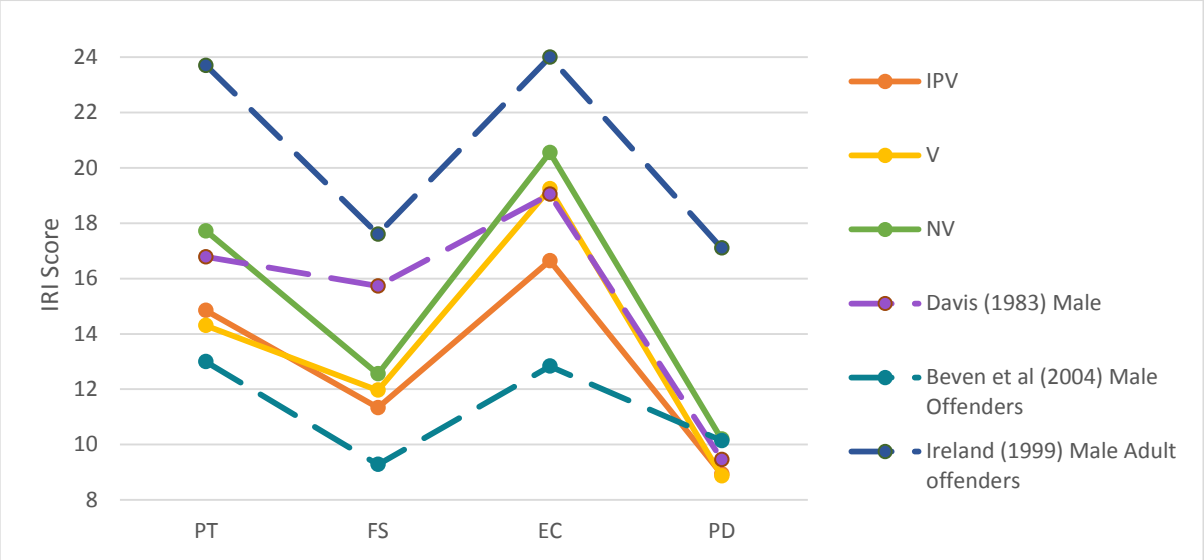
Table 14.
MANCOVA output for IRI and DANVA2-AF variables

	IPV	V	NV	MANCOVA				IPV/V	IPV/NV	V/NV
	Mean (SE)	Mean (SE)	Mean (SE)	*F*	*P*	η_p^2	OP	*P*	*P*	*p*
IRI				1.42	0.19	0.08	0.62			
PT	14.84 (1.01)	14.30 (1.26)	17.72 (1.26)	2.15	0.12	0.06	0.43	1.0	0.24	0.19
FS	11.33 (0.95)	11.96 (1.18)	12.55 (1.19)	0.33	0.72	0.01	0.10	1.0	1.0	1.0
EC	16.64 (0.95)	19.24 (1.19)	20.55 (1.19)	3.62	0.03	0.1	0.65	0.28	0.038*	1.0
PD	8.93 (0.84)	8.86 (1.05)	10.19 (1.05)	0.53	0.59	0.02	0.13	1.0	1.0	1.0
DANVA2-AF EMOTION				0.75	0.61	0.05	0.36			
Happy	0.57 (0.12)	0.33 (0.14)	0.42 (0.44)	0.92	0.41	0.03	0.20	0.58	1.0	1.0
Sad	1.73 (0.23)	1.60 (0.29)	1.61 (0.29)	0.09	0.91	0.00	0.36	1.0	1.0	1.0
Angry	3.20 (0.23)	2.92 (0.29)	2.52 (0.29)	1.67	0.30	0.05	0.34	1.0	0.22	1.0
Fearful	2.97 (0.27)	2.58 (0.33)	2.16 (0.33)	1.84	0.17	0.05	0.37	1.0	0.18	1.0
DANVA2-AF INTENSITY				0.93	0.3	0.04	0.38			
High	3.67 (0.34)	2.98 (0.42)	2.92 (0.42)	1.31	0.28	0.04	0.27	0.61	0.51	1.0
Low	4.80 (0.30)	4.46 (0.38)	3.83 (0.38)	1.98	0.15	0.06	0.4	1.0	0.15	0.79

Note. MANCOVA's Wilks' Lambda was used. All scores have been adjusted for covariates SDE and IM. SE = Standard Error; OP = Observed Power, IPV = Intimate Partner Violent offenders; V = Violent offenders; NV = Non-violent offenders; IRI = Interpersonal Reactivity Index; PT = Perspective Taking Subscale; FS = Fantasy Subscale; EC = Empathic Concern Subscale; PD = Personal Distress Subscale; * = significant result.

In order to understand how well the findings from the present study compared with normative data, outcomes from the IRI were compared with normative data presented in Chapter 3 (Graph 1).

Graph 1. IRI mean outcomes



The graph highlights the variability of normative data for both general and forensic populations, particularly for the fantasy scale whereby the all studies showed different mean scores and the three samples from the present study showing most similarity. Both Davis's (1983) community sample and Beven et al.'s (2004) forensic sample showed similar PD scores to the three offender groups in the present study. On the PT scale the IPV and V offender groups had scores higher than Beven et al.'s (2004) offender sample and lower than Davis's (1983) community sample. On the EC scale, all three samples from the present study had scores that most closely matched Davis's (1980) community sample.

Research Question 2: Are there significant differences in facial emotion recognition between offenders who have perpetrated intimate partner violence (IPV), offenders who have perpetrated violence (V) and offenders who have never perpetrated violence (IPV)?

All variables were normally distributed, the assumption of homogeneity was met and Box's M Test demonstrated that the correlation between the covariates did not differ significantly. MANCOVA results are presented here as for most analyses there was no

difference between MANOVA and MANCOVA. Where there was a difference, specific reference to the MANOVA has been made.

When comparing DANVA2-AF scores between groups, the ANCOVA found a non-significant multivariate effect, ($F(2, 67) = 1.543$, $p = 0.221$) indicating that there were no significant differences between groups on the total facial recognition error scores. Running two separate MANOVA's, to explore whether there were significant differences in emotion between groups, analysis revealed there was a non-significant multivariate outcome (Wilks Lambda = 0.905, $F(8, 124) = 0.794$, $p = 0.609$).

There was a non-significant multivariate outcome (Wilks Lambda = 0.585, $F(34, 98) = 0.886$, $p = 0.65$) between groups on DANVA2-AF scores. Post-hoc analysis (using Bonferroni) revealed no significant differences between groups, although the difference between IPV (m = 1.2) and NV (m = 0.5) offenders for incorrectly interpreting Fear as Sad came close to significance ($p = 0.052$). When conducting a MANOVA this interaction was significant ($p = 0.037$).

Research Question 3: Is there a relationship between self-reported empathy and facial emotion recognition?

The correlation between IRI scores and DANVA2-AF scores are detailed in Table 15 for the whole participant group and Table's 16, 17, and 18 for individual participant groups. The majority of correlations were not significant, for any of the participant groups. Of the IRI subscales, PD was the only scale that showed significant correlation with DANVA2-AF variables. When all participants were grouped together, there was a moderate positive correlation between PD and recognition of angry faces. That is the higher the level of personal distress, the higher the likelihood that the prisoner incorrectly recognised angry facial expressions. There was also a weak positive correlation between PD and recognition of

high intensity faces. That is the higher the level of personal distress, the higher the likelihood that the prisoner incorrectly identified faces with an emotion of high intensity.

When correlations were analysed for each offender group separately, different patterns emerged. The V group showed a positive strong correlation between PD and recognition of sad faces, demonstrating that the higher the level of distress the more errors they made on identifying sad faces. Both IPV and NV groups showed strong positive correlations between PD and recognition of angry faces, demonstrating that recognition of angry faces became less accurate with higher levels of PD. The IPV group showed strong correlation between PD and identification of fearful faces, indicating that the higher the level of personal distress, the more likely IPV offenders were to make errors identifying fearful faces. Interestingly, the NV group were the only offenders who showed strong positive correlation between PD and high and low intensity emotions.

Table 15.
Pearson's Partial correlation between IRI and DANVA2-AF output variables for whole participant group (N=70)

	Happy	Sad	Angry	Fearful	High Intensity	Low Intensity
PT	r = -0.17 (p = 0.18)	r = -0.11 (p = 0.39)	r = -0.06 (p = 0.64)	r = -0.33 (p = 0.79)	r = -0.16 (p = 0.20)	r = -0.01 (p = 0.96)
FS	r = -0.19 (p = 0.12)	r = -0.18 (p = 0.12)	r = 0.01 (p = 0.97)	r = -0.04 (p = 0.77)	r = -0.09 (p= 0.50)	r = -0.05 (p = 0.70)
EC	r = -0.04 (p = 0.74)	r = -0.02 (p = 0.9)	r = -0.19 (p = 0.34)	r = -0.14 (p = 0.26)	r = -0.14 (p = 0.26)	r = -0.10 (p = 0.40)
PD	r = 0.17 (p = 0.18)	r = 0.05 (p = 0.71)	r = 0.32 (p = 0.01)*	r = 0.19 (p = 0.12)	r = 0.23 (p = 0.05)*	r = -0.08 (p = 0.53)

Note. All scores have been adjusted for covariates SDE and IM. PT = Perspective Taking Subscale; FS = Fantasy Subscale; EC = Empathic Concern Subscale; PD = Personal Distress Subscale. * = significant correlation

Table 16.
Pearson's Partial correlation between IRI and DANVA2-AF output variables for IPV group (N=30)

	Happy	Sad	Angry	Fearful	High Intensity	Low Intensity
PT	r = -0.27	r = -0.158	r = -0.118	r = 0.09	r = -0.19	r = -0.68
	(*p* = 0.17)	(*p* = 0.42)	(*p* = 0.55)	(*p* = 0.66)	(*p* = 0.32)	(*p* = 0.73)
FS	r = -0.16	r = -0.355	r = 0.26	r = 0.16	r = -0.20	r = 0.15
	(*p* = 0.41)	(*p* = 0.06)	(*p* = 0.21)	(*p* = 0.41)	(*p*= 0.31)	(*p* = 0.44)
EC	r = 0.13	r = 0.18	r = 0.09	r = 0.06	r = 0.15	r = -0.18
	(*p* = 0.49)	(*p* = 0.35)	(*p* = 0.67)	(*p* = 0.75)	(*p* = 0.46)	(*p* = 0.37)
PD	r = 0.61	r = -.019	r = 0.63	r = 0.40	r = 0.31	r = 0.35
	(*p* = 0.76)	(*p* = 0.34)	(*p* = 0.00)*	(*p* = 0.04)*	(*p* = 0.11)	(*p* = 0.07)

Note. All scores have been adjusted for covariates SDE and IM. PT = Perspective Taking Subscale; FS = Fantasy Subscale; EC = Empathic Concern Subscale; PD = Personal Distress Subscale. * = significant correlation

Table 17.
Pearson's Partial correlation between IRI and DANVA2-AF output variables for V group (N=20)

	Happy	Sad	Angry	Fearful	High Intensity	Low Intensity
PT	r = 0.02	r = 0.20	r = 0.04	r = -0.01	r = -0.5	r = 0.28
	(*p* = 0.95)	(*p* = 0.42)	(*p* = 0.89)	(*p* = 0.99)	(*p* = 0.86)	(*p* = 0.27)
FS	r = -0.4	r = 0.41	r = -0.14	r = -0.20	r = 0.03	r = 0.02
	(*p* = 0.88)	(*p* = 0.09)	(*p* = 0.59)	(*p* = 0.41)	(*p*= 0.90)	(*p* = 0.93)
EC	r = 0.09	r = -0.11	r = -0.34	r = -0.09	r = -0.27	r = 0.01
	(*p* = 0.73)	(*p* = 0.66)	(*p* = 0.17)	(*p* = 0.73)	(*p* = 0.29)	(*p* = 0.99)
PD	r = 0.24	r = 0.64	r = 0.33	r = -0.09	r = 0.16	r = 0.18
	(*p* = 0.4)	(*p* = 0.01)*	(*p* = 0.18)	(*p* = 0.73)	(*p* = 0.65)	(*p* = 0.47)

Note. All scores have been adjusted for covariates SDE and IM. PT = Perspective Taking Subscale; FS = Fantasy Subscale; EC = Empathic Concern Subscale; PD = Personal Distress Subscale. * = significant correlation

Table 18.
Pearson's Partial correlation between IRI and DANVA2-AF output variables for NV group (N=20)

	Happy	Sad	Angry	Fearful	High Intensity	Low Intensity
PT	r = -0.05 (p = 0.85)	r = -0.10 (p = 0.67)	r = 0.05 (p = 0.85)	r = 0.09 (p = 0.72)	r = -0.11 (p = 0.66)	r = 0.15 (p = 0.56)
FS	r = -0.32 (p = 0.19)	r = 0.17 (p = 0.50)	r = -0.30 (p = 0.24)	r = 0.05 (p = 0.84)	r = 0.12 (p= 0.65)	r = -0.36 (p = 0.14)
EC	r = -0.44 (p = 0.70)	r = -0.22 (p = 0.37)	r = -0.14 (p = 0.59)	r = -0.17 (p = 0.49)	r = -0.25 (p = 0.32)	r = -0.26 (p = 0.31)
PD	r = 0.25 (p = 0.31)	r = 0.01 (p = 0.99)	r = 0.77 (p = 0.00)*	r = 0.42 (p = 0.08)	r = 0.60 (p = 0.01)*	r = 0.52 (p = 0.2)*

Note. All scores have been adjusted for covariates SDE and IM. PT = Perspective Taking Subscale; FS = Fantasy Subscale; EC = Empathic Concern Subscale; PD = Personal Distress Subscale. * = significant correlation

Discussion

The aim of the present study was to explore whether there were significant differences in empathy and emotion recognition between IPV, V and NV offenders. The rationale was to explore whether empathy plays a role in the aetiology of IPV and to contribute to an understanding of whether men who are violent towards their partners have a differing empathic profile to men who are violent outside of their relationship. Additionally, the research was interested in understanding the relationship between empathy and facial recognition in an offending sample.

Summary of Results

Research Question 1: Are there significant differences in self-reported empathy between offenders who have perpetrated intimate partner violence (IPV), offenders who have perpetrated violence (V) and offenders who have never perpetrated violence (NV)?

Results from the present study indicated that there were no significant differences between groups and effect sizes were small. It is worth noting however that IPV offenders

were less likely to experience compassion when others experienced adversity than NV offenders (as measured by EC). The similarities in empathic profiles between IPV, V and NV offenders would provide support for general theories of crime and violence (Felson, 2002; Gottfredson & Hirschi, 2007) that suggest IPV shares similar risk and protective factors as general violence and antisocial behaviour.

Relating these findings to previous literature exploring empathy in IPV males, Romero-Martínez, Lila, Sariñana-González, González-Bono and Moya-Albiol (2013) found that IPV perpetrators had higher PD and lower PT scores than community controls. Considering why differences in results may have occurred, it is important to note that previous studies have compared IPV males with community non-violent males and the present study compared empathy between offender subtypes. Therefore, it is not surprising that outcomes were different as the present study looked at what empathic structures distinguish IPV offenders from other offenders, rather than what distinguishes IPV offenders from non-offenders.

It is important to note that this does not mean that compared to non-offenders, offenders do not have empathic deficits, as this was not measured in the present study. Rather, this would suggest that if empathy deficits do exist in offenders, IPV and V offenders have similar intervention needs. These findings are therefore in line with research by Moffit, Krueger, Caspi and Fagan (2000) which suggest that whilst IPV and V offenders present different constructs of offending, they share a number of similar criminogenic needs. Given the small sample and effect sizes, further research is required to understand the robustness of these findings

Research Question 2: Are there significant differences in recognition of emotional facial expressions between offenders who have perpetrated intimate partner violence (IPV), offenders who have perpetrated violence (V) and offenders who have never perpetrated violence (IPV)?

Findings from the present study indicated that there were no significant differences in facial emotion recognition between groups. However, IPV offenders were more likely than NV offenders to interpret fearful faces as sad. Interestingly, a similar finding of a fearful emotion deficit in IPV offenders was found by Marsh and Blair (2008). Other studies have found that compared to community controls, offenders had significant impairment recognising negative emotions. However there was no difference between offender subtypes (Robinson et al., 2012). As the present study did not compare offender and non-offender groups, comparison with prior literature is limited as it is unclear whether all of the offenders in the present study had impaired facial recognition or not.

The results from the present study may be explained in line with Marshall and Marshall's (2011) theory of empathy in violent offenders. As stipulated in their three-stage model of empathy, if an offender has a neutral relationship with a subject they are able to empathise with the individual. This would suggest that in the present study, as the faces used were not known to the offender, empathic deficits were not present. This would also explain findings by Clements, Holtzworth-Munroe, Schweinle and Ickes (2007) of partner-specific empathy deficits in IPV males. It would be interesting for future research to examine emotion recognition in known faces amongst IPV, V and NV offenders.

Linking the findings of lower affective empathy and increased tendency to misinterpret fearful faces as sad, it may be that IPV offender's deficit in experiencing negative emotions of others (as measured by EC) underpins their deficit in recognising fear in others.

Considering how this may impact upon their IPV behaviour, if IPV offenders have a deficit in recognising fear and experiencing the negative emotions of their victims, they are less likely to understand the perspective of their victims and the impact of their violent behaviour. If offenders understood that their victims were fearful, they would understand that this was a direct impact of their behaviour. However, interpreting the emotion as sad may lead to perpetrators attributing blame to the victim (e.g. for being weak) which may be more aggravating for IPV offenders. Previous research has shown that a victims expression of negative emotions is likely to increase the level of violence perpetrated (Perry & Perry, 1974). This theory would also explain findings by Schweinle, Ickes and Bernstein (2002) and Schweinle and Ickes (2007) that IPV males were more likely to misidentify faces as critical/rejecting. Thus, such deficits may underpin a hostile attribution bias and may act as an IPV disinhibitor.

Research Question 3: Is there a relationship between self-reported empathy and recognition of emotional facial expressions?

Findings indicated that PD was the only IRI scale that correlated with the DANVA2-AF scores. Combining the offender groups, those with high levels of personal distress were less likely to recognise angry faces and faces that showed a high intensity emotion. Analysing group correlations separately however revealed differing correlations. Personal distress remained the only IRI scale that correlated with DANVA2-AF scores. For IPV offenders, the higher level of personal distress experienced the more likely they were to make errors identifying angry and fearful emotions. In V offenders, incorrectly identifying sad emotions was related to higher levels of personal distress. Higher level of distress experienced by NV

offenders was related to high error rates identifying angry, high intensity and low intensity emotions.

These findings are similar to those found in Seidal et al.'s (2013) study of empathy and emotion recognition in violent offenders. They found that PD was the only IRI scale to show a significant difference between groups and it was the only scale that significantly correlated with emotion recognition. It was found that PD scores significantly positively correlated with errors in identifying disgust. Further context to present results may be found in Covell, Huss and Langhinrichsen-Rohling's (2007) findings that the PD scale was correlated with higher levels of violence amongst their participant groups. Thus it may be hypothesised that in the present study, those with higher PD scores had higher levels of violence and significant deficit in recognising negative emotions.

The lack of correlation between the DANVA2-AF and three IRI scales is consistent with previous findings. Besille and Yuille (2010) found that cognitive measures of empathy did not correlate with facial emotion recognition. This may be due to the type of empathy that the assessment tools are assessing. Whereas the IRI required offenders to infer about their empathic capacity, the DANVA2-AF assesses dispositional empathy. The DANVA2-AF assesses whether participants can recognise the emotions of others whereas the IRI scales ask participants how they respond to situations. Considering Marshall and Marshall's (2011) theory of empathy, these are two very distinct processes. This is supported by prior literature that has shown that whilst differing offender subgroups did not show significant differences in emotional recognition tasks, they showed differing physiological responses (measured by cardiovascular hyperactivity and skin-conduce response). Violent and psychopathic offenders had much lower physiological responses when exposed to negative emotions than controls (Lobbestael et al., 2009; Seidal et al., 2013). Similar processes may have occurred in the present study. Therefore, lack of correlation may be due to the fact that the measures are

assessing different constructs. In this case, correlation would not be necessarily be expected. Future research would benefit from combining facial recognition tasks with physiological measurements in order to assess for these differing stages of empathy in offenders.

Additionally, the length of exposure in the DANVA2-AF task may have impacted upon the type of emotional processing. The DANVA2-AF faces were presented for approximately 2000ms (which is the standard time for the task). In Besille and Yuille's (2010) experiment, at 2000ms the EC scale of the IRI did not significantly predict facial emotion recognition accuracy, but did so at 50ms. Thus, the present findings add to previous research, which suggests that longer exposure durations (which facilitate top-down processing) are not correlated with affective empathy capacity. There is some evidence to suggest that when emotion recognition utilises automatic processing (at shorter exposure durations) it correlates with affective empathy, thus differing results may have been found in the present study if faces were shown for a shorter duration.

Considering what these findings are explaining, it may be that offenders own experience of high distress impaired their ability to recognise the emotional experiences of others. As explored in Chapter 2, increased affective distress may inhibit offender's empathic capacity. If this is the case, intervention to increase empathic capacity should focus first on supporting healthy emotional regulation. Once offenders are able to better self-regulate their emotions, they may be more responsive to emotionally and cognitively empathise with others.

Limitations

There are a number of limitations in the present study that may impact upon the interpretation of findings. Regarding methodological limitations, having three groups and a small sample size limits the applicability of findings. Further research is required to explore

the significant findings of this study. Recruiting opportunistically within the prison, although having its benefits, also risked that the sample was only made up of participants willing to converse with the researcher and mobilising within the prison. Therefore, there is a possibility that rather than be representative as intended, the sample did not represent offenders who stayed in their cells or were not comfortable talking to the researcher. Additionally, the sample was relatively small (n =70) considering there were three groups. Small samples increase the likelihood of making Type II errors and therefore ideally the research would have had a larger sample.

Prior literature indicated that IPV offenders may have a specific empathy deficit towards their partners (Clements, Holtzworth-Munroe, Schweinle & Ickes, 2007) and that empathic capacity may be influenced by the relationship between offender and subject (Marshall & Marshall, 2011). This hypothesis was not explored in the present study, but would be useful to explore in future research.

Regarding the measures used, despite the IRI being deemed the most appropriate assessment of empathy for the study, there has been question in the literature regarding how valid the IRI is amongst prison populations due to factors such as the negatively worded items (Beven, O'Brien-Malone & Hall, 2004). There is no research to indicate whether the IRI is an effective tool to distinguish between offender subtypes.

The DANVA2-AF has only 6 faces of each emotion. Considering the outcomes are accuracy rather than timing, the small amount of emotions for the amount of outcomes makes it difficult to observe a meaningful difference between groups. It may have also been informative to include a measure of physiological arousal in the study in order to measure response as well as recognition.

Another difficulty for the present study is that offender IQ and economic status were not accounted for. Jolliffe and Farrington (2004) found that effects of affective empathy did not stand when these factors were controlled for. As the only significant findings in the present study related to outcome of affective empathy, controlling for IQ and economic status may have impacted upon findings.

Conclusions

The present study provided some suggestion that compared to non-violent offenders, IPV offenders had deficits with affective empathy, namely empathic concern and interpretation of fearful faces as sad. Both of these deficits may inhibit an offender's use of violence by reducing their ability to understand the perspective of their victims and increasing their perception of hostility in victims.

Importantly, the research found no significant differences between IPV and V offenders. Practically, this would suggest that both IPV and V offenders have similar empathic deficits and treatment programmes could be designed to target both offender groups together rather than separately. Treatment targets may benefit from supporting violent and IPV offenders to better distinguish between negative emotions and to understand the emotional experiences of others.

The study is unable to inform as to whether the presence of empathic deficits are causal in the perpetration IPV. Further prospective research is required with larger participant numbers and across different prisons in order to investigate the true effect. Findings from the current research are unable to determine whether empathic deficits in offenders are significantly different for non-offenders, as a community non-offending sample was not included. Future research may also benefit from include a non-offending sample to better explore this difference.

Chapter 5

Discussion

This thesis aimed to explore the role of empathy in family violence, specifically IPV and CM. Empathy was considered an important construct to explore as it is an often cited determinant of the way in which humans behave towards each other and is therefore considered both a risk and protective factor for violence. Furthermore, in forensic practice, assessments of IPV and CM often include assessment of empathy and interventions include modules to increase empathic capacity.

Collectively, the body of work presented in this thesis integrates findings related to empathy and CM and explores the role of empathy in IPV offenders compared to other offender groups. It provides a specific focus on empathy in the role of family violence. This is in comparison to the majority of research which has included empathy measures as part of a wider assessment (e.g. of cognitive/affective functioning). Therefore, this thesis facilitates a more detailed analysis of empathy in the role of family violence, which helps to better inform practice and future research recommendations.

Summary of findings

Chapter 2

The literature review aimed to explore the relationship between empathy and CM. The review analysed 17 studies published between 1985 and 2014. 16 of these studies found that maltreating parents scored significantly lower on measures of empathy than non-maltreating parents, with a stronger relationship between CM and cognitive empathy than affective empathy. There were however differences amongst studies regarding which scales/assessments found significant differences. For example, regarding outcomes relating to the IRI personal distress, perspective taking and empathic concern subscales, 71%, 60% and 50% of studies respectively found a significant difference.

In addition to differences in outcomes relating to empathy, there was also disparity related to the quality of studies. Most of the studies had small sample sizes and had not considered which construct of empathy they were assessing. A concerning finding was that many of the studies assumed that their control group were non-maltreating rather than assessing this directly. None of the studies were completed in the UK and most of the studies were completed by one of two research institutes.

Considering implications for practice, the review highlights that maltreating parents appear to present with empathic deficits, thus the need to assess parental empathy in child protection assessments is supported by empirical evidence. However, reviewed literature is unable to inform: 1) which constructs of empathy are most important to assess and increase; 2) whether to assess and increase general empathy or child-specific empathy; 3) how important empathic deficits are in explaining CM perpetration. These are key considerations that should be explored in future research. It is also important for research to use both questionnaire and task-based assessments in order to provide a more informative empathic profile.

Chapter 3

The literature review in Chapter 2 identified that the IRI was one of the most commonly used assessments of empathy in CM. As a frequently used measure of empathy in both clinical and forensic settings (Konrath, O'Brien & Hsing, 2011), Chapter 3 explored the psychometric properties of the IRI. The analysis identified that the IRI has positive validity and reliability and benefits from having a range of normative data available. Other strengths of the IRI include its multidimensional assessment of empathy and its quick and easy administration.

Concerns were raised regarding the validity of the fantasy scale as the theory underpinning the scale (that empathy towards fictional characters is a distinct construct) is unsupported and each of the seven scale questions appears to assess the exact same construct. Accordingly, a large number of studies omit this scale. This was also found in studies reviewed in Chapter 2. Additional concern was raised regarding the personal distress subscale, with Alterman et al. (2013) suggesting that it relates to narcissism rather than empathy. Questions were also raised regarding the validity of the reverse items of the IRI, due to the increased level of processing required for such items. Preliminary research suggested that the negatively-worded items reduced the internal reliability of the measure in offending populations.

Considering practical implications, the review identified that the IRI is a valid and reliable measure of general empathy to be used in research and clinical settings. Adaptions have been made by researchers to make the assessment person-specific, although standardised data is not available for these. When using the IRI, it is important that researchers/clinicians intend to measure general trait empathy. Researchers/clinicians would benefit from using the IRI in conjunction with a task-based assessment in order to include assessment of state empathy.

Chapter 4

The empirical study presented in Chapter 4 aimed to examine the differences in empathy between IPV (n=30), violent (n=20) and non-violent (n=20) offenders. Previous studies interested in this research area had recruited participants from the community and no study had compared IPV offenders with other offender subtypes on measures of empathy. In line with the recommendations from Chapter 3, the study used the IRI and an emotion recognition task (DANVA2-AF) in order to provide a more encompassing assessment.

Results found that although there was not a significant overall group effect, IPV offenders had significantly lower empathic concern scores and were more likely to interpret fearful faces as sad than non-violent offenders. The study would have benefitted from including a non-violent non-offending group, in order to explore whether significant differences occurred between offenders and non-offenders. Correlating the IRI and DANVA2-AF scores, the PD scale showed significant relationships: those with higher levels of PD were less likely to correctly identify angry and high intensity emotions. However, contextualising these findings with findings from Chapter 3 and a finding that those with higher PD scores are more violent (Covell, Huss & Langhinrichsen-Rohling; 2007), it may be that this finding relates more to narcissism and violence potential than empathy.

The presented study identified that regarding empathy, IPV offenders did not significantly differ from other offenders on cognitive empathy and did not significantly differ from violent offenders on any measure. The research was limited by the sample being small in size and being recruited from a local male prison. Accordingly, findings cannot be generalised to the wider population without further supporting evidence.

It would be useful for further research to compare IPV and violent offenders on other criminogenic factors in order to better identify whether assessing and providing intervention to IPV offenders separately is justified. The current research has not informed whether capacity for victim-empathy is different from general empathic capacity in offenders. This would certainly be a useful aim of future research, which may benefit from using both qualitative and quantitative designs in order to understand whether lack of empathy was a feature of an offender's offence (e.g. index offence analysis). Models of empathy suggest that the relationship a person has with another determines the ability to empathise with them (Marshall and Marshall, 2011), thus suggests a difference would be found.

Synthesis of Findings

A key finding from the chapters is that of the role of emotional regulation. In the literature review, personal distress was generally found to be significantly different between maltreating and non-maltreating parents. In the research chapter, personal distress negatively impacted upon participants' ability to recognise negative emotions. Chapter 2 explored research that demonstrated that increased emotional arousal negatively impacts upon ability to process and access cognitive information. Thus, I suggest that emotional dysregulation acts as an inhibitor to empathic processing (as a state). This hypothesis would help to explain the inconsistent findings in the literature and in the Chapter 4 regarding empathic capacity of offenders and maltreating parents. Questionnaire based assessments only assess trait empathy and most participants are likely to complete assessments in a calm and controlled state. State empathy however is likely to be negatively impacted by emotional dysregulation and thus explains why maltreating parents/offenders are able to demonstrate empathy during interviews but do not exercise empathic capacity at the time of offending/high distress.

Considering intervention goals, supporting clients to effectively self-regulate their emotions appears to be a primary treatment need. This is likely to have a two-fold effect by increasing their empathic capacity and providing them with the resilience to understand the affect/perspective of others and the potential impact that their behaviour has had on victims.

An important component to consider is whether cognitive and affective empathy can be considered as separate phenomenon. It appears that although separate constructs, they mutually influence each other and thus measurements and separation of them is confounded. This would explain why factor analyses of the IRI (Chapter 3) combine perspective taking and empathic capacity scales. It appears logical that you must understand the perspective of somebody else in order to feel their affect and that feeling the affect of somebody else

increases perspective taking. This links to de Waal's (2008) theory that suggests emotional understanding is a unique process but a part of many human processes. Accordingly, it is likely to be a feature of many processes, including empathy, but considered a 'lower-order' aspect of 'higher-order' more complex human behaviours.

Future research

This thesis has helped to identify gaps in knowledge with relation to family violence. With regard to both CM and IPV, there is a lack of understanding about whether perpetrators general empathic capacity is different from empathy towards their family/victims. Three studies analysed in Chapter 2 (Francis & Wolfe, 2008; Perez-Albeniz & de Paul, 2004; Rodriguez, 2013), examining parental empathy towards their own children, found a significant difference between maltreating and non-maltreating parents, suggesting that a difference may be found. It is clear that empathic deficits exist amongst perpetrators of CM, although specifically what these deficits are is less clear. Some of this uncertainty has been due to studies using different measures of empathy and therefore not employing a consistent approach. The critical analysis of the IRI identified that using questionnaire-based assessments alone limits the effectiveness of studies, thus future research would benefit from combining both questionnaire and task-based assessments.

The study in Chapter 4 found some differences between IPV and NV offenders in affective empathy and emotion recognition. However it did not inform whether IPV offenders have a differing empathic profile to non-offenders. Thus future research would benefit from identifying how offender empathy differs from non-offender empathy, and how the empathy profile of IPV offenders fits in with this. Furthermore, research may benefit from moving from looking at IPV offenders as a homogenous group, to examining differences amongst IPV perpetrators (Covell, Huss & Langhinrichsen-Rohling, 2007). It is important for future

research to also consider the impact of emotional regulation on their outcomes. In order to be able to assess somebody's true spectrum of empathic capacity, it is important to understand how empathically they respond when they are in a non-threatening situation (trait empathy) and in a situation that evokes distress (state empathy).

It would be interesting for research to progress findings that suggest a high prevalence of co-occurring family violence (Bowen, 2000; Slep & O'Leary, 2005). Specifically related to this thesis, it would be informative to explore the empathic profiles of individuals who perpetrate both IPV and CM, and how these link to the perpetrator profiles (Paternal/Maternal, Hierarchical and Reciprocal) proposed by Dixon, Browne, Hamilton-Giachritis and Ostapuik (2010).

Despite this thesis exploring male perpetrated IPV and largely female perpetrated CM, it is important to acknowledge that both phenomena are gender-inclusive. The limited available data regarding female perpetrated IPV is likely to be due to a number of factors including the influence of the gendered theory on current practice, which proposes that 'gender is the most significant factor' in explaining IPV (Respect, 2008, P.1). This assertion is made despite an overwhelming opposing evidence base (Dixon, Archer & Graham-Kevan, 2012). As a consequences of the gendered theory influencing bias in current practice, males are more likely to be considered perpetrators by the criminal justice system for IPV and general violence compared to females (Hester, 2013), despite equal rates of perpetration suggested for both violence types (Thornton, Graham-Kevan & Archer, 2012). Research samples are often obtained from child protection services or women's shelters, which inevitably generates a 'clinical fallacy' (Straus & Gelles, 1999) and distorted statistics regarding IPV perpetration that cannot be applied to the general population. Similar processes are likely to be influential in the CM literature and practice also. Accordingly, it is important

for future research to obtain comparative data for both male and female samples, preferably taking a whole family perspective in order to capture co-occurrence of family violence.

Conclusions

This thesis identifies that empathy plays a role in family violence. The role of empathy in CM was strongly evidenced in the literature review which explored findings from several studies. Influence of empathy in IPV perpetration was not strongly evidenced in the research chapter. With regards to CM, cognitive empathy strongly differentiated between maltreating and non-maltreating parents in comparison to affective empathy. The opposite pattern was seen in IPV: IPV offenders showed a trend towards more impaired affective empathy than NV offenders. This tentatively suggests that the influence of empathy is different in CM and IPV. However, the small scale nature of the IPV study must be bore in mind. It is important that researchers interested in the role of empathy in family violence ensure that research is progressive, as identified in Chapter 2, research paradigms and questions today are similar to those posed forty years ago. Key questions to be explored are: 1) the extent to which empathy contributes to family violence perpetration and rehabilitation; 2) whether empathy is a risk and/or protective factor in family violence; 3) whether empathy in family violence is a general or victim-specific deficit.

References

Abel, G. G., Gore, D. K., Holland, C. L., Camp, N., Becker, J. V.,& Rathner, J. (1989). The measurement of the cognitive distortions of child molesters. *Annals of Sex Research, 2,*135–152. doi: 10.1007/BF00851319

Abidin, R. I. (1990). *Parenting Stress Index manual* (3rd ed.). Charlottesville, VA: Pediatric Psychology Press.

Allport, F. H. (1924). *Social Psychology*. Boston: Houghton Mifflin Co.

Alterman, A.I., McDermott, P.A., Cacciola, J.S., & Rutherford, M.J. (2003). Latent structure of the Davis Interpersonal Reactivity Index in methadone maintenance patients. *Journal of Psychopathology and Behavioral Assessment*, *25*, 257-265. doi: 10.1023/A:1025936213110

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Andrews, D. A., & Bonta, J. (2003). *The Psychology of Criminal Conduct* (3rd ed.) Cincinnati, OH: Anderson.

Appleyard, K., Egeland, B., van Dulmen, M., & Sroufe, A. (2005). When more is not better: the role of cumulative risk in child behavior outcomes. *Journal of Child Psychology, 46,* 235–245. doi:10.1111/j.1469-7610.2004.00351.x

Azar, S.T. (1986). A framework for understanding child maltreatment: An integration of cognitive behavioural and developmental perspectives. *Canadian Journal of Behavioural Science, 18*, 340 – 355. doi:10.1037/h0079961

Azar, S. T., Reitz, E. B., & Goslin, M. C. (2008). Mothering: thinking is part of the job description: application of cognitive views to understanding maladaptive parenting and doing intervention and prevention work. *Journal of Applied Developmental Psychology, 29,* 295–304. doi:10.1016/j.appdev.2008.04.009

Azar, S.T., Robinson, D.R., Hekimian, E., & Twentyman, C.T. (1984). Unrealistic expectations and problem-solving ability in maltreating and comparison mothers. *Journal of Consulting and Clinical Psychology, 52*, 687–691. doi:10.1037/0022-006X.52.4.687

Babcock, J.C., Green, C.E., & Webb, S.A. (2008). Decoding deficits of different types of batterers during presentation of facial affect slides. *Journal of Family Violence, 23,* 295 – 302. doi: 10.1007/s10896-008-9151-1

Barnett, G., & Mann, R.E. (2013). Empathy deficits and sexual offending: a model of obstacles to empathy. *Aggression and Violent Behaviour, 18,* 228 – 239. doi: 10.1016/j.avb.2012.11.010

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, 163 – 175. doi: 10.1023/B:JADD.0000022607.19833.00

Batson, C. D. (1991). *The Altruism Question: Toward a Social-Psychological Answer.* Hillsdale, NJ: Erlbaum

Batson, C. D. (2009). These things called empathy: Eight related but distinct phenomena. In J. Decety and W. Ickes (Eds.), *The Social Neuroscience of Empathy* (pp. 3-15). Cambridge: MIT Press.

Batson, C. D., & Ahmad, N.Y. (2009). Using empathy to improve intergroup attitudes and relations. *Social Issues and Policy Review, 3,* 141-177. doi:10.1111/j.1751-2409.2009.01013.x

Batson, C. D., & Coke, J. S. (1981). Empathy: A source of altruistic motivation for helping? In J. P. Rushton & R. M. Sorrentino (Eds.), *Altruism and Helping behavior* (pp. 167-187). Hillsdale, NJ: Erlbaum.

Batson, C.D., Sager, K., Garst, E., Kang, M., Rubchinsk, K., & Dawson, K. (1997). Is empathy-induced helping due to self-other merging? *Journal of Personality and Social Psychology, 73,* 495 – 509. doi: 10.1037/0022-3514.73.3.495

Baum, K.M., Logan, M. C., Walker, E.F., Tomlinson, H., & Schiffman, J. (1996). *Emotion Recognition in Adolescents with Schizotypal Personality Disorder.* Paper presented at the annual meeting of the Society for Research in Psychopathology. Atlanta: GA.

Bavolek, S. (1984). *Handbook for the AAPI: Adult-adolescent parenting inventory.* Park City, UT: Family Development Resources, Inc.

Bavolek, S. J., & Keene, R. G. (2001). Adult-Adolescent Parenting Inventory (AAPI-2): Administration and development handbook. Park City, UT: Family Development Resources, Inc.

Bellini, L.M., & Shea, J.A. (2005). Mood change and empathy decline persist during three years of internal medicine training. *Academic Medicine, 80,* 164 – 167. doi: 10.1097/00001888-200502000-00013

Belsky, J. (1980). Child maltreatment: An ecological integration. *American Psychologist, 35*, 320-335. doi:10.1037/0003-066X.35.4.320

Berman, G., & Dar, A. (2013) *Prison population statistics*. London: House of Commons Library. Retrieved from www.parliament.uk/briefing-papers/sn04334.pdf

Beven, J.P., O'Brien-Malone, A., & Hall, G. (2004). Using the Interpersonal Reactivity Index to assess empathy in violent offenders. *International Journal of Forensic Psychology, 1,* 33 – 41. Retrieved from: http://researchrepository.murdoch.edu.au/2698/1/Using_the_Interpersonal_Reactivity_Index_to_Assess_Empathy_in_Violent_Offenders.pdf

Biehal, N., & Parry, E. (2010). *Maltreatment and Allegations of Maltreatment in Foster Care. A Review of the Evidence.* The University of York. Retrieved from http://resources.leavingcare.org/uploads/377d77e8dafa3aa5fd060dcef3748cae.pdf

Bock, E.M., & Hosser, D. (2014). Empathy as a predictor of recidivism among young adult offenders. *Psychology, Crime and Law, 20,* 101 – 115. doi: 10.1080/1068316X.2012.749472

Bowen, K. (2000). Child abuse and domestic violence in families of children seen for suspected sexual abuse. *Clinical Paediatrics, 39,* 35-40. doi: 10.1177/000992280003900104

Brandon, M., & Lewis, A. (1996). Significant harm and children's experiences of domestic violence. *Child and Family Social work, 1,* 33 – 42. doi: 10.1111/j.1365-2206.1996.tb00005.x

Brandon, M., Sidebotham, P., Bailey, S., Belderson, P., Hawley, C., Ellis, C., & Megson, M. (2012). *New learning from serious case reviews: a two year report for 2009 – 2011*. London: Department for Education. Retrieved from http://webarchive.nationalarchives.gov.uk/20130401151715/https://www.education.gov.uk/publications/eOrderingDownload/DFE%20-%20RR226%20Report.pdf

Britton, P.C., & Fuendeling, J.M. (2010). The relations among varieties of adult attachment and the components of empathy. *The Journal of Social Psychology, 145*, 519 – 530. doi: 10.3200/SOCP.145.5.519-530

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*, 513-530. doi:10.1037/0003-066X.32.7.513

Browne, S., Harkins, L., & Beech, A.R. (2012). General and victim-specific empathy: associations with actuarial risk, treatment outcome and sexual recidivism. *Sexual Abuse: a Journal of Research and Treatment, 24,* 411 – 430. doi:10.1177/1079063211423944

Browne, K.D., & Herbert, M. (1997). *Preventing Family Violence*. Chichester: Wiley.

Bugental, D. B. (1998). *Parent Attribution Test: Revised manual.* Santa Barbara: University of California.

Bullock, K., Sarre, S., Tarling, R., & Wilkinson, M. (2010). *The Delivery of Domestic Abuse Programmes. An Implementation Study of the Delivery of Domestic Abuse Programmes in Probation Areas and Her Majesty's Prison Service. Ministry of Justice Research Series 15/10.* London: Ministry of Justice. Retrieved from http://www.justice.gov.uk/downloads/publications/research-and-analysis/moj-research/delivery-domestic-abuse-programmes.pdf

Butchart, A., Putney, H., Furniss, T., & Kahane, T. (2006). *Preventing child maltreatment: a guide to taking action and generating evidence*. Geneva: World Health Organisation. Retrieved from: http://whqlibdoc.who.int/publications/2006/9241594365_eng.pdf

Byrne, C. A., & Arias, I. (1997). Marital satisfaction and marital violence: moderating effects of attributional processes. *Journal of Family Psychology, 11*, 188-195. doi: 10.1037/0893-3200.11.2.188

Caddick, B., & Webster, A. (1998). Offender literacy and the probation service. *The Howard Journal, 37*, 137 - 147. doi: 10.1111/1468-2311.00086

Campbell, S.B., Brownell, C.A., Hungerford, A., Spieker, S.J., Mohan, R., & Blessing, J.S. (2004). The course of maternal depressive symptoms and maternal sensitivity as predictors of attachment security at 36 months. *Development and Psychopathology*, *16,* 231 – 252. doi:10.1017/S0954579404044499

Carey, J.C., Fox, E.A., & Spraggins, E.F. (1988). Replication of structure findings regarding the Interpersonal Reactivity Index. *Measurement and Evaluation in Counselling and Development, 21*, 102-105.

Carrasco, O.M.A., Delgado, E.B., Barbero, G.M.I., Holgado, T.F.P., & del Barrio, G.M.V. (2011). Psychometric properties of the Interpersonal Reactivity Index in Spanish child and adolescent population. *Psicothema, 23,* 824 – 831. Retrieved from: http://www.psicothema.com/PDF/3962.pdf

Casanueva, C., Martin, S.L., & Runyan, D.K. (2009). Repeated reports for child maltreatment among intimate partner violence victims: findings from the National Survey of Child and Adolescent well-being. *Child Abuse and Neglec*t, *33,* 84-93. doi:10.1016/j.chiabu.2007.04.017

Child Welfare Information Gateway (2004). *Risk and Protective Factors for Child Abuse and Neglect*. Washington, DC: US Department of Health and Human Services. Retrieved from http://www.childwelfare.gov/pubPDFs/riskprotectivefactors.pdf

Christopher, F. S., Owens, L. A., & Stecker, H. L. (1993). Exploring the darkside of courtship: A test of a model of male premarital sexual aggressiveness. *Journal of Marriage and the Family, 55*, 749 - 479. doi: 10.2307/352816

Cialdini, R.B., Brown, S.L., Lewis, B.P., Luce, C., & Neuberg, S.L. (1997). Reinterpreting the empathy-altruism relationship: when one into one equals oneness. *The Journal of Personality and Social Psychology, 73,* 481 – 494. doi:10.1037/0022-3514.73.3.481

Cleaver, H., Unell, I., & Aldgate, J. (2011). *Children's Needs – Parenting Capacity: The Impact of Parental Mental Illness, Problem Alcohol and Drug Use and Domestic Violence on Children's Development.* London: The Stationery Office. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/182095/DFE-00108-2011-Childrens_Needs_Parenting_Capacity.pdf

Clements, K., Holtzworth - Munroe, A., Schweinle, W., & Ickes, W. (2007). Empathic accuracy of intimate partners in violent versus nonviolent relationships. *Personal Relationships, 14,* 369 – 88. doi: 10.1111/j.1475-6811.2007.00161.x

Cliffordson, C. (2002). The hierarchical structure of empathy: dimensional organization and relations to social functioning. *Scandinavian Journal of Psychology, 43,* 49–59. doi: 10.1111/1467-9450.00268

Cohen, D., & Strayer, J. (1996). Empathy in conduct-disordered and comparison youth. *Developmental Psychology, 32,* 988–998. doi:10.1037/0012-1649.32.6.988

Comrey, A. L. (1978). Common methodological problems in factor analytic studies. *Journal of Consulting and Clinical Psychology, 46,* 648-659. doi: 10.1037/0022-006X.46.4.648

Conron, K.W., Beardslee, W., Koenen, K.C., Buka, S.L., & Gortmaker, S.L. (2009). A longitudinal study of maternal depression and child maltreatment in a national sample of families investigated by child protective services. *Archive of Paediatric Adolescent Medicine, 163,* 922 – 930. doi:10.1001/archpediatrics.2009.176.

Covell, C.N., Huss, M.T., & Langhinrichsen - Rohling, J. (2007). Empathic deficits among male batterers: A multidimensional approach. *Journal of Family Violence, 22,* 165 – 174. doi: 10.1007/s10896-007-9066-2

Critical Appraisal Skills Programme (2004). *Critical Appraisal Skills Programme: Making sense of evidence. 12 questions to help you make sense of a cohort study*. Milton Keynes: Milton Keynes Primary Care Trust. doi:10.3310/hta15160

Crittenden, P.M. (2006). A dynamic-maturation model of attachment. *Australian and New Zealand Journal of Family Therapy, 27,* 105 – 115. doi:10.1002/j.1467-8438.2006.tb00704.x

Davis, M.H. (1980). A multidimensional approach to individual differences in empathy. *Journal Supplement Abstract Services Catalog of Selected Documents in Psychology*, *10*, 85. Retrieved from: http://www.researchgate.net/profile/Mark_Davis18/publication/34891073_Individual_differences_in_empathy__a_multidimensional_approach_/links/0046352dee2296ea54000000.pdf

Davis, M.H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*, 113–126. doi: 10.1037/0022-3514.44.1.113

Davis, M. H. (1994). *Empathy: A Social Psychological Approach.* Madison, WI: Brown and Benchmark.

Davis, M. H., & Franzoi, S. L. (1991). Stability and change in adolescent self-consciousness and empathy. *Journal of Research in Personality, 25,* 70-87. doi: 10.1016/0092-6566(91)90006-C

Davis, M. H., Mitchell, K.V., Hall, J.A., Lothert, J., Snapp, T., & Meyer, M. (1999). Empathy, expectations, and situational preferences: Personality influences on the decision to participate in volunteer helping behaviors. *Journal of Personality*, *67*, 469-503. doi: 10.1111/1467-6494.00062

Day, A., Mohr, P., Howells, K., Gerance, A., & Lim, L. (2012). The Role of Empathy in Anger Arousal in Violent Offenders and University Students. *Journal of Offender Therapy and Comparative Criminology, 56,* 599 – 613. doi: 10.1177/0306624X11431061

Decety, J., & Jackson, P.L. (2006). A social-neuroscience perspective on empathy. *Current Directions in Psychological Science, 15,* 54 – 58. doi: 10.1111/j.0963-7214.2006.00406.x

De Corte, K., Buysse, A., Verhofstadt, L.L., Roeyers, H., Ponnet, K., & Davis, M.H. (2007). Measuring empathic tendencies: reliability and validity of the Dutch version of the interpersonal reactivity index. *Psychologica Belgica, 47,* 235 – 260. doi: 10.5334/pb-47-4-235

Department for Education (2013). *Working together to safeguard children: A guide to inter-agency working to safeguard and promote the welfare of children.* London: The Stationary Office. Retrieved from http://media.education.gov.uk/assets/files/pdf/w/working%20together.pdf

Department for Education (2014). *Statistical First Release: Children looked after in England (including adoption and care leavers) year ending 31 March 2014.* London: The Stationary Office. Retrieved from https://www.gov.uk/government/statistics/children-looked-after-in-england-including-adoption--2.

de Paul, J., Perez-Albeniz, A., Guibert, M., Asla, N., & Ormaechea, A. (2008). Dispositional empathy in neglectful mothers and mothers at high risk for child physical abuse. *Journal of Interpersonal Violence, 23*, 670-684. doi:10.1177/0886260507313532.

DeVellis, R. (1991). *Scale development: theory and applications*. London: Sage Publications.

de Vogel, V., de Ruiter, C., Bouman, Y., & de Vries Robbé, M. (2009). *SAPROF. Guidelines for the assessment of protective factors for violence risk. English version.* Utrecht: Forum Educatief.

De Waal, F.B.M. (2008). Putting the altruism back into altruism the evolution of empathy. *The Annual review of Psychology, 59, 279 – 300*. doi: 10.1146/annurev.psych.59.103006.093625

Dixon, L., Archer, J., & Graham-Kevan, N. (2012). Perpetrator programmes for partner violence: Are they based on ideology or evidence? *Legal and Criminological Psychology, 17,* 196–215. doi: 10.1111/j.2044-8333.2011.02029

Dixon, L., Browne, K., & Hamilton-Giachritsis, C. (2009). Patterns of risk and protective factors in the intergenerational cycle of maltreatment. *Journal of Family Violence*, *24,* 111-122. doi:10.1007/s10896-008-9215-2

Dixon, L., Browne, K.D., Hamilton-Giachritsis, C.E., & Ostapuik, E. (2010). Differentiating patterns of aggression in the family. *Journal of Aggression Conflict and Peace Research, 2*, 32-44. doi: 10.5042/jacpr.2010.0003

Dixon, L., & Graham-Kevan, N. (2011). Until Death Do They Part: Preventing Intimate Partner Homicide. *The Psychologist, 24,* 820-823. Retrieved from https://thepsychologist.bps.org.uk/volume-24/edition-11/until-death-do-us-part

Dixon, L., Hamilton-Giachritsis, C., & Browne, K. (2005). Attributions and behaviours of parents abused as children a meditational analysis of the intergenerational continuity of child maltreatment (part II). *Journal of Child Psychology and Psychiatry*, *46,* 58-68. doi:10.1111/j.1469-7610.2004.00340.x

D'Orazio, D.M. (2004). Letter to the editor: The Journal's Publication of Research that Incorrectly Employs Davis' Interpersonal Reactivity Index. *Sexual Abuse: A Journal of Research and Treatment, 16,* 173–174. Retrieved from: http://sax.sagepub.com/content/16/2/173.full.pdf

Douglas, K. S., & Reeves, K. A. (2010). Historical- Clinical - Risk- Management - 20 (HCR-20) Violence Risk Assessment Scheme: Rationale, application and empirical overview. In R. K. Otto, & K. S. Douglas (Eds.), *Handbook of violence risk assessment: International perspectives of forensic mental health* (1st ed.) (pp. 147-186). New York; London: Routledge.

Dutton, D.G. (1985). An ecologically nested theory of male violence towards intimates. *International Journal of Women's Studies, 8,* 404-413. Retrieved from: http://www.researchgate.net/publication/232569283_An_ecologically_nested_theory_of_male_violence_toward_intimates

Edleson, J.L. (1999). Children's witnessing of adult domestic violence. *Journal of Interpersonal Violence, 14,* 839 – 870. doi:10.1177/088626099014008004

Easterbrook, J.A. (1959). The effect of emotion on cue utilization and the organization of behaviour. *Psychology Review, 66,* 183 – 201.

Effective Public Health Practice Project. (2007). *Quality Assessment Tool for Quantitative Studies.* Retrieved from http://www.ephpp.ca/PDF/Quality%20Assessment%20Tool_2010_2.pdf

Egger, M., Dickerson, K., & Smith, G.D. (2001) Problems and limitations in conducting systematic reviews. In M. Egger, G. Smith, & D. Altman (Eds.), *Systematic Reviews in Health Care: Meta-analysis in Context.* (2nd ed., pp. 43 – 68). London, UK: BMJ Publishing Group.

Eisenberg, N. (1982). The development of reasoning regarding prosocial behavior. In N. Eisenberg (Eds.), *The Development of Prosocial Behavior* (pp. 219–249). New York: Academic Press. doi:10.1016/B978-0-12-234980-5.50014-6

Eisenberg, N. (1991). Values, sympathy and individual differences: toward a pluralism of factors influencing altruism and empathy. *Psychological Inquiry, 2*, 128-131. doi:10.1207/s15327965pli0202_5

Eisenberg, N., & Fabes, R.A. (1990). Empathy: conceptualization, measurement and relation to prosocial behaviour. *Motivation and Emotion, 14,* 131 – 149. Retrieved from http://link.springer.com/article/10.1007%2FBF00991640

Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin, 94*, 100-131. doi:10.1037/0033-2909.94.1.100

Eisenberg, N., & Miller, P. A. (1987). The Relation of Empathy to Prosocial and Related Behaviors. *Psychological Bulletin, 101*, 91–119. doi: 10.1037/0033-2909.101.1.91

Eisenberg, N., & Strayer, J. (1987). Critical issues in the study of empathy. In N. Eisenberg & J. Strayer (Eds.), *Empathy and Its Development* (pp. 292-316). Cambridge: Cambridge University Press.

Ekman, P., & Friesen, W. V. (1976). *Pictures of Facial Affect.* San Francisco: University of California.

Emde, R.N., Osofsky, J.D., & Butterfield, P.M. (1994). The IFEEL pictures, a new instrument for interpreting emotions. *Infant Mental Health Journal, 15*, 96 – 97. doi: 10.1002/1097-0355(199421)15:1<96::AID-IMHJ2280150111>3.0.CO;2-J

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Persepctive taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology, 87,* 327 – 339. doi: 10.1037/0022-3514.87.3.327

Felitti, V.J., Anda, R.F., Nordenberg, D., Williamson, D.F., Spitz, A.M., Edwards, V., Koss, M.P. & Marks, J.S. (1998) Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE) Study. *American Journal of Preventive Medicine, 14*, 245–258

Felson, R. B. (2002).*Violence and Gender Re-examined.* Washington, DC: American Psychological Association.

Fernandez, A.M., Dufey, M., & Kramp, U. (2011). Testing the psychometric properties of the interpersonal reactivity index (IRI) in Chile: Empathy in a different cultural context. *European Journal of Psychological Assessment, 27,* 179 – 185. doi: 10.1027/1015-5759/a000065

Fernandez, Y. M., Marshall, W. L., Lightbody, S., & O'Sullivan, C. (1999). The child molester empathy measure: Description and examination of is reliability and validity. *Sexual Abuse: A Journal of Research and Treatment, 11*, 17-31. doi:10.1177/107906329901100103

Feshbach, S. (1964). The function of aggression and the regulation of aggressive drive. *Psychological Review, 71*, 257-272. doi:10.1037/h0043041

Feshbach, N. D., & Caskey, N. (1985). *A new scale for measuring parent empathy and partner empathy: Factorial structure, correlates and clinical discrimination.* Unpublished manuscript. University of California: Los Angeles.

Feshbach, N.D., & Feshbach, S. (1969). The relationship between empathy and aggression in two age groups. *Developmental Psychology, 1,* 102 – 107. doi:10.1037/h0027016

Feshbach, N. D., & Roe, K. (1960). Empathy in six- and seven-year olds. *Child Development, 34,* 133-145. doi: 10.2307/1127365

Fincham, F.D., Paleari, F.G., & Regalia, C. (2002). Forgiveness in Marriage: The Role of Relationship Quality, Attributions and Empathy. *Personal Relationships*, *9*, 27 – 37. doi: 10.1111/1475-6811.00002

Forrester, D., & Harwin, J. (2006). Parental substance misuse and child care social work: findings from the first stage of a study of 100 families. *Child and Family Social work*, *11,* 325 – 355. doi:10.1111/j.1365-2206.2006.00415.x

Francis, K.J., & Wolfe, D.A. (2008). Cognitive and emotional differences between abusive and non-abusive fathers. *Child Abuse and Neglect, 32,* 1127-1137. doi:10.1016/j.chiabu.2008.05.007

Gallese, V. (2001). The "shared manifold" hypothesis: from mirror neurons to empathy. *Journal of Consciousness Studies, 8*, 33 – 50. Retrieved from: http://didattica.uniroma2.it/assets/uploads/corsi/33846/Gallese_2001.pdf

Geer, J.H., Estupinan, L.A., & Manguno-Mire, G.M. (2000). Empathy, social skills and other relevant cognitive processes in rapist and child molesters – Issues, theories and treatment of the offender. *Aggression and Violent Behaviour, 5,* 99 – 126. doi: 10.1016/S1359-1789(98)00011-1

Gibbs, J. (1987). Social processes in the causation and treatment of delinquency: The need to facilitate empathy as well as sociomoral reasoning. In W. Kurtins, & J. Getwits (Eds.), *Moral development through social interaction* (pp. 301–321). New York: Wiley

Gilbert, R., Kemp, A., Thoburn, J., Sidebotham, P., Radford, L., Danya, G., & MacMillan, H.L. (2009). Recognising and responding to child maltreatment. *Lancet, 373,* 167 – 180. doi: :10.1016/S0140- 6736(08)61707-9

Gilbert, R., Spatz-Widom, C., Browne, K., Fergusson, D., Webb, E., & Janson, S. (2009). Burden and consequences of child maltreatment in high-income countries. *The Lancet, 373*, 68–81. doi:10.1016/S0140-6736(08)61706-7

Glass, G.V., Peckham, P.D., & Saunders, J.R. (1972). Consequences of failure to meet assumption underlying the fixed effects analyses of variance and covariance. *Review of Educational Research, 42*, 237–288. doi: 10.3102/00346543042003237

Golding, J.M. (1999). Intimate partner violence as a risk factor for mental disorder: a meta-analysis. *Journal of Family Violence, 14,* 99 – 132. Retrieved from http://link.springer.com/article/10.1023/A:1022079418229#page-1

Gottfredson, M. R., & Hirschi, T. (2007). *A General Theory of Crime* (3rd Ed.). USA: Stanford University Press.

Gu, X., Hof, P.R., Friston, K.J., & Fan, J. (2013). Anterior insular cortex and emotional awareness. *The Journal of Comparative Neurology, 521,* 3371 – 3388. doi: 10.1002/cne.23368

Haas, B.W., Anderson, I.W., & Filowski, M.M. (2015). Interpersonal reactivity and the attribution of emotional reactions. *Emotion.* Advanced online publication. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/25706827

Haker, H., Schimansky, J., Jann, S., & Rössler, W. (2012). Self-reported empathic abilities in schizophrenia: a longitudinal perspective. *Psychiatry Research, 200,* 1028 – 1031. doi: 10.1016/j.psychres.2012.04.004

Hamby, S., Finkelhor, D., Turner, H., & Ormrod, R. (2010). The overlap of witnessing partner violence with child maltreatment and other victimizations in a nationally representative survey of youth. *Child Abuse and Neglect, 34,* 734-741. doi: http://www.sciencedirect.com/science/article/pii/S0145213410002127

Hanson, R.K., & Bussière, M.T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical psychology, 66,* 348 – 362. doi: 10.1037//0022-006X.66.2.348

Hanson, R. K., & Morton-Bourgon, K.E. (2004). *Predictors of sexual offender recidivism: An updated meta*-analysis (Corrections Research User Report No. 2004–02). Ottawa, Ontario: Public Safety and Emergency Preparedness Canada. Retrieved from https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/2004-02-prdctrs-sxl-rcdvsm-pdtd/2004-02-prdctrs-sxl-rcdvsm-pdtd-eng.pdf

Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology, 73*, 1154-1163. doi:10.1037/0022-006X.73.6.1154

Hare, R. D. (1991). *The Hare Psychopathy Checklist - Revised (PCL-R)* (2nd ed.). Toronto, Ontario, Canada: Multi-Health Systems.

Hart, S. D., Kropp, P. R., Laws, D. R., Klaver, J., Logan, C., & Watt, K. A. (2003). *The risk for sexual violence protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence*. Burnaby, BC: Mental Health, Law and Policy Institute, Simon Fraser University.

Hawk, S.T., Keijsers, L., Branje, S.J.T., Van der Graaff, J., de Wied, M., & Meeus, W. (2013). Examining the Interpersonal Reactivity Index (IRI) among early and late adolescents and their mothers. *Journal of Personality Assessment, 95,* 96 – 106. doi: 10.1080/00223891.2012.696080

Hester, M. (2013). Who does what to whom? Gender and domestic violence perpetrators in English police records. *European Journal of Criminology, 10,* 623 – 637. doi: 10.1177/1477370813479078

Hoffman, M.L. (2000). *Empathy and Moral Development. Implications for caring and justice*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511805851

Hoffman, C., Crnic, K.A., & Baker, J.K. (2006). Maternal depression and parenting: Implications for children's emergent emotion regulation and behavioral functioning, *Parenting: Science and Practice, 6,* 271–295. doi:10.1207/s15327922par0604_1

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology, 33*, 307-316. doi:10.1037/h0027580

Hojat, M., Mangione, S., Kane, G.C., & Gonnella, J.S. (2005). Relationships between scores of the Jefferson scale of physician empathy (JSPE) and the interpersonal reactivity index (IRI). *Medical Teacher, 27,* 625-628. doi: 10.1080/01421590500069744

Holmes, T. H., & Rahe, R. M. (1967). The Social Readjustment Rating Scale. *The Journal of Psychosomatic Research, 11,* 213–218. doi:10.1016/0022-3999(67)90010-4

Holtzworth-Munroe, A., & Stuart, G. L. (1994). Typologies of male batterers: Three subtypes and the differences among them. *Psychological Bulletin, 116,* 476–497. doi: 10.1037/0033-2909.116.3.476

Home Office (2013). *Domestic Violence and Abuse: New Definition.* Retrieved from https://www.gov.uk/domestic-violence-and-abuse#domestic-violence-and-abuse-new-definition

Ickes, W. (2009). Empathic accuracy: links to clinical, cognitive, developmental, social and physiological psychology. In J. Decety & W. Ickes (Ed). *The Social Neuroscience of Empathy* (pp. 114 – 120). Cambridge, MA: MIT Press. doi: 10.7551/mitpress/9780262012973.001.0001

Ireland, J.L. (1999). Provictim attitudes and empathy in relation to bullying behaviour among prisoners. *Legal and Criminological Psychology, 4,* 51 – 66. doi: 10.1348/135532599167789

Jaffee, S.R., Capia, A., Moffitt, T.E., & Taylor, A. (2004). Physical maltreatment victim to antisocial child: evidence of an environmentally mediated process. *Journal of Abnormal Psychology, 113,* 44 – 55. doi:10.1037/0021-843X.113.1.44

Jenkins, C.R., & Dillman, D.A. (1995). Towards a theory of self-administered questionnaire design. In L. Lyberg et al. (Eds.). *Survey Measure and Process Quality* (pp. 165-196). New York: Wiley. doi: 10.1002/9781118490013.ch7

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

Joliffe, D., & Farrington, D.P. (2007). Examining the relationship between low empathy and self-reported offending. *Legal and Criminological Psychology, 12,* 265 – 286. doi: 10.1348/135532506X147413

Jöreskog, K.G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika, 42*, 183 – 202. doi: 10.1007/BF02289343

Juni, P., Witschi, A., Bolsch, R., & Egger, M. (1999). The hazards of scoring the quality of clinical trials for meta-analysis. *The Journal of the American Medical Association, 11,* 1054 – 1060. doi:10.1001/jama.282.11.1054

Kahng, S.K., Oyserman, D., Bybee, D., & Mowbray, C. (2008). Mothers with serious mental illness: when symptoms decline does parenting improve? *Journal of Family Psychology, 22,* 162–66. doi:10.1037/0893-3200.22.1.162

Kendall-Tackett, K.A., & Eckenrode, J. (1996). The effects of neglect on academic achievement and disciplinary problems: A developmental perspective. *Child Abuse and Neglect, 20*,161–169. doi: S0145-2134(95)00139-5

Kim, J., Choi, S.S., Rhee, M.S., Kim, S.B., Joung, J.S., & Kim, E.H. (2012). Effect of sex offender treatment program on cognitive and emotional characteristics of mentally ill sex offenders. *Journal of Forensic Science, 57,* 1608 – 1613. doi: 10.1111/j.1556-4029.2012.02222.x

Klein, K.J., & Hodges, S.D. (2001). Gender differences, motivation and empathic accuracy: when it pays to understand. *Personality and Social Psychology Bulletin, 27*, 720 – 730. doi:10.1177/0146167201276007

Kline, P. (1986). *Handbook of test construction.* London: Methuen.

Kmet, L. M., Lee, R. C., & Cook, L., S. (2004). *Standard quality assessment criteria for evaluating primary research papers from a variety of fields.* Canada: Alberta Heritage Foundation For Medical Research, Retrieved from http://www.crd.york.ac.uk/crdweb/ShowRecord.asp?LinkFrom=OAI&ID=32004000313#.VUfkQ_lVikp

Knafo, A., C. Zahn-Waxler, C. Van Hulle, J. L. Robinson, & S. H. Rhee (2008). The developmental origins of a disposition toward empathy: Genetic and environmental contributions. *Emotion, 8,* 737-752.

Konrath, S.H., O'Brien, E.H., & Hsing, C. (2011). Changes in dispositional empathy in American college students over time: a meta-analysis. *Personality and Social Psychology Review, 15,* 180 – 198. doi: 10.1177/1088868310377395

Kropp, P.R., Hart, S.D., Webster, C.D., & Eaves, D. (1995). *Manual for the Spousal Assault Risk Assessment Guide (2nd Edition).* Vancouver: British Columbia Institute Against Family Violence

Krug, E.G., Dahlberg, L.L., Mercy, J.A., Zwi, A.B., & Lozano, R. (2002). *World report on violence and health.* Geneva: World Health Organisation. Retrieved from http://www.ayamm.org/english/Violence%20against%20women%204.pdf

Lansdown, G. (2005). *The evolving capacities of the child*. Florence: Unicef Innocenti Research Centre. Retrieved from http://www.unicef-irc.org/publications/pdf/evolving-eng.pdf

Larson, M.J., Fair, J.E., Good, D.A. & Baldwin, S.A. (2010). Empathy and error processing. *Psychophysiology, 47,* 415 – 424. doi: 10.1111/j.1469-8986.2009.00949.x

Lauterbach, O., & Hosser, D. (2007). Assessing empathy in prisoners – A shortened version of the Interpersonal Reactivity Index. *Swiss Journal of Psychology, 66*, 91 – 101. doi: 10.1024/1421-0185.66.2.91

Lawrence, E.J., Shaw, P., Baker, D., Baron-Cohen, S., & Davis, A.S. (2004). Measuring empathy: reliability and validity of the empathy quotient. *Psychological Medicine, 34,* 911 – 924. doi: 10.1017/S0033291703001624

Leon, I., Rodrigo, M.J., Quinones, I., Hernandez, J.A., Lage, A., Pardon, I., Bobes, M.A. (2014). Electrophysiological responses to affect stimuli in neglectful mothers. PLoS ONE, 9(1). Retrieved from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0087808. doi:10.1371/journal.pone.0087808

Letourneau, C. (1981). Empathy and stress: How they affect parent aggression. *Social Work, 26,* 383 – 389. doi:10.1093/sw/26.5.383

Lindsey, R.E., Carlozzi, A.F., & Eells, G.T. (2001). Differences in the dispositional empathy of juvenile sex offenders, non-sex-offending delinquent juveniles and non-delinquent juveniles. *Journal of Interpersonal Violence, 16,* 510 – 552. doi: 10.1177/088626001016006002

Lipps, T. (1897). *Raumästhetik und geometrisch-optische Täuschungen.* Leipzig: JA Barth.

Litvack-Miller, W., McDougall, D., & Romney, D.M. (1997). The structure of empathy during middle childhood and its relationship to prosocial behavior. *Genetic Social and General Psychology Monographs, 123*, 303-324.

Marsh, A.A., & Blair, R.J.R. (2008). Deficits in facial affect recognition among antisocial populations: a meta-analysis. *Neuroscience and Biobehavioural Reviews, 32,* 454 – 465. doi: 10.1016/j.neubiorev.2007.08.003

Marshall, W. L., Hudson, S. M., Jones, R., & Fernandez, Y. M. (1995). Empathy in sex offenders. *Clinical Psychology Review, 15*, 99 - 113. doi: 10.1016/0272-7358(95)00002-7

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

Martin, G. B., & R. D. Clark (1982) . Distress crying in neonates: Species and peer specificity. *Developmental Psychology, 18*, 3-9.

Mayer, J.D., & Geher, G. (1996). Emotional intelligence and the identification of emotion. *Intelligence, 22*, 89-113. doi: 10.1016/S0160-2896(96)90011-2

McElroy, E.M., & Rodriguez, C.M. (2008). Mothers with externalizing behaviour problems: cognitive risk factors for abuse potential and discipline style and practices. *Child Abuse and Neglect, 32,* 774 – 784. doi:10.1016/j.chiabu.2008.01.002

McGrath, R. J., Cumming, G. F., Burchard, B. L., Zeoli, S., & Ellerby, L. (2010). *Current practices and trends in sexual abuser management: The Safer Society 2009 North American Survey.* Brandon, VT: Safer Society Press. doi:http://www.safersociety.org/uploads/WP141-Current_Practices_Emerging_Trends.pdf

McIntire, K. A., Danforth, M. M., & Schneider, H. G. (1997). *Measuring cue perception: Assessment of the Reliability and Validity*. Poster presented at the meetings of the Southeastern Psychological Association, Atlanta, GA.

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality, 4,* 525-554. doi:10.1111/j.1467-6494.1972.tb00078.x

Mennen, F.E., & Trickett, P.K. (2011). Parenting attitudes, family environments, depression, and anxiety in caregivers of maltreated children. *Family Relations, 60,* 259 – 271. doi:10.1111/j.1741-3729.2011.00646.x

Miller, P.A., & Eisenberg, N. (1988). The relationship of empathy to aggressive and externalising/antisocial behaviour. *Psychological Bulletin, 103,* 324 – 344. Retrieved from: http://psycnet.apa.org/index.cfm?fa=buy.optionToBuy&id=1989-15868-001

Milner, J. S. (1986). *The Child Abuse Potential Inventory: Manual* (2nd ed.). Webster, NC: Psyctec.

Milner, J. S. (1989). Additional cross-validation of the Child Abuse Potential Inventory. *Psychological Assessment, 1,* 219-223. doi:10.1037//1040-3590.1.3.219

Milner, J.S. (1993). Social information processing and physical child abuse. *Clinical Psychology Review, 13,* 275 – 294. doi:10.1016/0272-7358(93)90024-G

Milner, J.S. (2003). Social information processing in high-risk and physically abusive parents. *Child Abuse and Neglect, 27,* 7 – 20. doi:10.1016/S0145-2134(02)00506-9

Milner, J.S., Halsey, L.B., & Fultz, J. (1995). Empathic responsiveness and affective reactivity to infant stimuli in high- and low-risk for physical child abuse mothers. *Child Abuse and Neglect. 19,* 767-780. doi:10.1016/0145-2134(95)00035-7

Ministry of Justice (2014). *Population Bulletin: Monthly – March 2014*. Retrieved from https://www.gov.uk/government/statistics/prison-population-figures-2014.

Mulloy, R., Smiley, W.C., & Mawson, D.L. (1999). The impact of empathy training on offender treatment. *Focus on Corrections Research, 11*, 15 – 18. Retrieved from http://www.csc-scc.gc.ca/research/forum/e111/e111d-eng.shtml

National Institute of Clinical Excellence (2009). *Costing statement: When to suspect child maltreatment.* Retrieved from https://www.nice.org.uk/guidance/cg89/resources

National Society for the Prevention of Cruelty to Children (2013). *The criminal justice response to child sexual abuse.* Retrieved from: http://www.nspcc.org.uk/globalassets/documents/research-reports/how-safe-children-2014-criminal-justice-response-child-sexual-abuse.pdf

National Society for the Prevention of Cruelty to Children (2014). Assessing parenting capacity: an NSPCC factsheet. Retrieved from: http://www.nspcc.org.uk/globalassets/documents/information-service/factsheet-assessing-parenting-capacity.pdf

Nomura, K., & Akai, S. (2012). Empathy with fictional stories: reconsideration of the fantasy scale of the interpersonal reactivity index. *Psychological Reports, 110,* 304 – 314.

Nowicki, S., Jr. (2010). *The Diagnostic Analysis of Nonverbal Acuracy-2.* Unpublished Manuscript. Department of Psychology, Emory University, Atlanta, Georgia.

Nowicki, S., Jr., & Carton, J. (1993). The Measurement of Emotional Intensity from Facial Expressions: The DANVA FACES 2. *Journal of Social Psychology, 133*, 749-750. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/00224545.1993.9713934?journalCode=vsoc20

Nunnaly, J. (1978). *Psychometric theory.* New York: McGraw-Hill.

O'Brien, E.D., Konrath, S.H., Gruhn, D., & Hagen, A.L., (2012). Empathic concern and perspective taking: linear and quadratic effects of age across the adult life span. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences,* 1 – 8. doi: 10.1093/geronb/gbs055

Odgers, A. (2014). *From broken attachments to earned security: The role of empathy in therapeutic change*. London: Karnac.

Office for National Statistics (2011). *2011 Census: KS201UK Ethnic group, local authorities in the United Kingdom.* Retrieved from http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-327143.

Oliver, L.D., Mitchell, D.G.V., Dziobek, I., MacKinley, J., Coleman, K., Rankin, K.P., & Finger, E.C. (2015). Parsing cognitive and emotional empathy deficits for negative and positive stimuli in frontotemporal dementia. *Neuropsychologia, 75,* 14 – 26. doi: 10.1016/j.neuropsychologia.2014.11.022

Oswald, P. A. (2003). Does the Interpersonal Reactivity Index Perspective Taking scale predict who will volunteer time to counsel adults entering college? *Perceptual and Motor Skills, 97*, 1184-1186. doi: 10.2466/PMS.97.8.1184-1186

Paterson, H., Reniers, B., & Vollm, B. (2009). Personality types and mental health experiences of those who volunteer for helplines. *British Journal of Guidance and Counselling, 37*, 459-471. doi: 10.1080/03069880903161419

Paulhus, D. L. (1994). *Balanced Inventory of Desirable Responding: Reference manual for BIDR Version 6.* Unpublished manuscript, University of British Columbia, Vancouver, Canada

Paulhus, D. (1998). *Paulhus Deception Scales (PDS): The Balanced Inventory of Desirable Responding— 7: User's manual.* North Tanawanda, NY: Multi-Health Systems.

Pears, K.C., & Capaldi, D.M. (2001). Intergenerational transmission of abuse: a two-generational prospective study of an at-risk sample. *Child Abuse and Neglect, 11,* 1439 – 1461. doi:10.1016/S0145-2134(01)00286-1

Perez-Albeniz, A., & de Paul, J. (2003). Dispositional empathy in high- and low-risk parents for child physical abuse. *Child Abuse and Neglect, 27,* 769-780. doi:10.1016/S0145-2134(03)00111-X

Perez-Albeniz, A., & de Paul, J. (2004). Gender differences in empathy in parents at high- and low-risk of child physical abuse. *Child Abuse and Neglect, 28,* 289-300. doi:10.1016/j.chiabu.2003.11.017

Perry, D.G., & Perry, L.C. (1974). Denial of suffering in the victim as a stimulus to violence in aggressive boys. *Child Development, 45,* 55 – 62. doi: 10.1111/j.1467-8624.1974.tb00558.x

Piaget, J. (1965). *The moral judgment of the child*. New York: Free Press.

Plotkin, R. (1983). *Cognitive mediation in disciplinary actions among mothers who have abused or neglected their children: Dispositional and environmental factors.* Unpublished manuscript, University of Rochester, New York, America.

Polaschek, D. L. L. (2003). Empathy and victim empathy. In T. Ward, D. R. Laws, & S. M. Hudson (Eds.), *Sexual deviance: Issues and controversies* (pp. 172 – 189). Thousand Oaks, CA: Sage Publications. doi:10.4135/9781483328751.n10

Radford, L., Corral, S., Bradley, C., Fisher, H., Bassett, C., Howat, N., & Collishaw, S. (2011). *Child abuse and neglect in the UK today.* London: NSPCC. Retrieved from: https://www.nspcc.org.uk/Inform/research/findings/child_abuse_neglect_research_PDF_wdf84181.pdf

Radloff, L.S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1,* 385-401. doi: 10.1177/014662167700100306

Respect. (2009). Position statement: Gender and domestic violence. Retrieved from http://www.respect.uk.net/data/files/respect gender dv position satatement.doc

Robinson, L., Spencer, M.D., Rhomson, L.D.G., Sprengelmeyer, R., Owens, D.G.C., Stanfield, A.C., … Johnstone, E.C., (2012). Facial emotion recognition in scottish prisoners. *International Journal of Law and Psychiatry, 35,* 57–61. doi: 10.1016/j.ijlp.2011.11.009

Robyn, S., & Fremouw, W.J. (1996). Cognitive and affective styles of parents who physically abuse their children. *American Journal of Forensic Psychology, 14*, 63-79. Retrieved from http://www.forensicpsychology.org/index.htm

Rodriguez, C.M. (2013). Analog of parental empathy: Association with physical child abuse risk and punishment intentions. *Child Abuse and Neglect, 37,* 493 - 499. doi:10.1016/j.chiabu.2012.10.004

Rodriguez, C.M., Cook, A.E., & Jedrziewski, C.T. (2012). Reading between the lines: Implicit assessment of the association of parental attributions and empathy with abuse risk. *Child Abuse and Neglect, 36,* 564-571. doi:10.1016/j.chiabu.2012.05.004

Rodriguez, C.M., & Richardson, M.J. (2007). Stress and anger as contextual factors and pre-existing cognitive schemas: predicting parental child maltreatment risk. *Child Maltreatment, 12,* 325 – 337. doi:10.1177/1077559507305993

Rodriguez, C.M., & Tucker, M.C. (2014). Predicting maternal physical child abuse risk beyond distress and social support: Additive role of cognitive processes. *Journal of Child and Family Studies, 24,* 1780 – 1790. doi: 10.1007/s10826-014-9981-9

Rogers, C. R. (1975). Empathic: An unappreciated way of being. *Counselling Psychologist, 5,* 2-10. doi:10.1177/001100007500500202

Rogers, K., Dziobrek, I., Hassenstab, J., Wolf, O.T., & Convit, A. (2007). Who cares? Revisiting empathy in Asperger syndrome. *Journal of Autism and Developmental Disorders, 37,* 709 – 715. doi: 10.1007/s10803-006-0197-8

Romero-Martínez, A., Lila, M., Sariñana-González, P., González-Bono, E., & Moya-Albiol, L. (2013). High testosterone levels and sensitivity to acute stress in perpetrators of domestic violence with low cognitive flexibility and impairments in their emotional decoding process: A preliminary study. *Aggressive Behaviour, 39,* 355 – 369. doi: 10.1002/ab.21490

Rosenstein, P. (1995). Parental levels of empathy as related to risk assessment in child protective services. *Child Abuse and Neglect, 19,* 1349-1360. doi:0.1016/0145-2134(95)00101-D

Russell, D., Springer, K., & Greenfield, E. (2010). Witnessing domestic violence in childhood as an independent risk factor for depressive symptoms in young adulthood. *Child Abuse and Neglect, 34,* 448- 453. doi: 10.1016/j.chiabu.2009.10.004

Sackett, D.L. (1979). Bias in analytical research. *Journal of Chronic Disease, 32,* 51 – 63. doi:10.1016/0021-9681(79)90012-2

Sanders, M.R., Pidgeon, A.M., Gravestock, F., Connors, M.D., Brown, S., & Young, R.W. (2004). Does parental attributional retraining and anger management enhance the effects of the Triple P-Positive Parenting Program with parents at risk of child maltreatment? *Behavior Therapy, 35,* 513-535. doi:10.1016/S0005-7894(04)80030-3

Sayers, A., (2007). Tips and tricks in performing a systematic review. *The British Journal of General Practice, 57,* 425. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2151802/pdf/bjpg57-759.pdf

Schweinle, W.E., & Ickes, W. (2007). The Role of Men's Critical/Rejecting Overattribution Bias, Affect, and Attentional Disengagement in Marital Aggression. *Journal of Social and Clinical Psychology, 26,* 173-198. doi: 10.1521/jscp.2007.26.2.173

Schweinle, W.E., Ickes, W., & Bernstein, I.H. (2002). Empathic inaccuracy in husband to wife aggression: the overattribution bias. *Personal Relationships*, *9*, 141 – 158. doi: 10.1111/1475-6811.00009

Seidel, E., Pfabigan, D.M., Keckeis, K., Wucherer, A.M., Jahn, T., Lamm, C., & Derntl, B. (2013). Empathic competencies in violent offenders. *Psychiatry Research, 210,* 1168 - 1175. doi: 10.1016/j.psychres.2013.08.027

Siu, M.H., & Shek, D.T.L. (2005). Validation of the interpersonal reactivity index in a Chinese context. *Research on Social Work Practice, 15,* 118 – 126. doi; 10.1177/1049731504270384

Smith, K., Osborne, S., Lau, I., & Britton, A. (2012). *Homicides, firearm offences and intimate violence 2010/11: supplementary volume 2 to Crime in England and Wales 2010/11*. London: Home Office. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/116483/hosb0212.pdf

Song, F.J., Parekh-Bhurke, S., Hooper, L., Loke, Y.K., Ryder, J.J., Sutton, A.J., … Harvey, I. (2009). Extent of publication bias in different categories of research cohort: a meta-analysis of empirical studies. *BMC Medical Research Methodology, 9,* 79. doi:10.1186/1471-2288-9-79

Spell, L. A. (1996). *Recognition of the Nonverbal Communication of Affect Following Traumatic Brain Injury*. Unpublished manuscript, Department of Speech-Language Pathology and Audiology, School of Public Health, University of South Carolina, Columbia.

Spielberger, C. D. (1988). *Manual for the State-Trait Anger Expression Inventory (STAXI).* Odessa, FL: Psychological Assessment Resources.

Straus, M.A., Hamby, S.L., Finkelhor, D., Moore, D.W., & Runyan, D. (1998). Identification of child maltreatment with the parent-child conflict tactics scales: Development and

psychometric data for a national sample of American parents. *Child Abuse and Neglect, 22,* 249–270. doi:10.1016/S0145-2134(97)00174-9

Straus, M.A., Hamby, S.L., McCoy, S.B., & Sugarman, D.B. (1996). The Revised Conflict Tactics Scales (CTS2). Development and preliminary psychometric data. *Journal of Family Issues, 17,* 283 – 316. doi: 10.1177/019251396017003001

Strayer, J. (1987). Affective and cognitive perspectives on empathy. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its Development* (pp. 47-80). Cambridge: Cambridge University Press.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643 – 662.

Thompson, R., Jones, D.J., Litrownik, A.J., English, D.J., Kotch, J.B., Lewis, T., & Dubowitz, H. (2014). Linking mother and youth parenting attitudes: indirect effects via maltreatment, parent involvement, and youth functioning. *Child Maltreatment, 19,* 23 – 246. doi: 10.1177/1077559514547263

Titchener, E.B. (1909). *Lectures on the Experimental Psychology of Thought Processes*. New York: Macmillan. Retrieved from http://psycnet.apa.org/books/10877/005

Tomarken, A.J., & Selin, R.C. (1986). Comparison of ANOVA Alternatives under Variance Heterogeneity and Specific Noncentrality Structures. *Psychological Bulletin, 99,* 90 – 99. doi: 10.1037/0033-2909.99.1.90

Verbeek, P. (1996). *Peacemaking in Young Children.* Unpublished manuscript, Department of Psychology, Emory University, Atlanta.

Verhofstadt, L.L., Buysse, A., Ickes, W., Davis, M., & Devoldre, I. (2008). Support provision in marriage: The role of emotional matching and empathic accuracy. *Emotion, 8,* 792 – 802. doi: 10.1037/a0013976

Vickers, A., & Smith, C. (2000). Incorporating data from dissertations in systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews on Evidence-Based Nursing, 1*, 165 – 184. Retrieved from http://onlinelibrary.wiley.com/journal/10.1111/%28ISSN%291741-6787

Vollm, B.A., Taylor, A.N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., Deakin, J.F., & Elliot, R. (2006). Neural correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a non verbal task. *Neuroimage, 29,* 90 – 98. doi:10.1016/j.neuroimage.2005.07.022

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., & Weil, L. (2006). Development of short forms of the Empathy Quotient (EQ-Short) and the Systemizing Quotient (SQ-Short). *Personality and Individual Differences, 41,* 929 – 940. doi:10.1016/j.paid.2006.03.017

Walby, S. (2009). *The Cost of Domestic Violence Update 2009*. Unpublished Manuscript, Lancaster University. Retrieved from www.lancs.ac.uk/fass/doc_library/sociology/Cost_of_domestic_ violence_update.doc

Ward, T., & Durrant, R. (2013). Altruism, empathy and sex offender treatment. *International Journal of Behavioural Consultation and Therapy, 8,* 66 - 71. doi: 10.1037/h0100986

Wastell, C.A., Cairns, D., & Haywood (2009). Empathy training, sex offenders and re-offending. *Journal of Sexual Aggression: An international, interdisciplinary forum for research, theory and practice, 15,* 149 – 159. doi: 10.1080/13552600902792599

Welch, B.L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika, 38,* 330 – 336. doi: 10.2307/2332579

Wiehe, V.R. (1985). Empathy and locus of control in child abusers. *Journal of Social Service Research, 9,* 17 – 30. doi:10.1300/J079v09n02_02

Wiehe, V.R. (2003). Empathy and narcissism in a sample of child abuse perpetrators and a comparison sample of foster parents. *Child Abuse and Neglect, 25,* 541-555. doi:10.1016/S0145-2134(03)00034-6

Wilcox, R.R. (2005). *Introduction to Robust Estimation and Hypothesis testing (2nd Ed.).* Burlington, MA: Elsevier.

Wilhelm, M. O., & Bekkers, R. (2010). Helping behavior, dispositional empathic concern, and the principle of care. *Social Psychology Quarterly, 73,* 11-32. doi: 10.1177/0190272510361435

Williams, K., Papadopoulou, V., & Booth, N. (2012). *Prisoners' childhood and family backgrounds: Results from the Surveying Prisoner Crime Reduction (SPCR) longitudinal cohort study of prisoners.* Ministry of Justice. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/278837/prisoners-childhood-family-backgrounds.pdf

Windham, A.M., Rosenberg, L., Fuddy, L., McFarlane, E., Sia, C., & Duggan, A.K. (2004). Risk of mother-reported child abuse in the first 3 years of life. *Child Abuse and Neglect*, *28,* 647 – 669. doi:10.1016/j.chiabu.2004.01.003

World Health Organisation (2013). *European report of preventing child maltreatment*. Retrieved from: http://www.euro.who.int/__data/assets/pdf_file/0019/217018/European-Report-on-Preventing-Child-Maltreatment.pdf

World Health Organisation (2014). *Child Maltreatment Fact Sheet 150.* Retrieved from: http://www.who.int/mediacentre/factsheets/fs150/en/

World Health Organisation (2015). *Chapter 4: Violent Crime and Sexual Offences - Intimate Personal Violence and Serious Sexual Assault.*. Retrieved from http://www.ons.gov.uk/ons/dcp171776_394500.pdf

Yamada, M., & Decety, J. (2009). Unconscious affective processing and empathy: An investigation of subliminal priming on the detection of painful facial expressions. *Pain, 143,* 71 – 75. doi:10.1016/j.pain.2009.01.028

Yarnold, P., Bryant, F., Nightingale, S., & Martin G. (1996). Assessing physician empathy using the Interpersonal Reactivity Index: A measurement model and cross-sectional analysis. *Psychology Health and Medicine, 1,* 207-221. doi: 10.1080/13548509608400019

Zahn-Waxler, C., J. L. Robinson, & R. N. Emde. (1992). The development of empathy in twins. *Developmental Psychology,* 28, 1038-1047.

Appendix 1: Risk and protective factors for child maltreatment, as identified by the Child Welfare Information Gateway (2004).

Ecological level	Risk factors	Protective factors
Individual	premature birth low birth weight antenatal exposure to toxins temperament: overly excitable or slow to warm up physical/cognitive/emotional disability serious illness childhood trauma anti-social peer group age aggressive behaviour attention deficits	good health and development above-average intelligence hobbies and interests good peer relationships personality factors easy temperament positive disposition active coping style positive self-esteem good social skills internal locus of control
Family	parental personality factors external locus of control poor impulse control depression/anxiety low tolerance for frustration feelings of insecurity; low self-esteem lack of trust insecure attachment with own parents childhood history of abuse high parental conflict, domestic violence family structure single parenthood high number of children social isolation, lack of support parental psychopathology substance abuse separation/divorce – especially high conflict divorce age high stress poor parent-child interaction negative attitude towards child inaccurate knowledge and expectations about child development	secure attachment warm parent-child relationship supportive family environment household rules parental monitoring extended family support and involvement, including childcare stable relationship with child good parental coping skills family expectations of pro-social behaviour high parental education
Social/ Environmental	low socio-economic status stressful life events lack of access to social support, including child care and social services parental unemployment homelessness social isolation/lack of social support exposure to racism/discrimination poor schools poor housing exposure to environmental toxins dangerous/violent neighbourhood community violence	mid to high socio-economic status access to health care and social services consistent parental employment adequate housing family religious faith participation good schools supportive adults outside of family

Appendix 2: Details of Database Search Strategies

2.1 PsycINFO (OVID) 1806 to April Week 3 2015

1	Empathy	9565
2	Empath*	23973
3	"Perspective taking"	2907
4	Child Abuse	23897
5	Emotional Abuse	2066
6	Physical Abuse	5131
7	Sexual abuse	23317
8	Child Neglect	3339
9	Child* abus*	27197
10	Emotion* abus*	3398
11	Child* neglect*	3752
12	Physical* abus*	8963
13	Sex* abus*	23633
14	Abus*	141435
15	Neglect*	33869
16	Parents	74532
17	Fathers	8943
18	Mothers	33744
19	Caregivers	20279
20	Parent*	227700
21	Father*	40508
22	Mother*	107172
23	Mom*	28701
24	Mum*	1158
25	Dad*	1698
26	Caregiver*	39280
27	1 OR 2 OR 3	25963
28	4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15	171329
29	16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26	353991
30	28 AND 29 AND 30	448
	Remaining after initial reading of abstracts	46

2.2 PubMed

(empathy* or "perspective") **AND** ("child* abus*" OR "emotion* abus*" OR "physical abus*" OR "sex* abus*"OR "child* neglect*" OR "emotion* neglect*" OR abus* OR neglect*) **AND** (parent* OR father* OR mother* or caregiver* OR carer* OR mom* OR mum* OR mommy OR mummy OR dad* OR daddy)

564 Results

27 Remaining after initial reading of abstracts

2.3 Web of Science

Topic=(empathy* OR perspective*) **AND** topic=("child* abus*" OR "emotion* abus*" OR "physical abus*" OR "sex* abus*"OR "child* neglect*" OR "emotion* neglect*" OR abus* OR neglect*) **AND** Topic=(parent* OR father* OR mother* or caregiver* OR carer* OR mom* OR mum* OR mommy OR mummy OR dad* OR daddy)

1252 Results
19 Remaining after initial reading of abstracts

2.4 Cochrane

(empathy* or "perspective") **AND** ("child* abus*" OR "emotion* abus*" OR "physical abus*" OR "sex* abus*"OR "child* neglect*" OR "emotion* neglect*" OR abus* OR neglect*) **AND** (parent* OR father* OR mother* or caregiver* OR carer* OR mom* OR mum* OR mommy OR mummy OR dad* OR daddy)

5 Results
0 remaining after initial reading of abstracts

2.5 Science Direct

ALL FIELDS=(empathy* OR perspective*) **AND** ALL FIELDS =("child* abus*" OR "emotion* abus*" OR "physical abus*" OR "sex* abus*"OR "child* neglect*" OR "emotion* neglect*" OR abus* OR neglect*) **AND** ALL FIELDS = (parent* OR father* OR mother* or caregiver* OR carer* OR mom* OR mum* OR mommy OR mummy OR dad* OR daddy)

6865 Results
11 remaining after initial reading of abstracts

2.6 Medline (OVID) 1946 to April Week 3 2015

1	Empathy	13223
2	Empath*	17281
3	"Perspective taking"	737
4	Child Abuse	24734
5	Emotional Abuse	0
6	Physical Abuse	0
7	Sexual Abuse	18762
8	Child Neglect	24734
9	Child* abus*	26628
10	Emotion* abus*	1003
11	Child* neglect*	497

12	Physical* abus*	3667
13	Sex* abus*	9405
14	Abus*	134785
15	Neglect*	34382
16	Parents	75385
17	Fathers	6438
18	Mothers	27458
19	Caregivers	22772
20	Parent*	340377
21	Father*	32651
22	Mother*	165038
23	Mom*	51552
24	Mum*	103067
25	Dad*	10226
26	Caregiver*	40699
27	1 OR 2 OR 3	17746
28	4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15	171827
29	16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26	678414
30	25 AND 26 AND 27	146
	Remaining after initial reading of abstracts	22

2.7 ProQuest (Applied Social Science Index & Abstracts (1987 – current) , ProQuest Nursing & Allied Health Source, Social Services Abstracts (1979 – current), Sociological Abstracts (1952 – current))

ALL FIELDS=(empathy* OR perspective*) **AND** ALL FIELDS =("child* abus*" OR "emotion* abus*" OR "physical abus*" OR "sex* abus*"OR "child* neglect*" OR "emotion* neglect*" OR abus* OR neglect*) **AND** ALL FIELDS = (parent* OR father* OR mother* or caregiver* OR carer* OR mom* OR mum* OR mommy OR mummy OR dad* OR daddy)

1315 results
17 remaining after initial reading of abstracts

Appendix 3: Email template for contacting researchers, and a list of researchers.

Email template

Dear *(Researcher)*,

I am currently conducting a systematic literature review for my doctoral project at the University of Birmingham UK, looking at whether there are empathy differences between maltreating and non-maltreating parents.

I am interested in an article you published to include in my review:

(Reference of article)

I have been unable to access the article online or through my library services, thus am emailing to inquire as to whether you would be able to provide me with a copy of the article? I would also be interested in the details of any other studies you are aware of that are relevant to my review.

Thank you for taking the time to consider my request

Yours Sincerely

Elizabeth Fitzmaurice

Researchers contacted

Dr Paul Miller, Associate Professor of Psychology, Department of Social & Behavioural Sciences, Arizona State University. Author of: Miller, P.A., & Eisenberg, N. (1988). The relation of empathy to aggressive and externalizing/antisocial behaviours. *Psychological Bulletin, 103,* 324 – 344.

Dr Vernon Wiehe, Now retired, Previously Professor of Child Welfare in the College of Social Work at the University of Kentucky. Author of: Wiehe, V.R. (1985). Empathy and locus of control in child abusers. *Journal of Social Service Research, 9,* 17 – 30.

Dr David Fontaine, Author of: Fontaine, D., & Nolan, P. (2012) Study of "hot" executive functions in a sample of parents who have been accused of physical abuse or Systematic search of online databases neglect. *Journal of Aggression, Maltreatment & Trauma. 21,* 1-18.

Appendix 4: Articles that were not available and were received via post from authors

Name	Action Taken
Wiehe, V.R. (1985). Empathy and locus of control in child abusers. *Journal of Social Service Research, 9,* 17 – 30.	• Searched online databases • Search performed by librarian • Inter-library loan requested to the British Library, copy has not been returned in time for current submission, but will be included in the final published version • Emailed author, articles sent via airmail

Appendix 5: Articles searched in full that did not meet the inclusion/exclusion criteria

Article Reference	How article was found	Reason for Being Excluded from the Review
Asla, N., de Paul, J., Perez-Albinez, A. (2011). Emotion recognition in fathers and mothers at high-risk for child physical abuse. *Child Abuse and Neglect, 35*, 712 – 721	Systematic search of online databases	The study looks at emotional recognition in parents at high- and low-risk of abuse, and not at empathy explicitly.
Baranowski, M.D., Schilmoeller, G.L., & Higgins, B.S. (1990). Parenting attitudes of adolescent and older mothers. *Adolescence, 25,* 781-790.	Systematic search of online databases	Compared empathy between adolescent and older mothers. The study did not focus on empathy differences between maltreating and non-maltreating parents.
Barr, R.G., Fairbrother, N., Pauwels, J., Green, J.M., Chen, M., & Rollin, B. (2014). Maternal frustration, emotional and behavioural responses to prolonged infant crying. *Infant Behaviour and Development, 37,* 652 – 664.	Systematic search of online databases	This study looks at mother's reactions to hearing a baby cry for prolonged periods of time. It links responses to anger and frustration, but not specifically to child maltreatment.
Bartle-Haring, S., Slesnick, N., Jasmin, C. (2015). Reciprocity in Adolescent and Caregiver Violence. *Journal of Family Violence, 30*, 149 – 159.	Systematic search of online databases	Did not measure empathy
Bartlett, J. D., & Easterbrooks, M.A. (2015). The moderating effect of relationships on intergenerational risk for infant neglect by young mothers. Child Abuse & Neglect, In Press, http://dx.doi.org/10.1016/j.chiabu.2015.02.018	Systematic search of online databases	The aim of the study was to explore intergenerational transmission of maltreatment. Although the study uses measures of empathy with neglectful and non-neglectful mothers, results relating to empathy are not described in detail
Belsky, J. (1993). Etiology of child maltreatment: a developmental-ecological analysis. *Psychological Bulletin, 114,* 414 – 434.	Systematic search of online databases	The article is the description of a theoretical model, and is not a study.
Cyr, C., Michel, G., Dumais, M. (2013). Child maltreatment as a global phenomenon: from trauma to prevention. *International Journal of Psychology, 48,* 141- 148	Systematic search of online databases	This is a review of cultural factors linked to child maltreatment

De Paul, J., & Guibert, M. (2008). Empathy and child neglect: A theoretical model. *Child Abuse & Neglect, 32*, 1063-1071.	Systematic search of online databases	The article is the description of a theoretical model, and is not a study.
de Paul, J., Perez-Albeniz, A., Ormaechea, A., Vergara, A., de Cadiz, B., & Torres-Gomez (2006) Aggression Inhibition in High- and Low-Risk Subjects for Child Physical Abuse: Effects of a Child's Hostile Intent and the Presence of Mitigating Information. *Aggressive Behavior, 32,* 216-230.	Systematic search of online databases	The study does not measure empathy.
Ethier, L.S., Lacharite, C., & Couture, G. (1995). Childhood adversity, parental stress, and depression of negligent mothers. *Child Abuse and Neglect, 19*, 619 -632.	Systematic search of online databases	The study is in French, and does not measure empathy.
Farrant, B.M., Devine, T.A.J., Maybery, T. (2012). Empathy, perspective taking and prosocial behavious: the importance of parenting practices. *Infant and Child Development, 21,* 175 – 188.	Systematic search of online databases	The study investigates empathy in non-maltreating parents and measures the outcomes with relation to child behaviours.
Feshbach, N.D. (1989). *The construct of empathy and the phenomenon of physical maltreatment of children*. In D. Cicchetti, & V. Carlson (Eds) Child maltreatment: Theory and research on the causes and consequences of child abuse and neglect. (pp. 349-373). New York: Cambridge University Press	Systematic search of online databases	This is not a study, and does not focus on empathy in maltreating parents.
Fontaine, D., & Nolan, P. (2012). Study of "hot" executive functions in a sample of parents who have been accused of physical abuse or Systematic search of online databases neglect. *Journal of Aggression, Maltreatment & Trauma. 21,* 1-18.	Systematic search on online databases	The study did not use a measure of empathy.
Frodi, A.M., & Lamb, M.E. (1980). Child abusers' responses to infant smiles and cries, *Child Development, 51,* 238 – 421.	Hand searching articles	The study only examined parents' emotional responses to child faces, and did not relate this to empathy.
Gonzalez, A. (2015). The role of maternal executive function. *Canadian Psychology, 56*, 46 – 53.	Systematic search on online databases	This is a review article and does not have an empirical study.
Gordon, M. (2003). Roots of Empathy: responsive parenting, caring societies.	Systematic search of	The article details the delivery of a programme

The Keio Journal of Medicine, 52, 236 – 243.	online databases	and does not measure empathy in maltreating parents
Goubert, L., Vervoort, T., Cano, A., Crombez, G. (2009) Catastrophizing about their children's pain is related to higher parent-child congruency in pain ratings: An experimental investigation. *European Journal of Pain, 13,* 196-201	Systematic search of online databases	The study does not investigate empathy in maltreating parents.
Haskett, M.E., Allaire, J.C., Kreig, S., Hart, K.C. (2008). Protective and vulnerability factors for physically abused children: effects of ethnicity and parenting context. *Child Abuse and Neglect, 32,* 567 – 576.	Systematic search of online databases	The study investigates empathy in abused children, and does not measure empathy in parents.
Helfer, R.E. (1987) The perinatal period, a window of opportunity for enhancing parent-infant communication: an approach to prevention. *Child Abuse and Neglect, 11,* 565 – 579.	Systematic search on online databases	Systematic review and summary of intervention programmes in the perinatal period. Does not measure empathy in maltreating parents
Howes, C., & Feshback, N.D. (1986). Responses of abused, neglected and nonmaltreated children to the behAviors of their peers. *Journal of Applied Developmental Psychology, 6,* 261 – 270.	Hand searching articles	The article is an unpublished thesis.
Jones, D.P. (1995). Parental empathy, emotionality and the potential for child abuse. *Child Abuse and Neglect, 19,* 765 – 766.	Systematic search of online databases	The article is a review of Milner, Halsey & Fultz (1995) and not an experimental study
Kaminski, P.L., Hayslip, B.J., Wilson, J.L., & Casto, L.N. (2008) Parenting attitudes and adjustment among custodial grandparents. *Journal of Intergenerational Relationships, 6,* 263-284.	Systematic search of online databases	The study does not measure empathy.
Knox, J. (2013). "Feeling for' and Feeling with': developmental and neuroscientific perspectives on intersubjectivity and empathy. *Journal of Analytical Psychology, 58*, 491 – 509.	Systematic search of online databases	This is a review paper of psychopathology from early childhood and not an empirical research paper
Kropp, J.P. & Haynes, O.M. (1987). Abusive and nonabusive mothers' ability to identify general and specific emotion signals of infants. *Child Development, 58,* 187-190.	Systematic search on online databases	Measured parents' responses to emotional faces, and did not measure empathy explicitly.

Lutenbacher, M. (2002).Relationships between psychosocial factors and abusive parenting attitudes in low-income single mothers. *Nursing Research, 51,*158-167.	Systematic search on online databases	Does not explicitly measure empathy.
Lutenbacher, M., & Hall, L.A. (1998). The effects of maternal psychosocial factors on parenting attitudes of low-income, single mothers with young children. *Nursing Research, 47,* 25-34.	Systematic search on online databases	Does not explicitly measure empathy.
Matos, A.L., Moleiro, C., & Dias, J.G. (2014). Clusters of abusive parenting; A latent class analysis of families referred to Child Protection Services in Portugal. *Child Abuse and Neglect, 38,* 2053 – 2061.	Systematic search on online databases	Does not explicitly measure empathy.
Melnick, B. & Hurley, J.R. (1969). Distinctive Personality Attributes of Child-Abusing Mothers. *Journal of Consulting and Clinical Psychology, 33,* 746-749.	Hand search of empathy articles	Does not have a measure of empathy.
Meyers, S.A., & Battistoni, J. (2003) Proximal and distal correlates of adolescent mothers' parenting attitudes. *Journal of Applied Developmental Psychology*, *24,* 33-49.	Systematic search on online databases	Does not relate empathy to child maltreatment.
Miller, P.A., & Eisenberg, N. (1988). The relation of empathy to aggressive and externalizing/antisocial behaviours. *Psychological Bulletin, 103,* 324 – 344.	Hand searching articles	The article is a literature review.
Perez-Albeniz, A., & de Paul Ochotorena, J. (2005). Empathy in Individuals at Risk for Child Physical Abuse: The Effects of Victim's Pain Cues on Aggression. *Aggressive Behavior, 31*, 336-349.	Systematic search of online databases	Looked at the difference between high- and low-risk parents' responses when giving pain shocks to adult participants. Does not directly look at empathy towards children.
Perez-Albeniz, A., & De Paul, J. (2005). Empathy in individuals at risk for child physical abuse: the effects of victim's pain cues on aggression. *Aggressive Behaviour, 31,* 336 – 349.	Systematic search on online databases	Assesses empathy to adult victims suffering, and not empathy towards children.
Perez-Albeniz, A., & De Paul, J.	Systematic	Looked at the difference

(2006). Empathy and Risk Status for Child Physical Abuse: The Effects of an Adult Victim's Pain Cues and an Adult Victim's Intent on Aggression. *Aggressive Behavior,32,* 421-432.	search of online databases	between high- and low-risk parents' responses when giving pain shocks to adult participants. Does not directly look at empathy towards children.
Peterson, J., & Hawley, D.R. (1998). Effects of stressors on parenting attitudes and family functioning in a primary prevention program. *Family Relations, 47,* 221-228.	Hand searching empathy articles	Does not include a measure of empathy
Risser, H.J., Swowronski, J.J., & Crouch, J.L. (2011). Implicit attitudes toward children may be unrelated to child abuse risk, *Child abuse and neglect, 35,* 514 – 523.	Systematic search of online databases	The study did not look at empathy.
Rodriguez, G., Rodrigo, M.J., Janssens, J.M., Triana, B. (2011). Quality of maternal thinking and mother-child interactions in at-risk contexts. *Scandinavian Journal of Psychology, 52,* 545–552	Systematic search on online databases	Does not measure empathy
Shahar, G. (2001). Maternal personality and distress as predictors of child neglect. *Journal of Research in Personality, 35,* 537 – 545.	Hand searching of studies	The study does not measure empathy in maltreating parents.
Solis-Camara, R. P., & Diaz Romero, M. (1991). Validity of the Parenting Inventory for adults and adolescents: Indices of child abuse. *Salud Mental, 14,* 11-16.	Systematic search on online databases	The article is in Spanish
Turcotte, Y., Lussier, Y., Bertrand, J., & Perron, A. (1997). The empathy of incestuous fathers and nonabusive mothers. *Revue Quebecoise de Psychologie, 18,* 169-187.	Systematic search on online databases	The article is in French

N = 39 studies

Appendix 6: Quality assessment tool (case-series and case-control)

Quality assessment for case-series, utilising CASP model

Screening Questions	Yes	No	Can't tell
Did the study address a clearly focused issue?			
Did the authors use an appropriate method to answer their question?			
Is it worth continuing?			

	Detailed Questions	Yes (2)	Partially (1)	No (0)	Can't tell (?)
1	Are the maltreating parent population in the study representative of the average maltreating parent population? • Is the sample biased by low response rates? • Were participants recruited from one particular source (if the study is retrospective, and utilising child protection referrals, score as partially)				
2	Were there a sufficient number of cases selected? • Score yes if there were above 120 participants • Score partially if there were above 50 participants • Score no if there were below 50 participants				
3	Was the study design appropriate for answering the research question?				
4	Have the authors accounted for the important confounding factors • Race, sex, marital status, carer status, number of children, age, socio-economic status, previous contact with social services, education • Score yes if they were accounted for in analysis				
5	Was empathy accurately measured? • Was the measure of empathy subjective or objective? • Are the measures standardized, reliable and validated? • Were the measurement methods similar in the cases and controls • If questionnaire based, were questionnaires completed in front of assessors?				
6	Was maltreatment status clearly defined and assessed utilising a standardized, reliable, valid measure? • Score yes if participants were proven to be maltreating by legal proceedings • Score partially if participants were measured as high- and low-risk using reliable measured. • Score no if there was no a system in place for ensuring the maltreatment status of parents, other than assumption.				
	What are the results of the study, with relation to empathy and maltreatment status?				
7	Is the analysis appropriate to the design? • Has the study conducted an appropriate level of analysis				
8	Were the outcomes clearly described with relation to the research question				

Quality assessment for case-control, utilising CASP model

Screening Questions	Yes	No	Can't tell
Did the study address a clearly focused issue?			
Did the authors use an appropriate method to answer their question?			

Is it worth continuing?				

	Detailed Questions	Yes (2)	Partially (1)	No (0)	Can't tell (?)
1	Are the maltreating parent population in the study representative of the average maltreating parent population? • Is the sample biased by low response rates? • Were participants recruited from one particular source (if the study is retrospective, and utilising child protection referrals, score as partially)				
2	Are the control group appropriate comparators for the maltreating parent group?				
3	Was maltreatment status clearly defined and assessed utilising a standardized, reliable, valid measure? • Score yes if participants were proven to be maltreating by legal proceedings • Score partially if participants were measured as high- and low- risk using reliable measures. • Score no if there was not a system in place for ensuring the maltreatment status of parents, other than assumption.				
4	Were there a sufficient number of cases selected in both the target and control populations? • Score yes if there are > 60 participants for each group • Score partially if there are > 20 participants for each group • Score no if there are < 20 participants				
5	Was the study design appropriate for answering the research question?				
6	Have the authors accounted for the important confounding factors between-groups (e.g. using matched-pairs design) • Race, sex, marital status, carer status, number of children, age, socio-economic status, previous contact with social services, education • Score yes if most of these factors were considered pre-assessment or in analysis • Score partially if most of these factors were considered post-assessment or analysis • Score no if there are significant group differences that could account for the findings				
7	Was empathy accurately measured? • Was the measure of empathy subjective or objective? • Are the measured standardized, reliable and validated? • Were the measurement methods similar in the cases and controls?				
	What are the results of the study, with relation to empathy and maltreatment status?				
8	Is the analysis appropriate to the design? • Has the study conducted an appropriate level of analysis				
9	Were the outcomes clearly described with relation to the research question				

Appendix 7: Quality Assessment Results

Quality Assessment Results for Case-Control Studies										
Score	**Reference**	**Question**								
		1	2	3	4	5	6	7	8	9
12/18 **66.67%**	De Paul et al. (2008)	2	1	1 *non-maltreating were not assessed	1	2	0	2	1	2
16/18 **88.89%**	Francis & Wolfe (2008)	2	2	2	1	2	1 *some differences on previous abuse	2	2	2
14/18 **77.78%**	Letourneau (1981)	2	2	1 *non-maltreating were not assessed	1	2	2	1	1	2
12/18 **66.67%**	Leon et al., (2014)	0	1	1	1	2	1	2	2	2
17/18 **94.44%**	Mennen & Trickett (2011)	2	2	1 * non-maltreating not assessed	2	2	2	2	2	2
15/18 **83.33%**	Milner et al. (1995)	2	2	1	0	2	2	2	2	2
14/18 **77.78%**	Perez-Albinez & de Paul (2003)	1 *low response rate	2	1	1	2	2	1	2	2
15/18 **83.33%**	Perez-Albinez & de Paul (2004)	1 *low response rate	2	1	1	2	2	2	2	2
9/18 **50 %**	Robyn (1996)	0	2	1	0	2	1	1	1	1
12/18 **66.67%**	Wiehe (1986)	1	1	1 *non-maltreating not assessed	1	1	1	2	2	2
15/18 **83.33%**	Wiehe (2003)	2	1	1 * foster parents assumed to be non-maltreating	2	2	1 *significant difference in education	2	2	2

Quality Assessment Results for Case-Series Studies									
Score	**Reference**	**Question**							
		1	2	3	4	5	6	7	8
15/16 **93.75%**	McElroy & Rodriguez (2008)	2	1	2	2	2	2	2 *not good for our research question	2
12/16 **75%**	Rodriguez & Richardson (2007)	1* from one school	1	2	1	2	1	2 *not good for our research question	2
10/16 **62.5%**	Rodriguez et al. (2012)	1	0	2	1	2	1	1	2
13/16 **81.25%**	Rodriguez (2013)	2	2	2	2	1	1	1	2
13/16 **81.25%**	Rodriguez & Tucker (2014)	2	1	2	2	2	1	1 *has a global empathy score	2
11/16 **68.75%**	Rosenstein (1995)	2	0	2	0	2	2	1	2
12/16 **75%**	Thompson et al., (2014)	1	2	2	1	2	0	2 * not good for our research question	2

Appendix 8: Data Extraction Form

Author:

Date of publication:

Title:

Name of publication Source:

Population

Parenting status of maltreating population:

Mothers Fathers Step-mothers Step-fathers Other (please state) _____

Sample size of maltreating population:

Age of sample size at time of initial assessment:

How was maltreating status determined?:

Type of maltreatment:

How were parents recruited?:

Intervention

How was empathy measured?:

Is this a standardised measure?:

What aspect of empathy was measured?:

Comparator

How was non-maltreatment status determined?:

Size of comparator sample:

How were comparators recruited?:

What variables were considered (e.g. what variables were matched)?:

Outcome

What statistical technique was used?:

Was there a significant relationship between empathy outcome and maltreating status?

If yes, briefly describe the relationship:

What were the study conclusions, with relation to empathy and maltreating parents?:

What were the study limitations?:

Appendix 9: References of articles included in the review

1. de Paul, J., Perez-Albeniz, A., Guibert, M., Asla, N., & Ormaechea, A. (2008). Dispositional empathy in neglectful mothers and mothers at high risk for child physical abuse. *Journal of Interpersonal Violence, 23*, 670-684. doi:10.1177/0886260507313532.

2. Francis, K.J., & Wolfe, D.A. (2008). Cognitive and emotional differences between abusive and non-abusive fathers. *Child Abuse and Neglect, 32,* 1127-1137. doi:10.1016/j.chiabu.2008.05.007

3. Letourneau, C. (1981). Empathy and stress: How they affect parent aggression. *Social Work, 26,* 383 – 389. doi:10.1093/sw/26.5.383

4. Leon, I., Rodrigo, M.J., Quinones, I., Hernandez, J.A., Lage, A., Pardon, I., Bobes, M.A. (2014). Electrophysiological responses to affect stimuli in neglectful mothers. PLoS ONE, 9(1). Retrieved from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0087808. doi:10.1371/journal.pone.0087808

5. McElroy, E.M., & Rodriguez, C.M. (2008). Mothers with externalizing behaviour problems: cognitive risk factors for abuse potential and discipline style and practices. *Child Abuse and Neglect, 32,* 774 – 784. doi:10.1016/j.chiabu.2008.01.002

6. Mennen, F.E., & Trickett, P.K. (2011). Parenting attitudes, family environments, depression, and anxiety in caregivers of maltreated children. *Family Relations, 60,* 259 – 271. doi:10.1111/j.1741-3729.2011.00646.x

7. Milner, J.S., Halsey, L.B., & Fultz, J. (1995). Empathic responsiveness and affective reactivity to infant stimuli in high- and low-risk for physical child abuse mothers. *Child Abuse and Neglect. 19,* 767-780. doi:10.1016/0145-2134(95)00035-7

8. Perez-Albeniz, A., & de Paul, J. (2003). Dispositional empathy in high- and low-risk parents for child physical abuse. *Child Abuse and Neglect, 27,* 769-780. doi:10.1016/S0145-2134(03)00111-X

9. Perez-Albeniz, A., & de Paul, J. (2004). Gender differences in empathy in parents at high- and low-risk of child physical abuse. *Child Abuse and Neglect, 28,* 289-300. doi:10.1016/j.chiabu.2003.11.017

10. Rodriguez, C.M. (2013). Analog of parental empathy: Association with physical child abuse risk and punishment intentions. *Child Abuse and Neglect, 37*, 493 - 499. doi:10.1016/j.chiabu.2012.10.004

11. Rodriguez, C.M., Cook, A.E., & Jedrziewski, C.T. (2012). Reading between the lines: Implicit assessment of the association of parental attributions and empathy with abuse risk. *Child Abuse and Neglect, 36,* 564-571. doi:10.1016/j.chiabu.2012.05.004

12. Rodriguez, C.M., & Richardson, M.J. (2007). Stress and anger as contextual factors and pre-existing cognitive schemas: predicting parental child maltreatment risk. *Child Maltreatment, 12,* 325 – 337. doi:10.1177/1077559507305993

13. Rodriguez, C.M., & Tucker, M.C. (2014). Predicting maternal physical child abuse risk beyond distress and social support: Additive role of cognitive processes. *Journal of Child and Family Studies*, *24,* 1780 – 1790. doi: 10.1007/s10826-014-9981-9

14. Rosenstein, P. (1995). Parental levels of empathy as related to risk assessment in child protective services. *Child Abuse and Neglect, 19,* 1349-1360. doi:0.1016/0145-2134(95)00101-D

15. Thompson, R., Jones, D.J., Litrownik, A.J., English, D.J., Kotch, J.B., Lewis, T., & Dubowitz, H. (2014). Linking mother and youth parenting attitudes: indirect effects via maltreatment, parent involvement, and youth functioning. *Child Maltreatment, 19,* 23 – 246. doi: 10.1177/1077559514547263

16. Wiehe, V.R. (1985). Empathy and locus of control in child abusers. *Journal of Social Service Research, 9,* 17 – 30. doi:10.1300/J079v09n02_02

17. Wiehe, V.R. (2003). Empathy and narcissism in a sample of child abuse perpetrators and a comparison sample of foster parents. *Child Abuse and Neglect, 25,* 541-555. doi:10.1016/S0145-2134(03)00034-6

Appendix 10: Participant Information Sheet

INFORMATION SHEET

Research study – participants needed

A research study is being carried out by staff at Birmingham and Coventry Universities to work out whether individuals who have been convicted of either a violent or non-violent offence have different abilities to take someone else's perspective and identify emotions. You have been given this sheet as your most recent conviction was either for a violent or non-violent offence.

If you are interested in taking part, you would need to complete some questionnaires which would take about half an hour, and then do a quick computer-based task, which would take about 15 minutes. You do not have to take part. If you are unsure whether you want to take part and would like to talk to someone about the project in more depth before you decide, please contact [redacted] and he will arrange for you to meet with one of the researchers and have a chat at an appropriate time.

Whether you decide to take part or not, all of the information that you provide will be stored in a locked filing cabinet and in password protected computer files. Nobody will be able to link your answers to your name as your name will be replaced with a code number. This makes sure that your information is kept private and confidential. This information will be stored and kept in this way for 10 years, in accordance with data retention requirements. Your sentence and conditions will not be affected if you decide that you do, or do not, want to take part.

If you do take part, it is possible that you might find some questions a bit upsetting. If this is the case you can ask to not complete the questionnaire, or not answer all of the questions. If at any point you no longer want to take part you can stop, there will be no consequences for you if you should decide to stop at any point. You can also ask to have the information that you have provided withdrawn from the study at any point up to 2 weeks after you have taken part in the study. If you wish to discuss further or make a complaint about any part of the research, please contact [redacted]

You will not be asked to disclose detailed information (e.g., such as names of victims, dates and location of crimes) about any crimes you have committed during this study. If you do disclose identifiable details of any offence that you have not been convicted of (e.g., such as names of victims, dates and location of crimes) to the researcher they have a duty to report this information to prison staff.

It is hoped that the information you provide will be the first small part of a much larger study looking at whether emotion recognition and empathy play a role in offending behaviour. It is possible that the results will be published in an academic journal, or presented at a conference.

DEBRIEF SHEET

Thank you for taking part in our study.

There is a lot of research which suggests that violent offenders have problems taking the perspective of, or empathising with, other people, and that they also have difficulties identifying emotions. The study that you have taken part in has looked at this. You were recruited either because you do, or do not, have a conviction for a violent offence. At the moment very little is known about whether this is the case and therefore we cannot provide any individual feedback about scores. It is also unclear whether, if these problems are identified, they are actually of any relevance to offending or risk.

What we are trying to do is to provide some initial information so that we can expand the study in order to try and answer these questions. It is possible that in the future we may have enough information to be able to change the way that we work with offenders, but we really don't know yet. We would be happy to provide you with a summary of our research findings once they are completed, please contact the person listed below to access this in approximately 6 months time.

If you have any concerns about the study, or it has raised issues for you, please ask us know now. If you have concerns later, please contact [redacted] in the prison. They will be able to advise you on how to best seek support for any issues raised through the prison service.

Appendix 11: Participant Consent Form

CONSENT FORM

Emotion recognition and empathy study

Please initial box

1. I confirm that I have read and understood the participant information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these questions answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving a reason and at no consequence to myself. ☐
3. I understand that any information I provide will be kept confidential and that my identity will be kept anonymous. ☐
4. If I disclose identifiable details of any offence that I have not been convicted of (e.g., such as names of victims, dates and location of crimes) to the researcher, I understand that they have a duty to report this information to prison staff. ☐
5. I understand that this research has been approved by the University of Birmingham and Coventry University's ethics committee and as such is carried out according to preapproved ethical standards. ☐
6. I understand that that the information I provide may be used and analysed for research purposes and the findings may be published in an academic journal. ☐
7. I understand that I can request that any information I provide will be destroyed upon request. ☐
8. I agree to take part in the above study. ☐

_____	_____	_____
Name of Participant	Date	Signature