

**Predictors of child abuse potential in a sample of parents at risk of
maltreatment: the role of epistemic trust, parenting stress, and parental
representations**

Sharim Ponticelli

May 2023

*Research submitted in partial fulfilment of the requirements for the degree of Doctor
in Clinical Psychology (DClinPsy), Royal Holloway, University of London*

Acknowledgements

I would like to express gratitude to my two supervisors, Dr Michelle Slead and Dr Gary Brown, for their ongoing support and encouragement; I have really enjoyed learning from your respective fields of expertise. A special thank you to Dr Michelle Slead for having provided me with guidance in developing my thesis, warm encouragement throughout and additional training opportunities that have enriched both my research and clinical thinking. A special thank you also to Dr Gary Brown for having provided solid supervisory support in my hours of need. Thank you so much to the Royal Holloway course staff for always being kind and approachable, including Ms Deborah Phillips for her support in our (multiple) library sessions.

I am also grateful to the Supporting Parents Project team for having warmly welcomed me, showing me the ropes for the data collection and for having entertained stimulating and friendly discussions. I would like to extend gratitude to the small group of parents I had the opportunity to interview who shared their precious experiences with me. I would also like to thank the two Parent Champions from the Anna Freud Centre who kindly took time to learn about my project, showing refreshing enthusiasm and providing invaluable input.

To my fellow trainees, it has been such a pleasure to do this journey with you. Thank you for the support, humour, and wisdom that you have shared with me.

Finally, I would like to thank my wonderful partner Pablo, for being so supportive and fiercely proud and encouraging of my path, and amongst other things, providing me with your hard earned (delicious) takeaways and endless source of hugs. Thank you to my special Italian, English and Spanish family for believing in me, lightening my load and being a source of affection, encouragement, and giggles. Thanks so much to my lovely friends for also being there, dispatching much needed humour, and cheering me on. You have all helped see me through.

Table of Contents

Acknowledgements.....	2
Lay summary	4
Paper 1 - A systematic review on evaluating validity and responsiveness of the CAPI among maltreating parents of children aged 0-12	8
Abstract.....	8
Introduction.....	10
Method.....	16
Results.....	28
Discussion	45
Conclusion	59
Paper 2 - Predictors of child abuse potential in a sample of parents at-risk of maltreatment: the role of epistemic trust, parenting stress, and parental representations	60
Abstract.....	60
Introduction.....	62
Methods.....	71
Results.....	84
Discussion	89
Conclusion	104
Paper 3 - Integration, Impact and Dissemination.....	105
Integration.....	105
Impact.....	112
Dissemination	115
References	118
Appendix A	140
Appendix B	141
Appendix C	143
Appendix D	145
Appendix E	161
Appendix F.....	169
Appendix G	178
Appendix H	184
Appendix I.....	191
Appendix J.....	195

Lay summary

Systematic review

Parental physical child abuse is a devastating societal problem. It can have a range of physical and psychological consequences throughout life and has burdensome health and social care costs to society. The Child Abuse Potential Inventory (CAPI) is a tool that was designed for researchers and professionals working within child maltreatment to assess and prevent child abuse by providing an estimate of how likely caregivers are to physically mistreat their child. To ensure that the CAPI performs well and that it adequately detects child abuse risk among parents at high-risk of maltreating their child/ren, this review aimed to assess two of its properties: (1) 'validity', that is, whether the CAPI measures what it aims to measure and (2) 'responsiveness', which refers to whether the CAPI can detect changes in child abuse risk among high-risk parents who have taken part in parenting interventions aimed at reducing such risk. To address these questions, this review first assessed the quality of studies reporting on these two CAPI properties among high-risk and maltreating caregivers. Second, it examined whether study findings were in support of, or against CAPI validity and responsiveness. Last, it assessed the overall strength of the evidence obtained.

This review included twenty-eight studies published in English language journal articles between 1986 and January 2023. Most of the studies had been carried out in the USA, and over half included mothers only. Most articles (67%) reported on CAPI validity while the rest reported on its responsiveness. Articles reporting on CAPI validity were mostly of adequate if not very good quality, whilst just over half of those reporting on CAPI responsiveness were of adequate or very good quality. The study findings largely supported CAPI validity in three ways: first, CAPI scores were linked to scores from other instruments measuring similar factors (for example, parental distress, sensitivity, and empathy); second, CAPI scores consistently distinguished between groups of abusive and non-abusive caregivers; third, CAPI scores distinguished between other known caregiver groups (for example, parents with and without histories of childhood abuse). The findings for CAPI

responsiveness were mixed: whilst CAPI almost consistently detected small-to-medium decreases in child abuse risk after several types of parenting interventions, treatment differences were not detected to the same extent between caregivers who took part in parenting interventions compared to those in services as usual. Overall, the strength of evidence on CAPI validity was rated as moderate because although all studies reported data that could be used to understand CAPI's validity, this was not the primary aim for most studies, therefore weakening the evidence. The overall strength of evidence for CAPI responsiveness was rated as moderate too, though this was mainly due to inconsistent findings across studies.

These findings support the use of CAPI as a tool in research and to detect changes after parenting interventions, however more studies directly assessing CAPI validity and responsiveness are required and furthermore, among more diverse groups of maltreating parents to increase the evidence strength. To fully recommend its use in science and child protection settings, stronger ways of measuring how well the CAPI detects maltreatment are also recommended, such as assessing how strongly it associates with 'known instances of child physical abuse'.

Empirical project

Ample research suggests that individuals who were abused during childhood are more likely to maltreat their own children as adults. One suggestion for this 'cycle of violence' is that abuse and trauma in the parents' backgrounds interferes with the development of key social emotional competencies, thereby putting these parents at higher risk of repeating the abuse. One such competency refers to the capacity to be open to socially communicated knowledge, that is, 'epistemic trust'. The second, refers to the capacity to perceive others as psychological beings, and therefore make sense of their behaviour in terms of needs, wishes, desires and other psychological states, a capacity known as 'mentalization'. Researchers have argued that high-risk parents who struggle with

epistemic trust and mentalization are more likely to maltreat their child, if for instance, they are not able to trust and absorb what the child is communicating to them, or if they are unable to accurately read their child's behaviour and thereby jump to hostile conclusions. Also, high-risk parents are more likely to have parental stress, which is a well-established risk factor for increased risk of maltreatment.

This study, therefore, aimed to investigate whether low levels of mentalizing, parental stress and three different outlooks to socially communicated knowledge (epistemic trust, mistrust, and credulity) would be linked to the risk of parental child abuse among high-risk parents with histories of childhood maltreatment. The study used data that were collected as part of another study investigating the effectiveness of a parent programme for maltreating parents; the current study used data that were collected before the intervention occurred. One hundred and ten parents with children's social care involvement filled out questionnaires and were interviewed about a specific child, their relationship to that child, and how they view themselves as parents. Researchers then transcribed the interviews word-for-word and rated the interview material in terms of how risky parents' views were; that is, how much their views reflected an accurate representation of the child, showing capacity to mentalize and understand themselves and their children as psychological beings with needs, wishes, and desires, versus how much their views were distorted due to their own trauma backgrounds, preventing them from accurately and appropriately reading and supporting their child and relationship.

The study findings showed that whilst epistemic trust was not meaningfully linked to the risk of parents abusing their child, parents with higher levels of credulity and mistrust were at higher risk of mistreating their child. Additionally, the findings suggested that having more distorted and risky parent representations, which reflected low levels of mentalizing, was also associated with higher levels of child abuse risk. Finally, higher levels of parental stress were linked to higher levels of child abuse risk. When these parental risk factors were looked at together, as already seen in other studies, parental stress appeared to be

particularly important. Among the new parental factors tested here, the one that appeared to potentially be most important for child abuse risk was low parental mentalizing. Overall, these findings suggest that parental stress, epistemic mistrust, and credulity may be relevant parental factors to consider in the context of parents at high-risk of maltreatment. They also suggest that among these new characteristics tested here, low mentalizing may be particularly important.

This study advances the field of child maltreatment by identifying factors that link with parental child abuse risk, which had not been considered before. Some caution should be used when interpreting these findings as this study had several limitations. For this reason, future research would be beneficial, for example, to replicate the current study findings using larger samples, and following parents over time to detect possible mechanisms and causes. Nonetheless, the study is an important starting point for future investigations and has implications for clinical, social care and future policy.

A Systematic review on evaluating validity and responsiveness of the CAPI among maltreating parents of children aged 0-12

Abstract

Parental child physical abuse is a pervasive public health problem associated with numerous short- and long-term consequences. The Child Abuse Potential Inventory (CAPI) is one of the most widely used and validated measures in child protection and research settings to assess the caregiver's likelihood of physically abusing their child, which is referred to as Child Abuse Potential (CAP) and measured using the CAPI Abuse scale. Despite the widespread use of the CAPI Abuse scale, there is an informational gap concerning several of its psychometric properties, namely construct validity and responsiveness, which to date, have also not been assessed exclusively among high-risk caregivers. This systematic review, thus, aimed to evaluate CAPI Abuse scale construct validity and responsiveness among parents at high-risk of maltreating their children aged 0-12. Articles published in English between 1986 and January 2023 in the databases of PsycINFO, PubMed and Web of Science, and, which reported data on CAPI psychometric properties among maltreating caregiver samples, were evaluated using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) guidelines for systematic reviews of patient reported outcome measures. Twenty-eight articles directly and indirectly reported on construct validity (i.e., convergent, criterion, and known-group validity) and responsiveness. Most studies reported on convergent validity. Study quality was overall very good/adequate. Findings on CAPI validity and responsiveness were overall rated sufficient and insufficient, respectively. Validity findings revealed that CAP related to numerous theoretically similar constructs and crucially that CAP scores distinguish between maltreating and normative caregiver groups. Further CAPI was responsive to post-treatment changes for several tertiary interventions, though it was less sensitive to differences between intervention groups. The overall quality of validity and responsiveness evidence was moderate. Nonetheless, further psychometric testing and to a higher methodological

standard is required to determine CAPI Abuse scale performance among high-risk parent populations.

Introduction

For children across the world, few social problems cause greater harm to their health than parental child abuse and neglect. The potential for lifelong physical and emotional consequences is significant for all types of child maltreatment, not least child physical abuse (CPA; Centres for disease control and prevention, 2022). CPA is often conceptualised using the World Health Organization (WHO, 2020) definition as: “the intentional use of physical force against a child that results or has high likelihood of resulting in harm for the child’s health, survival, development or dignity”.

Despite worldwide recognition that CPA constitutes a violation of children’s rights (Gray et al., 2016; Pinheiro, 2006), reports indicate that it has reached epidemic levels. Indeed, estimates from a meta-analysis suggest that each year there are 41,000 homicide deaths of children under the age of 15 (WHO, 2022), and when informant and self-reports are combined, worldwide prevalence rates reach 22.6% (Stoltenborgh et al., 2015). These findings delineate the serious and pervasive nature of CPA at a global scale.

CPA prevalence data are even more alarming when considered together with its well-established and potentially tragic outcomes. A large body of evidence indicates that CPA can have a detrimental impact on all aspects of child development, thus posing victims at a higher risk of experiencing a broad range of difficulties across the lifespan. Findings from a systematic review and meta-analysis concluded that CPA (as well as child emotional abuse and neglect) is associated with several mental disorders, drug abuse, suicide attempts, sexually transmitted infections, risky sexual behaviour, and physical ill health throughout later life (Norman et al., 2012). Consequently, CPA also incurs enormous societal costs in the form of social and health care expenditures (Conti et al., 2017) and decreased quality of life (Wang & Holton, 2007). These findings have underscored the importance of understanding the mechanisms and correlates of child maltreatment.

Given the societal burden and global estimates of CPA, assessment tools that are capable of screening for potential CPA are crucial for several reasons. First, screening

measures aid the decision-making process in child protection settings; second, such tools help better define prevalence estimates and third, they provide helpful research instruments to evaluate the effectiveness of caregiver interventions and of policies aimed at reducing child maltreatment, as well as to identify CPA risk factors and putative causal pathways. In this vein, a significant number of screening instruments have been developed to assess parenting competency and to screen for maltreatment. Such measures include, for example, the Parenting Stress Index (PSI; Abidin, 1995), the Conflict Tactics Scale (CT; Straus, 1979), the Identification of Parents at Risk for Child Abuse and Neglect (IPARAN; Bouwmeester-Landweer, 2006) and the Child Abuse Potential Inventory (CAPI; Milner, 1980, 1986, 1994). Despite providing valuable information about parenting difficulties more broadly, no other instrument sensitively screens for specific forms of maltreatment, providing an estimate of risk, in the way that the CAPI was designed to do for CPA.

The Child Abuse Potential Inventory (CAPI)

The CAPI is one of the most frequently used scales in the scientific literature on child maltreatment (Georgieva et al., 2022), with over 300 studies documenting its use, and it is frequently adopted by clinicians in child protection settings (Laulik et al., 2015). The CAPI is a self-report instrument for parents and/or caregivers (these will be referred to here, interchangeably). The CAPI was constructed and validated in North America using samples that included caregivers suspected of CPA and known physical child abusers (Milner, 1986). As mentioned, the CAPI is the only screening measure to provide an estimate of caregiver's *probability of physically mistreating or abusing their child*, a construct referred to as Child Abuse Potential (CAP). The CAP is estimated by assessing female and male parents and primary caregivers' characteristics, namely psychological and interpersonal difficulties, which have been theoretically and empirically linked to individuals with a history of CPA perpetration (Milner, 1986). Indeed, elevated CAPI scores have been found to associate with adverse parental behaviour, such as punitiveness, rejection, irritability, and controllability

(Monroe & Schellenbach, 1989) and physical discipline (Chilamkurti & Milner, 1993). Parents with high CAPI scores have also been identified as having higher aggressiveness and a positive attitude toward physical punishment (Milner, 1994).

The CAPI comprises a 77-item Abuse scale and three validity scales from which three response distortion indices can be calculated. The Abuse scale score is used for evaluating CAP, which is derived from a weighted scoring procedure of its six factors: Distress, Rigidity, Unhappiness, Problems with Child and self, Problems with family, and Problems from others. The three validity scales (i.e., lie, random response, and inconsistency) and response distortion indices (i.e., faking good, faking bad, and random response) were designed to aid the detection of potentially abusive caregivers in situations in which they are keen to present positive impressions of themselves, for instance in the context of parenting capacity assessments. Research evidence indicates that the CAPI has reliable psychometric properties (Milner, 1986). The CAPI scales have obtained good internal consistency estimates across a variety of samples and consistent factorial structures (Walker & Davies, 2010). With regards to predictive validity, Walker and Davies (2010) concluded that there was little evidence available to evaluate it. A more recent psychometric review of CAPI psychometrics aimed at clinicians indicated that there continues to be a lack of evidence for CAPI's predictive validity owing to a dearth of prospective studies (Laulik et al., 2015). With regards to cross-cultural validity, however, translated versions of the CAPI have been successfully validated across cultures (Milner & Crouch, 2012).

Systematic review rationale

Despite the wealth of research relating to the CAPI and its frequent use in clinical settings, recently, a systematic review on child abuse screening tools concluded that further scrutiny of the CAPI's psychometric properties is necessary (Georgieva et al., 2022).

Focussing on the psychometric properties of the CAPI, as opposed to other child maltreatment measures, is important for several reasons. First, as mentioned, it is one of the

most recurrently used instruments in research and child protection settings. Second, it is one of the easiest instruments to administer owing to its 'yes', 'no' dichotomous format and year four reading level; this contributes to its useability in research and clinical settings. Third, its validity scales, aimed at assessing social desirability, make it especially valuable in child protection settings, where parents at high risk of maltreating their child/ren are keen to present themselves as 'good' and would otherwise go undetected. Last, it is the only measure to provide a risk estimate, which furthermore, is for a specific form of abuse.

In general, examining an instrument's measurement properties is important to the extent that the quality of a measure is largely dependent on its psychometric properties (Karanicolas et al., 2009). Broadly speaking, psychometric properties include the following: reliability (i.e., the extent to which scores for respondents who have not changed are the same for repeated assessments), validity (i.e., the extent to which an instrument measures the construct it purports to measure), and responsiveness (i.e., the ability of the instrument to detect change over time and between groups in the construct measured; Prinsen et al., 2018).

The already mentioned systematic review by Georgieva et al. (2022) examined the measurement properties of the CAPI along with four of the other most used instruments for measuring child maltreatment (i.e., Childhood Trauma Questionnaire, Maltreatment and Abuse Chronology of Exposure, and the Identification of Parents at Risk for child Abuse and Neglect, and Psychosocial Screening Tool). Georgieva and colleagues (2022) examined the following criteria: internal consistency, measurement error, content validity, structural validity, hypothesis testing, cross-cultural validity, criterion validity, responsiveness, and interpretability. From the only two studies they could find that directly examined the psychometric properties of CAPI, the reviewers reported that there was limited evidence of internal consistency, hypothesis testing for construct validity and criterion validity. This was because although the results related to these psychometric properties were evaluated as 'good', the methodological quality of the articles were deemed poor. Additionally, the review

identified a lack of information concerning the remaining CAPI measurement properties, namely its structural validity, test-retest reliability, measurement error, measurement invariance and responsiveness. In conclusion, among the five instruments, the CAPI was thus deemed as having the weakest evidence as only three of its psychometric properties could be examined, all of which had limited evidence.

Another issue concerns the target population in which the CAPI has been validated. The abovementioned, older critical review by Walker and Davies (2010) had examined CAPI measurement properties among studies using high-risk parent samples, however, it also included ones conducted in community samples. Therefore, to date, no review has focused on CAPI performance specifically among maltreating parents and caregivers. Accumulation and evaluation of evidence concerning the valid interpretation of data from the CAPI (i.e., its construct validity) for this specific population is important when considering the significant impact that actions taken in response to CAP score inferences have in terms of child safeguarding, family provision and for the development of tertiary prevention services aimed at children experiencing CPA. To this end, evaluating the CAPI's 'responsiveness' or treatment sensitivity, that is, its ability to detect changes in child abuse potential in high-risk caregivers undergoing treatment as part of parenting interventions (Youngstrom et al., 2017) is also paramount. Indeed, CAP scores are recurrently used in the effectiveness literature to assess the impact of parenting interventions (Georgieva et al., 2022; Slead, Fearon et al., 2021). Thus, examining the CAPI's sensitivity to treatment change would support future studies in validly drawing conclusions about the effectiveness of their treatments for high-risk parent populations. This is particularly important when considering the current limited availability of help for the most high-risk families in the UK (Alink, 2020; Barlow et al., 2006; Mulcahy et al., 2014).

Owing to these reasons, collection and analysis of validity and responsiveness evidence is necessary to argue for, or perhaps against, the proposed interpretation and use of CAP test scores among maltreating and high-risk parent populations. Hence, it is

important to assess whether the studies that examine and use CAPI possess the methodological quality and strength of evidence to support its psychometric properties.

Study aims

The overall aim of this systematic review was to critically appraise the methodological quality and psychometric properties of published research articles using the CAPI Abuse scale among maltreating or parents at high-risk of maltreating their children aged 0-12, by utilizing the COnsensus-based Standards for the Selection of health status Measurement INstruments (COSMIN) for systematic reviews of patient rated outcome measures (see Appendix A) (Mokkink et al., 2018; Prinsen et al., 2018). Accordingly, articles which directly or indirectly reported data on the validity of the CAPI Abuse scale, were considered for inclusion in this review. The psychometric properties assessed in this review (i.e., construct and criterion validity and responsiveness) are detailed in Table 1. Due to the size, scope, and complexity of reporting, and since another study concerning other measurement aspects of the CAPI will be conducted elsewhere (P. Martin, personal communication, October 10, 2022), the remaining psychometric properties were beyond the scope of the present review.

To the researcher's knowledge a review as the one proposed here has not been conducted before. The already mentioned systematic review by Georgieva and colleagues (2022) exclusively assessed validity studies between 2010 and 2020, that is, only studies which directly assessed psychometric properties of the CAPI, and without focussing on a specific caregiver population as proposed here. Further, a review by Laulik et al. (2015) reporting on psychometric findings was not performed systematically and did not review CAPI responsiveness (i.e., treatment sensitivity).

Method

Eligibility criteria

Articles which directly analysed or indirectly reported on the psychometric properties of the CAPI Abuse scale were considered for inclusion. The specific measurement properties examined were construct and criterion validity, and responsiveness, with their definitions and relative subdomains summarized in Table 1. Accordingly, only studies which met the following criteria were included:

1. Original articles published in English in peer review journals.
2. Articles published from 1986 (i.e., the publication year of the final edition of the CAPI; Milner, 1986) to January 2023.
3. Studies involving parents or caregivers (including non-biological parents) of children aged 0-12 years old who were determined to be at risk by child protective services (CPS) and constituted at least a significant proportion of the study sample (i.e., 20%).
4. Primary empirical outcome studies using quantitative and mixed study design methodologies (where quantitative data were extractable) which considered any variable statistically analysed specifically in relation to the CAPI Abuse scale in CPS involved caregivers. Studies could have the following design: correlational, longitudinal, comparative, and experimental.
5. Intervention studies for parents at high-risk of child maltreatment that included the mean difference in change scores for CAPI Abuse scale scores between different subgroups and ones that included the mean difference in change scores for CAPI Abuse scale before and after an intervention.

The exclusion criteria were:

1. Studies involving participants who were defined as 'high-risk' caregivers solely based on their demographic characteristics (e.g., younger age group or low socio-economic status).
2. Qualitative studies.
3. Studies using the Brief version of the CAPI (BCAPI; Ondersma et al., 2005), 'shortened' versions of the CAPI, or translated versions which had not been psychometrically and/or cross-culturally validated.
4. Studies using only the factor subscales of the CAPI Abuse scale, that is, Distress, Rigidity, Unhappiness, Problems with child and self, Problems with family, and Problems from others, subscales as well as studies only using CAPI validity scales.
5. Grey literature articles (e.g., doctoral theses, conference papers, press articles), systematic reviews or meta-analyses of articles and last, study protocols.

Though a review on the CAPI has already been published from 1986 to 2010 (Walker & Davies, 2010), the current review included articles published starting from CAPI's inception since no review has collated all studies using the CAPI Abuse scale in maltreating samples only. Additionally, psychometric properties may vary across generations (Georgieva et al., 2021), so a wider time frame may allow exploration of CAPI behaviour across time. Children aged 0-12 years old were the focus of this review as children this age usually spend a significant amount of time and frequently interact with their caregivers (Larson & Richards, 1991). Last, it was deemed that focusing on the original version of the CAPI would ensure appropriate comparisons between studies.

Table 1*Psychometric properties and definitions*

Psychometric property	Domain and definition
Validity	The degree to which an instrument measures the construct(s) it purports to measure
Convergent validity	How closely a test is related to other tests that measure the same (or similar) constructs.
Criterion validity	The ability of the test to accurately differentiate between abusive and non-abusive parents.
Known-group validity	The ability of a test to distinguish among distinct groups where differences are expected a priori.
Responsiveness (or sensitivity to treatment change)	The ability of an instrument to detect change over time in the construct to be measured
Before and after comparison	The ability of the test to detect changes before and after an intervention.
Between groups comparison	The ability of the test to detect changes between different treatment conditions.

Search strategy

A systematic literature search was conducted in PsycINFO, Web of Science, and PubMed electronic databases. Four groups of keywords were identified: the 'Child Abuse Potential Inventory (CAPI)', 'quantitative studies', 'children', and 'high-risk'. The first group specified the examined instrument: CAPI. Though COSMIN guidelines suggest for the literature search to be built more broadly around the constructs of interest, i.e., CPA rather than specific instruments, the purpose of this review was not to identify all existing measures for CPA, which furthermore, has already been done elsewhere (Saini et al., 2019), but to specifically focus on the CAPI. Accordingly, keywords were constructed directly on the CAPI. The second group of keywords focussed on the study design, i.e., quantitative, as it was deemed that this would help search for studies that both directly and indirectly assessed for

CAPI validity; studies that indirectly reported on CAPI validity were, for instance epidemiological or outcome studies that included the CAPI Abuse scale score as an outcome variable rather than directly aiming to study its psychometric properties. The third group had its focus on the collective of interest, that is children aged 0 to 12. The fourth group had its focus on the parent population, specifying parents or caregivers at high-risk of maltreating their child/ren as indicated by child protection service (CPS) or child welfare involvement. These key words are shown in Table 2.

The search was performed in April 2022, and subsequently in August 2022 after the term ‘CAP Inventory’ was introduced to improve the search. A final search was run in January 2023 to identify any additional papers. Filters were applied for language (English), date (1986) and type of article (Journal article).

Table 2

Terms used for Boolean search

CAPI		Quantitative		Children		High-risk
“Child Abuse Potential Inventory” OR CAPI OR “CAP Inventory”	AND	Quantitative OR experimental OR quasi-experimental OR longitudinal OR cross-sectional OR prospective OR retrospective OR “cohort study” OR “case-control” OR “intervention design” OR correlation OR relationship	AND	Child* OR Childhood OR "Middle Childhood" OR "Young Child" OR Infant* OR Infancy OR New-born*	AND	“high-risk” OR “high risk” OR “at risk” OR “at-risk” OR safeguard* OR “child protect*” OR “child welfare*” OR “child in need” OR “child-in-need” OR

OR association	“social
OR factor OR	care**” OR
influence OR	“support
interact OR	service**”
"correlational	OR
study" OR	“welfare”
regression OR	
predict OR “mixed	
methods” OR	
“mixed-methods”	

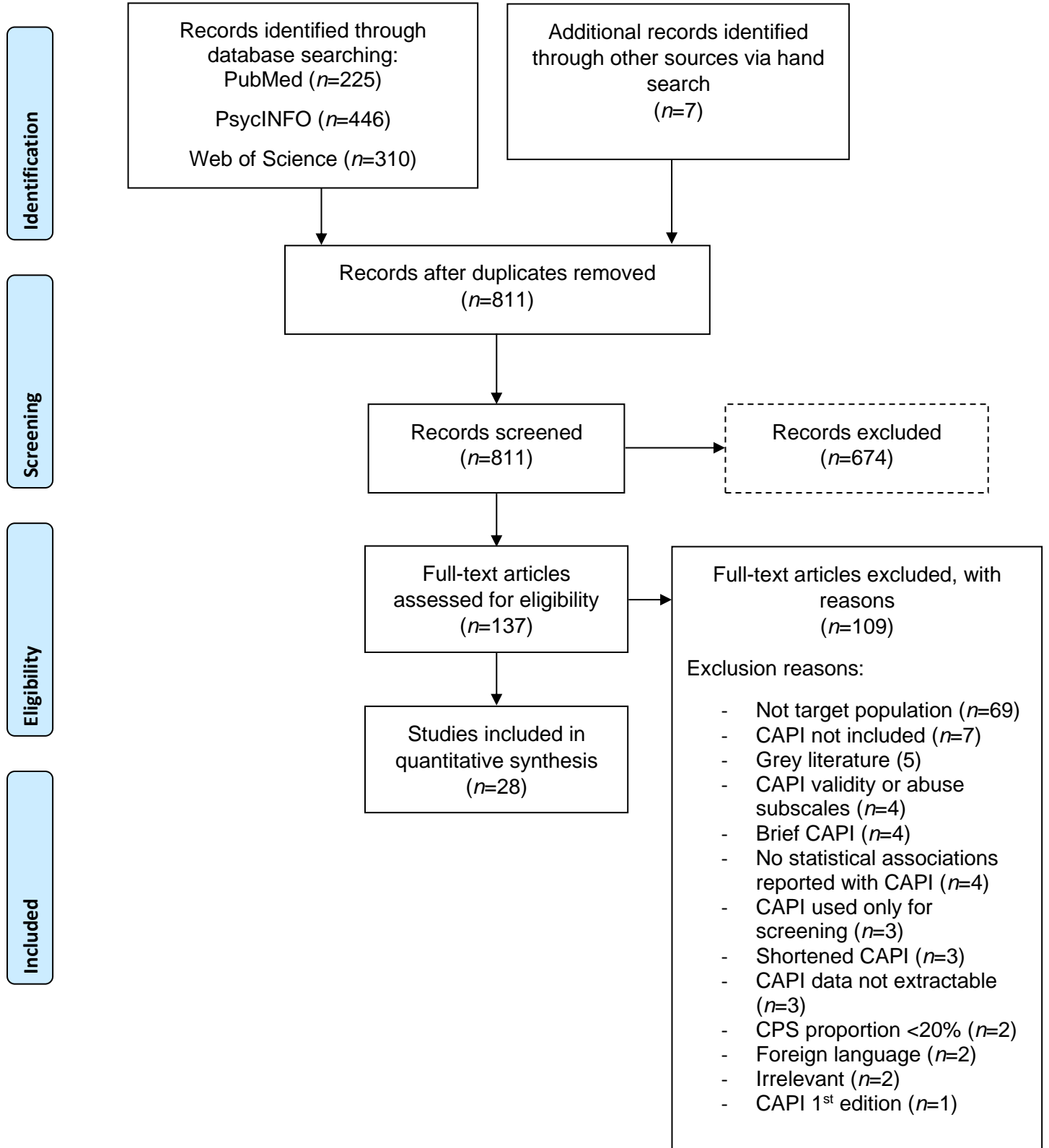
Data collection

To facilitate impartiality in the research process, a protocol was developed and published on Prospero systematic review database (2022- CRD42022383665) prior to commencing searches. The protocol contained suggested search terms and guidelines for identifying and screening articles. This was necessary given the involvement of additional reviewers in the screening process.

Duplicates across the three databases were merged by the primary researcher. The eligibility screening of all titles and abstracts identified with the search strategy was independently conducted by a primary researcher and a secondary reviewer (a clinical psychology trainee) in August 2022. Articles that were considered to have met eligibility criteria by either reviewer were retrieved in full text. Full text versions of the articles that passed the initial screening were assessed independently by both reviewers in September 2022 and reasons for excluding articles were recorded (Figure 1). In case of disagreement, it was discussed with a third reviewer. There were disagreements in two studies, which were solved by consensus. Cohen’s Kappa indicated a good inter-rater agreement ($\kappa = .83$) (Sim & Wright, 2005). A final literature search was conducted in January 2023 where an additional 16 papers were obtained, none of which met inclusion criteria.

Figure 1

Adapted Prisma flow diagram of the study selection (Moher et al., 2009)



Study selection and data extraction

This systematic review followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA; Moher et al., 2015). The results of the searches were transferred into an electronic reference manager (Mendeley Desktop, version 1.19.8) which eliminated duplicate entries. A total of 811 studies were identified, from which 28 were selected for this review (see Figure 1). The information extracted from the eligible studies included the following: names of study authors, publication date, sample characteristics, country, study design, methodological quality of the studies, and psychometric quality of the CAPI Abuse scale.

Quality assessment

The methodological quality of included studies was assessed independently by the primary author and second reviewer using COSMIN checklist for the systematic reviews of Patient-Reported Outcome Measures (PROMs) (Mokkink, Prinsen et al., 2018; see Appendix A). These guidelines are specifically for PROMS, which are defined as measures that directly assess an aspect of the patient's health status by the patient themselves, without any interpretation (Prinsen et al., 2018), and thus apply to the CAPI. The first part of the COSMIN manual focuses on standard procedures in conducting systematic reviews which are in accordance with Cochrane and PRISMA guidelines (Higgins et al., 2019; Moher et al., 2015). The second part guides the evaluation of the measurement properties of the PROM, and the final part refers to the evaluation of the interpretability and feasibility of the PROM, as well as formulating and reporting the systematic review. As per COSMIN manual guidelines, the current review therefore followed three steps to assess the following: 1) methodological quality of studies, 2) psychometric properties of each single study and 3) overall quality of the evidence relating to the psychometric property (see Appendix A). Each step will be described in turn, below.

Step 1 – Assessment of methodological quality of the studies

Methodological quality was assessed in two main groups using the standards included in the COSMIN checklist that were relevant to the measurement properties of interest: Hypothesis testing for construct validity (including convergent and known-group validity), criterion validity and responsiveness (including comparison before and after an intervention and comparison between intervention groups) (see Appendix B). All items on the checklist were scored on a four-point scale ranging from ‘very good’ to ‘inadequate’. The general rating for each study was determined by the lowest score given to an item in the section pertaining to the psychometric property that the study was reporting on (see Appendix B) (Mokkink et al., 2018; Prinsen et al., 2018; Terwee et al., 2018).

Adaptations to checklist standards, which concerned the statistical methods of the included studies, were implemented in this review; study quality was evaluated as being ‘very good’ regardless of whether the article reported the distribution of the sample, since parent/caregiver samples in the included studies were unlikely to have a normal distribution relative to normative community samples due to their ‘high-risk’ nature. As all the included studies comprised this caregiver population, it was deemed that between study comparisons would not be impacted.

Adaptations were also implemented for another item on the checklist concerning criterion and known-group validity; since this review included studies which indirectly analysed psychometric properties of the CAPI Abuse, i.e., studies that used the CAPI as an outcome measure to test research hypotheses rather than analyse its psychometric properties. Therefore, studies which identified subgroups or compared change between subgroups were deemed ‘very good’ irrespective of whether effect sizes (ES) were reported, so long as sufficient data were included in the article (i.e., the standard deviation and sample size) for the primary reviewer to calculate ES. Studies reporting on responsiveness, at a minimum had to have a relatively well-matched control group to receive ‘very good’ methodological ratings.

Last, when multiple comparator instruments were used and multiple analyses were performed in a study, each analysis was considered as a separate ‘study’ as per COSMIN guidelines (Mokkink et al., 2018). This is to address the fact that some of the comparator instruments were of adequate quality while others were more doubtful. Inter-rater agreement was determined using Cohen’s weighted κ (Cohen & Humphreys, 1968) which indicated a good inter-rater agreement ($\kappa=.825$).

Step 2 – Assessment of quality of the CAPI Abuse scale

The psychometric quality of the CAPI Abuse scale was assessed with the updated criteria for good measurement properties according to the COSMIN manual (Mokkink et al., 2018). Specifically, CAPI hypothesis testing of construct validity (convergent, and known-group validity), criterion validity, and responsiveness (before and after; between intervention subgroups) were evaluated, with criteria shown in Appendix C. Given that the working definition of criterion validity used in this review was similar to that of known-group validity (see Table 1), the same COSMIN criteria were applied to both. The quality of these measurement properties was evaluated on a three-point scale: sufficient (+), insufficient (-), and indeterminate (?).

To interpret the results of studies on hypotheses testing for construct validity, the reviewers formulated a set of hypotheses *a priori* on the relationships expected between CAP scores and other similar variables, as well as expected differences in CAP scores between subgroups (see Appendix C and D). With regards to convergent validity, CAP scores were hypothesised to have a strong relationship ($r \geq .50$) to variables such as those consisting of ‘known instances of child physical abuse’ or subgroups with substantiated child abuse claims; these findings were also considered to attest the quality of CAPI Abuse scale criterion validity. Any instrument that was taken to measure another aspect of child abuse (e.g., neglect, emotional abuse) or which proposed to measure physical abuse and was

using the CAPI Abuse scale as the 'gold standard', was considered to be measuring a related but dissimilar construct and, therefore, a lower, or medium-size correlation was expected with the CAP score (i.e., $0.1 < r < 0.5$); these findings were used to assess the quality of CAPI Abuse scale convergent validity (see Appendix D).

Similarly, it was considered that any other instrument measuring aspects of the parent-child relationship, parent characteristics or child characteristics would also be measuring, related but dissimilar constructs (Milner et al., 2022), and therefore a lower correlation was expected with CAP scores (i.e., $0.1 < r < 0.5$) (see Appendix D); these findings were also taken as evidence of CAPI's convergent validity. Given the heterogeneity of the included studies, it was deemed appropriate to assess consistency with reviewer hypotheses based on ES descriptors (rather than statistical significance). This was to ensure that small-sampled studies, which were unlikely to gain statistical significance (Columb & Atkinson, 2016) could still contribute to the quality assessment. However, in the context of multivariate analyses, the statistical significance of the relationships was used to help assess consistency with the included study authors' hypotheses, given that multiple variables within a model were being controlled for.

To evaluate the quality of criterion validity of the CAP scores, hypotheses concerning expected differences between abusive and normative parent subgroups were formulated; it was expected that maltreating caregiver groups would have statistically significant higher CAP scores relative to normative caregiver groups ($p < .05$) (see Appendix D). Similarly, to evaluate known-group validity it was expected that distinct groups, such as trauma exposed, and non-trauma exposed caregivers would also differ significantly ($p < .05$) on CAP scores (see Appendix D), in line with empirical review evidence suggesting that parental history of maltreatment is one of the strongest predictors of child maltreatment (van IJzendoorn et al., 2020). Last, the results of all studies on convergent, criterion and known-group validity of the CAPI Abuse scale scores were then taken together, and it was calculated whether 75% of the results on each property were in accordance with the hypotheses.

With regards to responsiveness, a sufficient (+) rating was given to single study results if data on CAPI Abuse change scores before and after an intervention or change scores between intervention subgroups were available to enable calculation of the standardised mean difference (SMD); the SMD also needed to confirm at least a medium effect size (i.e., Hedges' $g \geq .50$; Cohen, 1988). SMDs that were below a medium effect size were assigned insufficient ratings (-) (i.e., Hedges' $g < .50$; Cohen, 1988). Single study results that did not provide sufficient data to enable SMD calculations were given an indeterminate (?) rating (see Appendix C). Hedges' g estimates were chosen for this review rather than Cohens d , as the latter tend to overestimate SMD in small sample sizes (Cohen, 1988). All results on the responsiveness of CAPI Abuse scale scores from included studies were quantitatively pooled into an overall rating of responsiveness per measure (Prinsen et al., 2018). As per individual studies, an overall sufficient rating on responsiveness for the CAPI Abuse scale was given if the pooled SMD had at least a medium ES (i.e., Hedges' $g \geq .50$; Cohen, 1988). Hedges' g for both single study results and pooled study results was calculated using SPSS version 27. As suggested by Borenstein and colleagues (2010), a random effect model was used for both the within and between group meta-analysis as moderate heterogeneity across studies was detected (i.e., Higgins' $I^2 \geq 50\%$; Higgins et al., 2003).

Additionally, a narrative synthesis approach was used to describe data relating to CAPI Abuse responsiveness and concerning statistical associations between CAP and other constructs. In terms of responsiveness, specifically, this allowed consideration of additional aspects of the data, such as whether CAPI Abuse scores consistently picked up on treatment change in line with study authors' predictions and if so, for which types of intervention. Further, it was considered whether for studies reporting no change in CAP scores following intervention, whether a lack of change was also observed in the other study outcome measures.

Step 3 – Strength of evidence of the CAPI Abuse scale

The quality of evidence, that is, the entire body of evidence used for overall ratings on the validity and responsiveness subtypes of the CAPI Abuse scale, was graded using a modified Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach, which ranged from ‘high’ to ‘very low’ evidence (Mokkink, Prinsen et al., 2018). The modified GRADE approach assumes that the initial quality of evidence used for overall ratings is of high quality. The quality of evidence is subsequently downgraded by one-to-three levels (to moderate, low, or very low) when there are serious (-1: one level down), very serious (-2: two levels down), or extremely serious (-3: three levels down) concerns across the evidence.

Accordingly, the strength of evidence for the CAPI Abuse scale was graded considering four factors: (1) risk of bias (i.e., methodological quality of studies); (2) inconsistency (i.e., unexplained inconsistency of results across studies); (3) imprecision (i.e., total sample size of the available studies); and (4) indirectness (i.e., evidence from different populations other than the population of interest in the review) (Mokkink, et al., 2018; Prinsen et al., 2018).

Results

The flow diagram of included studies is presented in Figure 1. A total of 981 articles were identified through the search of three databases and via reference checking, which after de-duplication left 811 articles to be screened based on title and abstract. After the exclusion of 674 which did not meet criteria, a total of 137 full-text articles were assessed for eligibility, of which 24 all met inclusion criteria. Reference checking of the included articles identified seven additional articles, of which four met inclusion criteria. As a result, 28 articles were included in this review.

Characteristics of included studies

Study samples sizes ranged from 13 to 2,175 participants. Samples are described in detail in Tables 3 and 4. The countries where the studies were conducted were not very diverse: 23 were performed in North America, four in Australia and one in Finland. Despite this, across most included studies, samples were ethnically diverse. In terms of study design, sixteen studies were cross-sectional, five were cohort, and seven were controlled trials (of which five had a randomised design). All studies used questionnaires to collect data; among these, several used observational measures as well as semi-structured, and structured clinical interviews and child abuse risk analogue tasks.

Focusing on the samples of the included studies, over half ($n=16$) included mothers only, whilst the remaining studies included both male and female caregivers, among which two examined potential gender differences in mothers and fathers. The studies' primarily female gender distribution may limit the generalizability of the findings mainly to female caregiver populations. Three of the studies were performed with adolescent samples and three included parents living in foster care. Most participants were youth and adult, with mean ages ranging from 17 years old (Budd et al., 2000) to 40 years old (Haapsalo & Altonen, 1999), where most participants, were on average 29 years old. Children's mean age ranged from eight months to 14 years old. In terms of the parents' profile, 23 studies had

a maltreating sample only, and five had a mixed sample (i.e., maltreating, and matched controls or maltreating and non-maltreating community parents).

Of the final 28 articles, 19 directly analysed as well as indirectly reported on CAPI Abuse scale convergent, criterion and known-group validity (see Table 3), whilst the remaining nine studies only reported data relating to CAPI Abuse scale responsiveness (see Table 4).

Table 3*Description of included articles reporting on the validity of the CAPI Abuse scale*

Authors (year); country	<i>N</i>	Sample characteristics	Mean age (years)	Study design
Bradshaw et al. (2011); USA	82	Mothers (45% White; 26% Black) referred for treatment by caseworkers for child abuse, neglect, and comorbid substance abuse disorder.	Mother: 29 Child: 3	Cross-sectional
Budd et al. (2000); USA	75	Adolescent mothers (89% Black) who were in foster care served by the Illinois CPS.	Mother: 17 Child: 8.2 (months)	Cross-sectional
Budd et al. (2006); USA	49	Adolescent mothers (86% Black) from the study above by Budd et al. (2000) most of whom were still in foster care.	Mother: 17 Child: 8.2 (months)	Retrospective longitudinal*
Caliso & Milner (1992); USA	90	Mothers (65% White; 36% Black) including normative non-abusive with ($n=30$) and without childhood history of physical abuse ($n=30$) and mothers identified as physical abuse perpetrators ($n=30$) by child welfare agencies.	NR	Cross-sectional
Caliso & Milner (1994); USA	78	Mothers (63% White; 37% Black) including physical child abusers who reported a childhood history of abuse ($n=26$), non-abusive mothers who reported a childhood history of abuse ($n=26$), non-abusive mothers with no childhood history of abuse ($n=26$).	NR	Cross-sectional
Craig & Sprang (2007); USA	1,680	Maltreating caregivers (50.1% female; 87.2% White) with open substantiated cases of abuse and neglect, consecutively assessed at a university outpatient-based clinic for comprehensive treatment for maltreating families.	Parent: 33.7 Child: NR	Cross-sectional
Donohue et al. (2016); USA	77	Mothers (50.6% White; 20.8% Black) referred for treatment of substance abuse and child neglect by the county's Department of Family Services.	Mother: 29 Child: 3.8	Cross-sectional

Donohue et al. (2017); USA	18	Mothers (39% Black; 33% White) referred by county CPS agency caseworkers for family behavioural treatment to assist substance use disorder and child neglect.	Mother: 29.28 Child: 3.22	Cohort study
Donohue et al. (2019); USA	80	Mothers (48.8% White; 23% Black) referred for behavioural treatment for drug abuse and child neglect by CPS.	Mother: 28.88 Child: 3.83	Cross-sectional
Haapsalo & Altonen (1999); Finland	50	Mothers (ethnicity not reported) with previous CPS involvement due to child abuse or neglect ($n=25$) and matched sample of Finnish mothers from the community with no current or prior CPS involvement ($n=25$).	Mother: CPS: 40, control: 39 Child: CPS: 12, control: 11	Cross-sectional
Haskett et al. (1995); USA	41	Parents (34 mothers, 7 fathers; 60% Black; 40% White) enrolled in a university-based multifamily group intervention for maltreating high-risk parents and their young children.	Parent: 30.1 Child: 6	Cross-sectional
Hien et al. (2010); USA	152	At-risk community sample of mothers (71% Black) belonging to minority population recruited via OB/GYN clinic in public hospital; recent past (34%) or current (9%) open cases with CPS.	Mother: 37.83 Child: NR	Cross-sectional
Holden et al. (1989); USA	87	Parents (80% White) consecutively referred for a wide range of parenting problems, recruited from community agency specialising in intervention with maltreating parents - mothers (70%); over 75% were court referred –remaining caregivers referred for reasons other than confirmed child abuse/neglect.	NR	Cross-sectional
Kilpatrick (2005); Australia	103	Parents majority of White Celtic origin, 88% Mothers, ($n=50$ registered maltreating, $n=32$ matched distressed, and $n=21$ matched controls).	Parent: NR Child: 8.6	Cross-sectional
Plant et al. (2016); USA	72	Mothers (47% White; 25% Black) referred for behavioural treatment for child neglect and substance abuse by caseworks from local CPS agency.	Mother: 29.04 Child: 3.91	Cross-sectional
Rinehart et al. (2005); USA	536	Mothers (56% White; 27% Hispanic; 20% Black) participating in the Substance Abuse and Mental Health Services Administration's (SAMHSA's) national Women Co-occurring Disorders and Violence Study (WCDVS); ($n=88$ were mandated to participate by CPS).	Mother: site 1: 31, site 2: 33 Child: NR	Cross-sectional

Sprang et al. (2005); USA	208	Maltreating parents (45% male, 55% female; 60.6% White, 30% Black) with open substantiated cases of abuse and neglect referred to the Comprehensive Assessment and Training Service (CATS) project by state public child welfare organization.	Parent: 28 Child: 6.46	Cross-sectional
Rodriguez & Silvia (2022); USA	114	Two samples (AIMS-P: $n = 38$; TRIPLE-F: $n = 76$) of mothers (73.7% and 75% Black, respectively). AIMPS-P comprised CPS-involved mothers, mandated to take part in designated parent training program; TRIPLE-F comprised matched control sample of mothers drawn from four-wave longitudinal study tracing factors contributing to child abuse risk across time.	Mother: sample 1: 29.54, sample 2: 27.66 Child: sample 1: 2.54, sample 2: 6.33	Cross-sectional
Urgelles et al. (2012); USA	26	Mothers (38.5% White; 26.9% Black) referred by CPS to a 6-month treatment outcome study after having evidenced an incident of child neglect and illicit drug use within the past 4 months; participants had to have had at least 2 episodes of Emergency Prevention Management (EPM) and they received at least one Family Behaviour Therapy session.	Mother: 30.12 Child: NR	Cohort

Note. CPS: child protective services; NR: not reported; * Study used longitudinal analysis between baseline CAPI Abuse scores and PSI-SF at T2 and a cross-sectional analysis between CAPI Abuse scores and other variables at T1

Table 4*Description of included articles reporting on the responsiveness of the CAPI Abuse scale*

Article	N	Study purpose	Intervention name	Study design	Study population	Mean age
Chaffin et al. (2012); USA	2,175	Test SafeCare Homebased services with parents in CPS vs Homebased services as usual in a scaled-up implementation.	SafeCare Home-based services	RCT	Parents/caregivers (91% female; 67% White; 16% Native American; 9% Black) enrolled in a state-wide system operated by community-based agencies under contract with CPS.	Parent: 29.4 Child: NR (Preschool age)
Donohue et al. (2014); USA	72	Evaluate the effectiveness of family-based behavioural therapy in substance abusing mothers referred from CPS for child abuse and neglect.	Family Behaviour Therapy (adapted to meet unique needs of families referred by CPS for substance abuse)	Controlled	Mothers (47.2% White, 25% Black) referred for treatment of substance abuse and child neglect by the county's Department of Family Services living with the target child.	Parent: 29.04 Child: 3.92
Harnett & Dawe (2008); Australia	13	Evaluate the effectiveness of home-based intensive treatment program in families referred by CPS.	Parents Under Pressure (PUP)	Cohort	Parents (65% female) of 10 families (two identified as Indigenous Australians) were referred for treatment by a nongovernment agency (Parents Under Pressure Program; PUP) that provides	Parent: 32 Child: 4.4

					services for parents and children of families involved in CPS.	
Hubel et al. (2018); USA	294	Evaluate the effectiveness of SafeCare among diverse group of at-risk parents.	SafeCare	RCT	Subsample consisting of adolescent parents (98 % female; 65.2% White; 12.9% Native American) with pre-school aged children drawn from a state-wide controlled trial of the 'SafeCare' home visiting model in Oklahoma; referred to community-based agencies under contract with the child welfare system.	Parent: 19.6 Child: NR (Preschool age)
Kolko et al. (1996); USA	55	Evaluate the effectiveness of child and parent CBT or Family Therapy relative to families who receive routine community service.	Child and Parent CBT, Family Therapy	RCT	Maltreated children (47% Black, 22% White) and their parents and guardians who received either a project treatment protocol or routine community services; CPS case workers referred 87% cases following incidents classified as physical abuse, maltreatment, or severe frequent use of physical discipline.	Parent: NR Child: 8.6
Sanders et al. (2004); Australia	98	Evaluate whether Parental attributional retraining and anger management enhance the effects of the Triple-P-Positive Parenting programme.	Enhanced group behavioural family intervention programme vs standard care group parent	RCT	Mixed sample of parents (ethnicity not reported) (92/94% female) including families at risk of child maltreatment referred via CPS, family doctors etc. to an enhanced group behavioural family intervention that specifically targeted parents' negative attributions regarding their	Parent: 34 Child: 4.4

			training programme		child's and their own behaviour and parents' anger control deficits.	
Thomas & Zimmer-Gembeck (2011); Australia	150	Evaluate the effectiveness of Parent Child Interaction Therapy (PCIT) among mothers at high risk of maltreating child/ren.	PCIT	RCT	Female caregivers (ethnicity not reported) referred from child protection authorities, identified as suspects of maltreatment by other professionals, or self-identified because of significant child behaviour problems and stress.	Parent: 33.5 Child: 5
Timmer et al. (2005); USA	307	Examine the effectiveness of Parent Child Interaction Therapy (PCIT) with maltreating parent-child dyads.	PCIT	Controlled	Parent-child dyads (89% female; 53% White, 21% Hispanic/Latino) with and without child maltreatment history. Children of dyads with maltreatment history were referred by CPS social worker, were 2-8 years old and had high externalizing problems (parent-child dyads: $n=193$ with child maltreatment history; $n=114$ without such history).	NR
Vorhies et al. (2009); USA	25	Evaluate the effectiveness of a residential programme with wrap around services for pregnant and parenting foster care youth.	Transitional Living Programme (TLP)	Cohort	Mothers (88% Black) who are wards of the state who took part in Transitional Living Programme for foster care youth with severe mental illness who are pregnant and parenting.	Parent: 18.8 Child: 1.35 (T1); 1.53 (T2)

Step 1 - Methodological quality

As already mentioned, when multiple comparator instruments were used in relation to the CAP score and therefore multiple corresponding analyses were performed within a single study, each analysis was considered a separate 'study' as per COSMIN guidelines (Mokkink et al., 2018). This resulted in a combined total of 160 'studies' extracted from the 19 included articles reporting data on construct and criterion validity and from the nine included articles reporting data on responsiveness. The methodological quality of each 'study' was assessed using the COSMIN risk of bias checklist (Prinsen et al., 2018) with adaptations made to certain standards as per detailed in the methods section. Tables E1-5 in Appendix E present an overview of all methodological quality ratings per 'study' extracted from the 19 articles reporting on construct and criterion validity and the nine articles reporting on responsiveness, respectively.

Among the 160 analyses obtained across all 28 articles, the majority 72.5% (116/160) reported data on convergent validity, whilst a smaller proportion also considered criterion 7.5% (12/160), and known-group validity 5.6% (9/160), and responsiveness 14.4% (23/160). All nine articles concerning CAPI responsiveness conducted 'before and after intervention' hypothesis testing, and the majority (7/9) also conducted 'comparisons between subgroups' hypothesis testing insofar as they were controlled studies. In terms of convergent validity, only one study correlated the CAP score to what could be considered a 'gold standard' of child physical abuse, namely child maltreatment severity as defined by a state CPS coding system in the USA (Sprang et al., 2005). With regards to studies reporting on what was defined here as criterion validity, comparisons on CAPI Abuse scores between maltreating caregivers and matched controls were conducted.

Concerning the methodological quality ratings, in total, 92% (107/116) of studies reporting on convergent validity were scored as having 'very good' or 'adequate' methodological quality, whilst the remaining studies were considered 'inadequate'. The latter ratings were mostly due to the use of instruments which had not been adequately validated

in the study target population or due to conducting statistical analyses that typically require larger samples sizes. Similarly, all the analyses concerning criterion and known-group validity were scored as having very good or adequate methodological quality. For responsiveness, over half the articles reporting data on this measurement property 55% (5/9) were scored as having doubtful methodology. This was primarily due to a lack of reporting on whether the different conditions were adequately matched on demographic characteristics or study outcome variables. Furthermore, two reported on small non-controlled studies, which conducted parametric analysis on insufficient sample sizes.

Step 2 - Psychometric quality of the CAPI Abuse scale

The quality of convergent, criterion, and known-group validity and responsiveness of the CAPI Abuse scale were examined according to the updated COSMIN criteria for good measurement properties as shown in Appendix D (Mokkink et al., 2018; Prinsen et al., 2018; Terwee et al., 2018). A psychometric rating was applied to each psychometric property by study (see Tables E1-5 in Appendix E).

To examine convergent validity of the CAPI Abuse scale among high-risk parents, additional analyses were conducted (see Appendix D). First, the measures that were used to test convergent validity against the CAPI Abuse scale were examined. Second, the reported correlation coefficient or regression outcome was compared to the reviewers' hypotheses concerning what the expected correlation would be. Similar steps were performed to report on the criterion and known-group validity of the CAPI abuse scale among high-risk parents (see Appendix D); the reported *p*-value obtained when comparing groups on CAPI Abuse scale scores were compared to the reviewers' hypotheses concerning expected differences.

For convergent validity, 102 of the 116 studies (88%) performed across included articles received sufficient ratings. For criterion validity, 11 of the 12 studies (92%) performed across included articles received sufficient ratings; one study was rated indeterminate owing

to a lack of reviewer a priori hypotheses concerning expected differences in CAPI Abuse scores between specific trauma subtypes among maltreating parents. For known-group validity, five of the nine studies (55.5%) performed across included articles received sufficient ratings, whilst three received indeterminate ratings due to reviewers, again, not having formulated an a priori hypothesis concerning expected CAPI Abuse score differences for gender and specific subtypes of trauma. Importantly, only one study reported on criterion validity by comparing CAPI Abuse scale scores with maltreatment severity as coded using CPS state records reporting on substantiated cases of abuse (Sprang et al., 2005); this measure of maltreatment severity can be considered to approximate what the reviewers had formulated as a 'gold standard' a priori. Interestingly, some studies used CAPI Abuse as the 'gold standard' measure to establish criterion validity for the instrument they were validating (i.e., Bradshaw et al., 2011, Kilpatrick, 2005, and Rodriguez & Silvia, 2022).

In terms of responsiveness, two studies with inadequate methodological quality had smaller effect sizes. Although results from methodologically doubtful analyses may be biased, these results were still included when pooling results from all analyses, given that pooled results downgrade the quality of evidence in terms of risk of bias (Mokkink et al., 2018). Of all ten ratings on *before and after intervention* responsiveness, only one received an indeterminate rating due to less robust data being reported, which therefore prevented effect size calculations (see Table E4, Appendix E). All other analyses received either a sufficient (4/10) or insufficient (5/10) rating. Of all 13 ratings on *between intervention and control group* responsiveness only two received an indeterminate rating due to the data reported being insufficient for the reviewer to calculate effect sizes. All other analyses received either a sufficient (6/13) or insufficient (5/13) rating (see Table E5, Appendix E).

Statistical associations with CAPI Abuse scale

The studies included here, reported a range of statistical associations with CAPI Abuse scale scores among high-risk and maltreating parents. The most frequently studied factors were parent-related, followed by social contextual and child-related factors. Among the parent-related factors the following domains were identified: affective (e.g., depression, anger arousal, and emotional distress), cognitive (e.g., reading achievement), behavioural (e.g., drug and alcohol use) and parenting-related variables (e.g., parental empathy, parenting attitude, infant stimulation, observed parenting behaviour, maltreatment severity). The following contextual-related factors were also studied: ethnicity, socio-economic status, social problems, parents' previous history of childhood abuse. The only child-related factors were parent and teacher perception of child adjustment.

The risk factors that showed a large association with CAP scores included the following: childhood history of violence in the parents' family of origin, social problems (i.e., criminal conviction, alcohol and drug abuse, institutionalization for mental health issue), emotional distress, observed parenting behaviour, parental emotional empathy, social satisfaction, and parent perception of child adjustment behaviour problems. Medium size relationships to CAP scores were found for the following variables: maltreatment severity, socio-economic status, parents' history of childhood physical abuse, academic achievement, content in their child across various domains, frustration intolerance, current adult sexual abuse in parents, and teacher reported child adjustment. The remaining factors (see Table D1, Appendix D) showed small size relationships to CAP scores.

In terms of multivariate analyses, the following factors were found to significantly explain CAP even when controlling for a range of different demographic, affective, cognition and trauma related covariates: academic achievement and average satisfaction in social support (Budd et al., 2000); mother, father, and sibling support, as well as mother and sibling punishment (Caliso & Milner, 1994); parents' childhood sexual and physical abuse as well as adult physical and sexual abuse (Craig & Sprang, 2007); mothers' hard drug and Marijuana

use (Donohue et al., 2017); social problems (i.e., criminal convictions, substance abuse disorder, mental health issues requiring medication and institutionalization) (Haapsalo & Altonen, 1999); parents' anger arousal and reactivity (used as a measure of emotion regulation difficulties; Hien et al., 2010); and last, emotional distress (Rinehart et al., 2005). Together these multivariate results (see Appendix D) tentatively suggest that these parent-related factors may have an important role in predicting CAP. Further, given that in many cases they confirmed study author hypotheses, they also received positive ratings, thus contributing to the evidence base on CAPI validity (see Table E1, Appendix E).

Criterion and known-group validity

Across all studies comparing maltreating caregivers to normative non-abusive parents, the former scored significantly higher on CAP scores, with effect sizes ranging from small to medium (see Table E2, Appendix E). Furthermore, the CAPI scores were able to distinguish between known groups of parents with and without childhood histories of abuse, with results showing that, in line with theory, the former consistently scored higher on the CAPI Abuse scale relative to parents without childhood histories of victimization. Two studies comparing CAP scores based on parent gender revealed small effect size differences between female and male caregivers, with the former scoring significantly higher. Other comparisons on CAP scores revealed that substance abusing, and clinically depressed parents scored significantly higher on CAPI Abuse than normative parents. Collectively, the fact that CAPI scores could significantly distinguish between maltreating and normative parent groups and other known-groups of caregivers, supports its criterion and known-group validity, respectively.

Responsiveness

The CAPI Abuse scale demonstrated responsiveness to before and after treatment change fairly consistently, as indicated by the decrease in CAP scores following multiple types of parent interventions, which included Family behaviour therapy, the Parents under Pressure (PUP) programme, Parent Child Interaction Therapy (PCIT), the standard and enhanced versions of family behavioural interventions adapted from the Triple-P parenting programme, child and parent CBT interventions, and family therapy.

Three studies did not strongly support the post-treatment responsiveness of the CAPI Abuse scale. In both the study by Kolko et al. (1998), which evaluated the effectiveness of two treatment types (i.e., parent and child CBT and family therapy) and Hubel et al. (2018), which assessed SafeCare intervention efficacy, the effect of time showed only marginally significant improvements in CAPI scores ($p < .08$, and $p < .09$, respectively). In the study by Vorhies et al. (2009), assessing the effectiveness of a Transitional Living Programme for adolescent mothers in foster care, no significant effect of time was noted in the CAPI abuse scores. Though not statistically significant, Kolko et al. (1998) did note that parents in the CBT and family therapy conditions reported a consistent reduction in CAPI Abuse scores over time and the interpersonal factor subscales of the CAPI Abuse scale did reveal significant reductions. However, the overall lack of statistically significant findings on the CAPI Abuse scale contrasts with time effects observed on other outcome measures used in the study, such as parent-to-child violence. Importantly, however, in both the Hubel et al. (2018) and Vorhies et al. (2009) study, for the most part, the interventions did not result in significant improvements for the other outcomes employed either. Indeed, though Hubel and colleagues (2018) found improvements in depression in both treatment conditions, no significant differences were found before and after treatment in terms of child welfare recidivism, which is consonant with the lack of significant reductions in CAPI scores.

Similarly, Vorhies and colleagues (2009) did not observe significant reductions for any of the other treatment measures (i.e., Brief Symptom Inventory, Parent Opinion

Questionnaire, and the Parental Stress Index) except for educational attainment, which constituted one of the main targets of the programme.

Mixed findings were obtained for between intervention group responsiveness. Whilst expected differences in CAPI Abuse scores were obtained in three of the controlled studies included here (i.e., Donohue et al., 2014; Kolko et al., 1998; Sanders et al., 2004), such differences were not observed for the remaining three controlled studies (i.e., Chaffin et al., 2012; Hubel et al., 2018; Thomas & Zimmer-Gembeck, 2011). Importantly, Hubel et al. (2018) did not observe any significant differences between treatment conditions on child maltreatment recidivism either, whilst Chaffin et al. (2012) and Thomas & Zimmer-Gembeck (2011) observed significant between group differences on child maltreatment recidivism, and child behaviour parental stress and parent child interactions, respectively, which favoured the treatment condition.

Step 3 - Strength of evidence

The strength of evidence for each measurement property (i.e., convergent, criterion and known-group validity, and responsiveness) was assessed using the GRADE analysis. This involved assigning an overall 'quality of evidence' rating as presented in Tables 5 and 6, for convergent validity, criterion and known-group validity, and responsiveness. These tables also include the overall psychometric rating per measurement property. This was obtained by pooling results shown in the Tables E1-5 in Appendix E, as per COSMIN guidelines (Mokkink et al., 2018), to obtain an overall rating for each psychometric measure of the CAPI Abuse scale. According to the modified GRADE approach, it can be considered that the CAPI Abuse scale has moderate evidence for convergent validity, moderate evidence for criterion and known-group validity, and moderate and low strength evidence for before and after and between groups responsiveness, respectively, among at-risk parent populations.

Table 5

Step 3 - Overall ratings on pooled study results and quality of evidence on convergent, criterion and known-group validity

Quality of evidence (Convergent validity)				Overall rating on pooled results	Overall quality of evidence (reasons)
Pooled results					
Risk of bias	Inconsistency	Imprecision	Indirectness		
No concern: Multiple studies of adequate methodological quality	Small concern: 12%	No concern: Pooled sample size ($n = 3,479$)	No concern: All studies addressing target population of this review (high-risk/maltreating caregivers)	+	Moderate: small inconsistency across studies; most studies were outcome rather than psychometric studies
Quality of evidence (Criterion & Known-group validity)				Overall rating on pooled results	Overall quality of evidence (reasons)
Pooled results					
Risk of bias	Inconsistency	Imprecision	Indirectness		
No concern: Multiple studies of adequate methodological quality	Small concern: 24%	No concern: Pooled sample size ($n = 2,173$)	No concern: All studies addressing target population of this review (high-risk/maltreating caregivers)	+	Moderate: small inconsistency across studies; most studies were outcome rather than psychometric studies

Note. The modified GRADE approach was used for grading the quality of summarized evidence on hypothesis testing for convergent, criterion and known-group validity; an overall positive (+) rating was assigned on the basis that at least 75% of study results were in line with reviewer hypotheses.

Table 6

Step 3 - Overall ratings on pooled study results and quality of evidence on responsiveness

Quality of evidence (before and after treatment Responsiveness)					Overall rating on pooled results (overall rating on adjusted pooled results)	Overall quality of evidence (reasons)
Pooled results						
Risk of bias	Inconsistency ^a	Imprecision	Indirectness	Hedges' g effect size (95% CI; I ²)		
No concern: Multiple studies of adequate methodological quality	Serious concern: Moderate heterogeneity in results across studies (I ² = 49%)	No concern: Pooled sample size (n = 1,014)	No concern: All studies addressing target population of this review	0.56 (.32 .80; 49%);	+	Moderate: partly inconsistent results across studies
Quality of evidence (between group Responsiveness)						
Pooled results						
Risk of bias	Inconsistency ^a	Imprecision	Indirectness	Hedges' g Effect size (95% CI; I ²)	Overall rating on pooled results (overall rating)	Overall quality of evidence (reasons)
No concern: Multiple studies of adequate methodological quality	Serious concern: high heterogeneity in results across studies (I ² = 84%)	No concern: pooled sample size (n = 976)	No concern: All studies addressing target population of this review	-0.55 (-.36 -.07; 84%)	-	Low: inconsistent results across studies

Note. The modified GRADE approach was used for grading the quality of summarized evidence on responsiveness; ^a to evaluate inconsistency in results across studies reporting on responsiveness, *I-squared* (I²) statistic was used, which is the percentage of the total variability in a set of effect sizes across studies due to heterogeneity; values of less than 50%, 50% to 74%, and higher than 75% indicate low, moderate, and high heterogeneity, respectively (Higgins et al., 2003).

Discussion

The aim of this systematic review was to evaluate the quality of construct validity (convergent and known-group validity), criterion validity and responsiveness (comparison before and after interventions and between intervention subgroups) of the CAPI Abuse scale among parents at high risk of maltreating their child/ren.

This review identified 28 articles that reported on these measurement properties of the CAPI Abuse scale within the target population. The identified individual articles contained 160 studies, the methodological quality of which was generally adequate or very good. Most of the studies provided convergent validity data, and supported the construct, criterion, and known-group validity of the CAPI Abuse scale in this regard. However, with regards to the strength of the validity data, most studies indirectly assessed the CAPI Abuse scale psychometric properties and, as such, could not ultimately be considered high-quality overall. Helpfully, most of the included articles reported hypotheses *a priori* which aided reviewers in establishing hypotheses on which to rate the findings against. With regards to CAPI responsiveness, a proportionately smaller number of studies reported on this property. The moderate quality evidence for before and after intervention responsiveness suggested that the CAPI Abuse scale may be adequate in detecting post-treatment effects of parenting interventions aimed at reducing parental child abuse. On the contrary, the low-quality evidence for the responsiveness data detecting change between intervention and control groups, suggested that the evidence supporting CAPI Abuse scale's ability to detect treatment effects between conditions is currently insufficient. Overall, these findings indicate that further research directly assessing and further reporting on the CAPI Abuse scale responsiveness properties is required before its use in intervention research among high-risk parents can be fully endorsed. In terms of its construct and criterion validity, the evidence supports its use in maltreatment research and potentially clinical assessment, though future validity studies which further evaluate its criterion validity using gold standard measures and constructs are strongly encouraged.

Convergent validity

As often seen in the child maltreatment literature, most of the included studies assessed the relationship between Child Abuse Potential (CAP) and parent-related factors (Milner et al., 2022; WHO, 2020). Theoretical links between parental characteristics and CAP were generally confirmed showing that, as expected, CAP was related to a wide range of parent-related factors such as parental depression, psychopathology, anger arousal, drug use, parents' own history of abuse, as well as parenting variables, such as parents' satisfaction in the parent-child relationship, parental stress, empathy and so on. Interestingly, only two studies examined the CAPI Abuse score in relation to child characteristics, namely parent and teacher perceived child adjustment; however, this is not representative of the larger number of child characteristics, such as age, gender, attachment style, and temperament, which are known to be related to child physical abuse risk (WHO, 2022). Though studies generally provided adequate justification for why variables would be correlated with CAP scores and what hypotheses they had concerning these associations, no study reported hypotheses concerning the expected magnitude of the association. As such, caution is warranted when extrapolating inferences on CAPI validity from these findings, since they do not relate directly to validity for predicting child abuse.

Collectively, the findings suggest that parents who endorse norms and beliefs typically observed in known child physical abusers, have more psychopathology and social isolation, carry out neglectful parenting actions more frequently, perceive their child/ren as having more internalizing and externalizing behaviour problems, are less able to exhibit empathy, respond empathically or come up with empathic reasons for their child/ren's behaviour; make more negative attributions of their children's behaviour; have childhood histories of abuse of their own, and are more likely to have a concurrent number of social problems such as alcohol and drug abuse, mental health problems, as well as criminal convictions. The correlational findings here also suggest that parents with high abuse potential scores score lower on arithmetic and reading abilities, report more emergencies,

exhibit less infant stimulation within the home, have higher inappropriate expectations of child behaviour, have less satisfaction in their children's education activities and in their overall happiness. Last, they are more likely to occupy lower socio-economic (SES) backgrounds.

The multivariate findings suggest that high-risk parents' academic achievement, social support, emotional distress, anger arousal, hard drug use and experiences of childhood sexual abuse as well current adult experiences of abuse, may be of particular relevance given that they accounted for variance in CAP even after controlling for other well-known salient factors, such as parental stress, demographic characteristics and other trauma types.

Importantly, common risk factors were reported in this systematic review with two other meta-analyses on child physical abuse risk factors (Milner et al., 2022; Stith et al., 2009). These factors include negative parent-child interactions, parent cognition, parenting stress, parental childhood history of maltreatment, psychopathology, depression, parental empathy, problem solving, and negative child behaviour attributions, social isolation, perceived child problems, inappropriate expectation of child behaviours.

Furthermore, the review findings are largely consistent with the evidence base on child maltreatment more generally, which has found that parental characteristics and socio-economic background are some of the strongest predictors of child maltreatment (Mulder et al., 2018; van IJzendoorn et al., 2020). Indeed, in terms of parental-related factors, child maltreatment models frequently refer to parent-related characteristics, such as parental history of childhood maltreatment, parent cognition, parenting stress and parental perception and attributions of child behaviour, which may thus constitute important markers in CPA (Assink et al., 2018). However, caution must be used when considering these findings since most studies adopted cross-sectional or case-controlled designs. This is a limitation due to the large number of risk factors proposed at community and cultural levels that do not necessarily appear in cross-sectional type research. Consequently, these findings may

present with an inflated emphasis on parent-related individual and relational factors, which limits the extent to which they can be used to understand competing models.

Nonetheless, the finding that SES and parental history of child abuse are associated with CAP is consistent with a recent umbrella synthesis which identified these factors as two of the strongest predictors of child maltreatment (van IJzendoorn et al., 2020). Factors underlying the association between SES and CAP may involve having an undesirable home environment, or the fact that parents from lower income and educational backgrounds have less parenting knowledge and social resources and lower ability to cope with anxiety and stress in their parenting roles (Marsh et al., 2020). The association between SES and CAP may be complex, especially among maltreating caregiver samples, and may not be solely explained via correlational findings. Future studies should thus further investigate the role of SES in CAP.

Overall, though numerous theoretical links were confirmed empirically, thus providing indirect evidence in support of CAPI construct validity, only one study assessed the relationship between CAP and what could be considered a gold standard of maltreatment, i.e., maltreatment severity of parents with substantiated cases of abuse known to CPS (Sprang et al., 2005). Surprisingly, contrary to reviewer hypotheses, the study authors obtained a medium rather than large effect size relationship between CAP and maltreatment severity. Given the lack of additional studies using gold standard criterion measures it is not possible to discuss the full importance of these findings. The fact that very good inter-rater reliability scores had been obtained for the CPS coding system used by Sprang and colleagues (2005), arguably, provides evidence against CAPI criterion validity. However, importantly, the authors did report higher mean CAPI scores that were above CAPI clinical cut off for extreme levels of maltreatment and consecutively lower sets of scores for severe, moderate and mild forms of abuse, respectively. The lack of statistical analyses concerning group differences, however, prevented extraction concerning the size and significance of these CAPI differences obtained for each maltreatment severity classification. Furthermore,

no study provided predictive validity evidence, to assess whether CAP would be associated to gold standard constructs in the future. As such, future correlational studies are strongly recommended to fill these gaps to better attest the criterion validity of the CAPI Abuse scale.

Criterion validity

The finding that CPS groups obtained significantly higher CAP scores relative to community sample parents can be considered to contribute to the evidence base on CAPI Abuse scale criterion validity, insofar as these results suggest that the CAPI was successfully able to discriminate between the two groups. These findings are consonant with original criterion validity study findings described by Milner (1986), in which parent samples with active and substantiated cases of child abuse were correctly identified over comparative non-abusing parent groups.

Importantly, the five studies included here reporting on CAPI Abuse scale criterion validity all used matched control groups. This increases the robustness of the findings to the extent that significant differences reported are more likely to result from actual differences in parental attitudes and beliefs known to be associated with higher risk of child physical abuse, as opposed to other sample differences. Notwithstanding, across the included studies, groups were not perfectly matched on all demographic variables. In Rodriguez and Silvia (2022) for instance, parents in the maltreating sample had more children and of older age groups relative to the non-maltreating sample. Importantly, however, the authors controlled for this difference in their analyses. This was not the case for Holden et al. (1989) who did not control for significant differences across groups on mothers' educational level, which as discussed in the previous section, is a factor which has been found to be strongly related to CAPI Abuse scale scores. Similarly, Haapsalo and Altonen (1989) did not control for significant differences in mothers' education levels, socio-economic status, mental health, criminal and substance abuse problems. As a result, in both these studies the large CAP score differences obtained between groups may in part be biased by differences on these

other demographic factors. This was less of an issue in the studies by Caliso and Milner (1992; 1994), which did not find significant differences in age, education, and number of children nor significant differences in ethnic backgrounds. Importantly, Caliso and Milner (1992; 1994) found small to medium effect size differences between groups, thus corroborating the criterion validity of the CAPI Abuse scale.

Known-group validity

Most studies examining the CAPI Abuse scale's known-group validity used t-tests or ANOVAs to assess differences between distinct groups on CAPI Abuse scale scores. The majority reported whether differences were significant rather than effect sizes, however for most cases this could be calculated by the reviewer. Many articles included clear hypotheses concerning differences that they were expecting. Despite this, it would have been beneficial for anticipated differences to be quantified in advance in terms of effect size. This would have provided more robust known-group validity data since another criterion would have been applied to the evidence base. Nevertheless, CAPI Abuse scale scores were tested for known-group validity against parental history of childhood abuse, parental trauma history, gender, substance abuse disorder and depression. Expected significant differences were obtained for all domains.

The finding that trauma exposed groups scored significantly higher on the CAPI Abuse scale overall reinforces the evidence base concerning the intergenerational transmission of abuse hypothesis, which suggests that parents with their own experiences of childhood victimization are more likely to perpetrate child abuse later in life (Assink et al., 2018). Interestingly, the two studies which compared female and male caregivers both obtained significantly higher CAP scores for the former. This result tentatively suggests that female caregivers may have higher child abuse potential and is consistent with a previous meta-analytic finding on gender differences in CPA perpetrators (Behl et al., 2003). Behl and colleagues (2003) had found that females were included significantly more frequently than

males in Child Physical Abuse (CPA) perpetration articles. Their meta-analytic results were congruent with the perpetration prevalence literature at the time, which indicated a larger proportion of CPA perpetrators among mothers compared to fathers (US Department of Health and Human Services, 1999). In line with these older prevalence data, more recent US prevalence statistics indicate that there is a higher proportion of child fatalities perpetrated by mothers relative to fathers and that 1, 452, 099 children were abused and neglected by a mother acting alone compared to 661,129 children who were abused and neglected by a father acting alone (US Department of Health and Human Services, 2006). These statistics are in line with even more recent data, which indicate a higher percentage of CPA perpetrated by mothers (51.7) versus fathers (47.2) (US Department of Health and Human Services, 2021). Statistics that were based on US state submissions to the National Child Abuse and Neglect Data System (NCANDS) of alleged child abuse and neglect also indicated a higher prevalence of abuse perpetrated by mothers (210, 746 cases) relative to fathers (132, 363) (Statista, 2021). Interestingly, findings from the 2015-2016 Crime Survey for England and Wales found that 29% of victims were abused by their mothers, whilst 39% were abused by their fathers (UK Census Families and households in the UK, 2021). Nonetheless, most of the available prevalence data summarised here indicate a higher proportion of CPA and child fatalities among mothers relative to fathers.

There are important caveats, however, to such gender prevalence differences in abuse. Indeed it has been noted that such differences may be an artefact produced by single parent studies (Behl et al., 2003); a study which had accounted for the absence of fathers in the home pointed out that fathers are equally likely, if not even more likely, to be the perpetrators of CPA (Nobes & Smith, 2000). As commented by Behl and colleagues (2003), a deliberate oversampling of males may thus be helpful in reducing participation bias when examining gender-based differences in child abuse. Furthermore, from the US prevalence data summarised above, it is not possible to discern whether a larger proportion of mothers perpetrate CPA, and more children are killed by mothers, simply because more

women than men are primary carers of children. In the UK, for example, in 2014, 91% of single parents were women (UK Census, 2021). Additionally, single mothers are more likely to be poor (Office for National Statistics, 2019), which increases parental stress levels and therefore risk of perpetrating abuse (Budd et al., 2006).

Thus, it is not possible to extrapolate from descriptive statistics alone, and from the findings in the current review, whether mothers *per se* are at a higher risk of perpetrating child abuse. Gender differences in child abuse potential may be further explained by the fact that women perceive their role differently relative to men and may have different perceptions, coping abilities, and experiences of stress and differences in terms of willingness to seek help (Fang et al., 2022). There is not enough research to comment on the strength or importance of these gender-based differences in child abuse potential obtained here, as they were reported only in one or two studies. Future studies using the CAPI Abuse scale would thus benefit from further exploring potential gender-based differences when controlling for the abovementioned confounders and ensuring an equal, if not overrepresentation of men, in the study.

Responsiveness

Examination of responsiveness was carried out in a relatively small number of studies. However, four of these were RCTs which constitutes a methodological strength of these findings, since random allocation of study samples to either an intervention or control group minimises selection bias and ensures similar sample characteristics across groups. Accordingly, RCTs are the most powerful study designs for estimating unbiased effect sizes of an intervention (Yoon et al., 2022). However, as mentioned, relatively few RCTs have been carried out on the effectiveness of interventions to reduce the risk of child physical abuse as measured by the CAPI Abuse scale, and specifically, among maltreating parent or caregiver samples. Indeed, among the nine available studies on responsiveness, four were RCTs of which only three provided sufficient data to obtain estimated effect sizes. Two other

studies were low quality cohort studies which represent a methodological shortcoming, since they conducted less robust statistical analyses, such as paired *t*-tests and reporting only *p*-values. The latter are considered less appropriate for measuring responsiveness. First, *p*-values can only confirm whether 'before and after' and 'between group' differences in treatment scores are statistically significant (i.e., not due to chance), which does not speak to whether the estimated difference is clinically meaningful. Second, unlike effect sizes, *p*-values depend on sample size (Columb & Atkinson, 2016). For this review, which included studies that indirectly examined psychometric properties of the CAPI Abuse scale, studies were rated as adequate or very good if they provided sufficient data to calculate Hedges' *g*, which is the preferred indicator of responsiveness in the COSMIN risk of bias checklist. Notwithstanding, future studies using the CAPI Abuse scale would be encouraged to report on effect sizes to contribute to a higher quality evidence base on this measure.

Overall, the quality of evidence in relation to CAPI post-treatment (i.e., before and after) change responsiveness was moderate, and the findings suggest that CAPI Abuse scale can detect changes at a group level following several types of tertiary parent interventions. The overall quality of evidence between subgroup responsiveness findings was low due to inconsistent results across studies. The finding on heterogeneity is partly consistent with a previous meta-analysis by Chen and Chan (2016), which examined the effects of preventative parenting interventions aimed at reducing child maltreatment. Their meta-analysis obtained a wide variation of effect sizes among studies that used the same measure; the authors found that sample (e.g., country and gender) as well as intervention (e.g., dosage) characteristics contributed to significant differences between studies. Though in the present review, study samples did not vary so much in terms of gender and country, since they were mostly female caregiver samples from north America or Australia, samples did vary in terms of drug abuse habits, parent, and child age as well as in terms of the duration and types of interventions received, such as PCIT, SafeCare, and family behavioural therapy and so on. Thus, future research should also focus on elucidating what

variables contribute to the heterogeneity of effect sizes across studies on parenting interventions, specifically for the CAPI Abuse scale. This will help better determine CAPI sensitivity to between intervention group differences.

Regardless of differences between studies, the lack of treatment-change responsiveness observed in some of the included studies may be due to the intervention and study limitations. This is supported by the fact that for several studies in which no significant treatment changes were observed in CAPI Abuse scale scores, also did not obtain differences on other primary outcome measures (e.g., parental depression, child maltreatment recidivism). As such, in these cases it is possible that the studies were simply underpowered to detect significant treatment effects or as argued by some study authors, that the intervention did not target what it aimed to (Hubel et al., 2018). For studies in which treatment effects were obtained on other primary outcome measures, it possible that the mechanism of change was not targeting factors underpinning CAPI. Given the relatively small number of intervention studies examining CAPI change scores in maltreating parent populations, future research could perform analyses correlating CAPI outcomes with other primary outcome measures used in in the intervention to clarify these inconsistent findings between outcome measures.

Limitations

One of the strengths of the systematic review is the fact that study inclusion and methodological quality were assessed by two independent raters and potential disagreements were resolved by the field supervisor, making the inclusion process more reliable. Another strength was the inclusion of high-quality studies according to COSMIN guidelines (Mokkink, Prinsen et al., 2018). Furthermore, this review adds to the existing literature by presenting a focussed evaluation on three important psychometric properties of one of the most widely used child abuse instruments among high-risk parents by using the updated COSMIN manual. Focusing on a high-risk parent population is useful for tertiary

prevention efforts and policies targeting maltreating parent populations. Though most included studies used self-report measures to collect data, some studies also used observational tools and coded responses to open ended questions. This is important to the extent that observation data can provide more objective information compared to self-reports. Studies also conducted checks on inter-rater reliability, thus further increasing the validity of the findings.

However, several limitations should also be noted. First, as mentioned, this review focused on evaluating two main psychometric properties of the CAPI Abuse scale, i.e., construct and criterion validity and responsiveness. Though beyond its scope, this review cannot, therefore, provide a full picture of the psychometric properties of the CAPI Abuse scale, as the remaining psychometric properties (i.e., reliability, content, structural-, and cross-cultural validity) were not assessed.

Second, only studies published in English were included, resulting in the exclusion of studies published in any other language, which could have resulted in missing valuable information on its psychometric properties. Third, the use of the COSMIN checklist (Mokkink et al., 2018), which, although appropriately validated and well-structured, as critiqued by McKenna and Heaney (2021), entails subjectivity in the interpretation of the information provided from the studies and how to apply the guidelines as per COSMIN manual (Prinsen et al., 2018). For example, standards for convergent validity state that for the statistical methods to have been applied appropriately, it required the article to include the distribution of scores or mean scores. However, the distribution of constructs depends on the population under study, which, as is the case for this review examining parents at high-risk of maltreating their children, unlikely presents a normal distribution on the constructs being measured.

This review considered outcome studies as eligible for inclusion, which is not in line with recommendations in the COSMIN guideline (Mokkink et al., 2018; Prinsen et al., 2018). This adaptation was made to maximise the likelihood of obtaining all relevant findings on

CAPI Abuse construct validity and responsiveness within a more focussed sample of parents and caregivers at high risk of maltreating their child/ren. Consequently, this review considered the testing carried out within the outcome studies as 'psychometric testing' even when this was not the original aim of the study conducting that test. As a result, this impacted the quality assessment of the studies; indeed, COSMIN quality assessment standards require the studies to have formulated a priori hypotheses concerning effect sizes, though this is not necessarily relevant in an outcome study examining relationships between variables for instance. This, in turn, required adaptations of the standards, which would have increased subjectivity during the quality evaluation. Including outcome studies may have also increased subjectivity during the study selection phase, due to the higher number of papers eligible for inclusion.

In terms of responsiveness, this review reported only on two aspects of the construct approach for responsiveness (comparison before and after an intervention and comparison between subgroups; Mokkink et al., 2010). The other aspect (i.e., comparison with other outcome measures) was not possible given that no longitudinal study (including intervention studies) with at least two measurements reporting either the relationship between the change scores on CAPI Abuse scale and instruments assessing similar constructs, were not identified in the selection phase. Thus, longitudinal research on the CAPI Abuse scale among high-risk parents is required to fill this gap.

Furthermore, more than half the studies were conducted on female caregiver samples and in some of the remaining studies females were overrepresented. These findings point to a general trend in parenting-related studies in which fathers are generally underrepresented (Fang et al., 2021). This is in contrast however with the changing culture of parenthood, which over the past few decades has seen an increase in caregiving fathers. The findings in this review are, therefore, predominantly applicable to a high-risk female population. In future studies, representative samples of mothers and fathers are thus

recommended to allow further examination of the CAPI's psychometric properties among male caregivers.

Finally, though samples in the included studies comprised ethnically diverse high-risk caregivers, all studies (except one in Finland) were performed in a very limited number of western countries. This is important insofar as psychometric properties of CAPI Abuse within high-risk caregiver populations remain unexamined in a big proportion of countries and cultures which might present different patterns of understanding and experiencing of this complex topic. Indeed, for some cultures a good child education may be more synonymous with physical chastisement (Vachon et al., 2015). These findings reflect ones found in a meta-analysis examining prevalence estimates of child abuse across the world, which pointed out that child maltreatment research is mainly concentrated in western cultured countries (Stoltenborgh et al., 2015). Whether the construct validity and responsiveness of the CAPI Abuse scale is generalizable to high-risk caregiver populations in non-western countries still requires further investigating.

Recommendations

Regarding hypothesis testing for construct validity, convergent validity testing would benefit from studies having a priori hypotheses concerning the size of the correlation they would expect to find. In addition, to better estimate unbiased effect sizes on convergent validity, future research on child abuse potential using CAPI Abuse scale should examine CAPI risk factor covariation. Indeed, the issue around collinearity among child physical abuse risk factors has recently been raised by Milner et al. (2022). Collinearity and multicollinearity refer to when two risk factors and more than two child physical abuse risk factors, respectively, are highly correlated. When these conditions are present, it is likely that one risk factor can be used to explain the variation in the other risk factor. Although it appears as though there were more than one risk factor, they may not all represent independent child physical abuse risk factors but rather a single underlying risk dimension. Validity data using

findings from correlations where covariation is present may be at risk of over-estimating convergent validity. Thus, more multivariate designed studies are encouraged to account for possible collinearity effects when assessing possible CAP risk factors. In this vein, future studies should also consider accounting for the effects of child characteristics among maltreating parent samples given the paucity of research obtained here.

With regards to the criterion and known-group validity of CAPI Abuse, future studies would benefit from more tightly matched control groups or where possible to control for any obtained differences. Crucially, research should continue to evaluate CAP criterion validity using gold standard measures such as 'known instances of child physical abuse' given that there was only one study that applied this standard, and which did not obtain sufficient results.

Last, from the findings on methodological quality of included studies on responsiveness among parents at high-risk of maltreating their child, more research is encouraged to report and calculate effect size in addition to *p*-values, which is also in line with APA's reporting standards. Building on the responsiveness findings obtained here, future research could consider the extent to which the CAPI Abuse scale is able to assess clinically significant change; this recommendation is consonant with that suggested by Walker and Davies (2010), which highlighted the need for there to be further clarity on what the changes in CAP scores after an intervention reflect. In this way, clinically significant change could be examined by calculating CAPI Abuse scale's reliable change index, which indicates whether post-treatment changes in CAP scores are more than would be expected by chance by controlling for error variance in the CAPI (Jacobson & Truax, 1991).

Conclusion

This review aimed to evaluate the methodological quality and psychometric properties of studies using the Child Abuse Potential Inventory (CAPI) Abuse scale among parents at high-risk of maltreating their children aged 0-12. The included studies, which were mostly conducted in North America and directly and indirectly reported on CAPI measurement properties, provided evidence in support of CAPI construct- and criterion validity and responsiveness. In line with theoretical predictions, the CAPI Abuse scale related to multiple parent-related and social-contextual factors, thus supporting its construct validity. Crucially, several studies showed that there were significant differences on CAPI Abuse scale scores between maltreating parents and non-abusive counterparts, thereby supporting its criterion validity; CAPI Abuse scale scores also distinguished between other known groups, such as parents with and without childhood histories of trauma, respectively. Intervention studies reporting on responsiveness showed that post-treatment changes could be observed on the CAPI Abuse scale scores, however, the findings were more mixed for differences between treatment conditions. Whilst the CAPI Abuse scale should continue to be used in research and clinical settings among high-risk parents, further studies are required to directly examine CAPI Abuse psychometric properties, namely comparing it to gold standards and using stronger methodological designs before its use can be fully endorsed in epidemiological and intervention research.

Predictors of child abuse potential in a sample of parents at-risk of maltreatment: the role of epistemic trust, parenting stress, and parental representations

Abstract

Theory and research suggest that parents who were abused as children are more likely to repeat abuse in the following generation. One model of explanation suggests that this cycle is partly owing to disruptions in the acquisition of mentalizing abilities, to low levels of epistemic trust and to further disruptions in mentalizing brought about by high parental stress observed in high-risk parents. The present study aimed to assess whether parental epistemic trust domains, parental risk representations and parental stress would associate with the risk of parental child abuse, using the Child Abuse Potential (CAP) Inventory. Secondary data analysis was conducted on baseline data collected as part of a parenting intervention aimed at reducing child maltreatment in high-risk parents. One hundred and ten parents with child protection service involvement were interviewed on the Parent Development Interview (Slade et al., 2004) about their views on themselves as parents, a specific child, and their relationship to that child. The transcripts were coded using the Assessment of Representational Risk (Sleed, Isosävi et al., 2021). Participants also completed measures assessing their epistemic trust, mistrust and credulity and their perceived parental stress levels. Adverse childhood experiences data collected at post intervention were also analysed. Epistemic trust showed a small, non-significant negative relationship to CAP, whilst higher levels of epistemic mistrust and credulity were significantly associated with higher levels of CAP. Higher risk in parental representations was also significantly associated with higher CAP. Last, parental stress showed a large, positive correlation to CAP. When analysed together only parental stress predicted CAP, and there was a trend of representational risk approaching significance as a CAP predictor. Overall, the findings suggest that parental stress is a salient factor in CAP and that parents' epistemic stance, and parental representations may be relevant to CAP among high-risk

parents with childhood histories of maltreatment. The results have implications for treatment models, social care, and future research.

Introduction

Child maltreatment (CM) in the family context is defined as any interactions (or lack of) on the part of the caretaker which result in harm to the child's health or physical, mental, spiritual, moral, or social development in the context of the society in which the child grows up (WHO, 1999). Sadly, CM is relatively common affecting children worldwide (Stoltenborgh et al., 2015) and can have lifelong impacts (Winter et al., 2022). Investigating potential pathways and factors related to CM has thus become a priority for many researchers, in efforts aimed to develop treatment programs and policies that can help decrease its occurrence, and ultimately the damage that it yields (van IJzendoorn et al., 2020).

In terms of understanding putative mechanisms of CM, it is widely accepted that it is a complex phenomenon, best understood by adopting a whole system perspective (DePasquale et al., 2019). Belsky (1993) noted the complex multilevel structure of CM risk factors in his Ecological Integration Model of Child Abuse. This model locates risks at four different "levels of analysis", each varying in distance to the child. Parent- and child-related factors occupy the level closest to the child; the second level comprises family-related factors and the third and fourth levels represent risk factors present in the community and wider system, respectively, to which the family belongs. An interactive accumulative effect between risk and protective factors is suggested at each level (Baldwin et al., 2020), therefore making it difficult to accurately predict and identify children who may be at risk of CM. It has been argued, however, that given how proximal parent-related risk factors are to the child, such factors are likely to determine CM occurrence most strongly (Assink et al., 2018). Consistent with this argument, in two meta-analytic reviews the largest effect sizes (ES) were obtained for a variety of parent-related risk factors (Mulder et al., 2018; Stith et al., 2009). Such parental factors included poor mental health, parental stress, alcohol and/or drug abuse, and poor parenting skills. Similarly, a relatively recent umbrella synthesis of meta-analytic reviews examining the antecedents of CM, also derived the strongest ES for parent-level factors, followed by intimate partner violence and socio-economic status (van

IJzendoorn et al., 2020). Together, these review findings point to the importance that parent-related factors may have in explaining potential underlying mechanisms of CM.

Parenting behaviours are significantly influenced by a parent's own experiences of childhood (Assink et al., 2018). This notion is advanced by the 'intergenerational transmission theory', which states that parents raise their children in ways that resemble their own upbringing (van IJzendoorn, 1992). As a result, positive as well as abusive parenting practices can be passed down from one generation to the next.

Drawing on this theory, an extensive body of research has investigated the impact of child abuse and neglect in terms of its intergenerational continuity and transmission, which is also referred to as the 'cycle of violence'. Research findings generally support the idea that a history of CM increases the chances of maltreatment perpetration in the following generation. For example, families with childhood histories of victimization show a significantly higher number of risk factors, including a higher incidence of punitive parenting styles, than families without abusive childhoods (Dixon et al., 2005). Another study examined the intergenerational transmission of abuse among a smaller sample of teen, adult low-resource, and adult high-resource first-time mothers. The authors found that higher levels of past exposure to emotional and physical abuse were associated with decreased maternal responsiveness and increased endorsement, and propensities for abusive behaviour (Bert et al., 2009). Finzi-Dottan and Harel (2014) found that the risk for maltreating children was six times greater when parents had experienced CM themselves. Overall, these empirical findings among others (e.g., Jaffee et al., 2013; Schofield et al., 2013; Thornberry et al., 2013; Widom et al., 2015) support the principle that 'maltreatment begets maltreatment'.

Meta-analytic review findings also converge toward this theory. In one meta-analysis on studies performed in families of parents who experienced maltreatment in their own childhood, the odds of CM were almost three times those observed in families of parents without CM histories (Assink et al., 2018). Similarly, in the already-mentioned umbrella synthesis of meta-analyses on the antecedents of CM (van IJzendoorn et al., 2020) 'parental

experiences of CM' was derived as the risk factor with the strongest ES. Taken together, the review findings further corroborate the intergenerational transmission of maltreatment hypothesis that individuals with childhood histories of abuse are at an increased risk of maltreating their own children in adulthood.

The empirical literature on the intergenerational transmission of abuse, however, is not without methodological weaknesses (Jaffee, 2017; Thornberry et al., 2012). Limitations such as small sample sizes, the lack of valid data on abuse and neglect across the generations (Widom, 1989), non-representative study populations, and differences between studies, which adopted either a retrospective or prospective design, have made it difficult to establish valid prevalence estimates of abuse transmission (Dixon et al., 2005). Importantly, however, regardless of the variability in transmission estimates reported, the data point to an important fact: most individuals with maltreatment history do not themselves go on to become perpetrators. Therefore, though it is likely that abuse gets played out across the generations, the cycle is not *inevitable*. Indeed, across most studies, the majority of adults who were abused as children did not go on to perpetrate abuse or neglect. For example, in an English prospective study, which followed up new-borns during the first five years of life, only 7.6% of parents with a history of abuse were found to have abused and/or neglected their child (Browne & Herbert, 1997).

Many factors are likely to shield or counteract risk and reduce the likelihood of perpetrating abusive and neglectful behaviours within the next generation. An emerging model of explanation, which is gaining importance in the current understanding of intergenerational cycles of maltreatment, lies within the concepts of mentalization, epistemic trust, and attachment (Asen & Fonagy, 2017). According to this developmental pathology framework, these constructs underpin the lack of resilience observed in caregivers who continue the cycle of violence (Byrne et al., 2019; Fonagy et al., 2017; Slead, Fearon et al., 2021).

Epistemic trust refers to a person's trust that what they are being shown and taught is trustworthy, relevant to them and generalizable to other people and settings (Orme et al., 2019). This capacity has been deemed to provide humans with an evolutionary advantage by addressing the 'learnability problem' (Csibra & Gergely, 2006, 2009, 2013); that is, the fact that humans enter a world populated by objects, customs, attributes, and human minds which are 'epistemically opaque' (i.e., whose purpose, function or meaning is not immediately obvious from appearance). Without being open and trusting toward the knowledge shared by others, humans would be left with the difficult, time consuming and often impossible task of working cultural knowledge out for themselves. Thus, theoretical, and empirical research proposes that epistemic trust is a crucial evolutionary adaptation; it opens the channel for a fast transmission of intergenerational knowledge, owing to which individuals can positively benefit from their social environment. As a result, humans can take in huge amounts of sophisticated, cultural information and develop resilience owing to their high social functioning (Fonagy et al., 2017).

Importantly, however, though humans are evolutionarily primed to be open to information from others, this is not the default setting. Individuals still require capacity to judge whether what they are being told is accurate or not. Human complexity is such that some people may not always be accurate and may instead be unreliable due to a lack of knowledge themselves or even deliberate attempts to deceive. This is supported by studies showing that already early on, infants instinctively look for cues from their primary caregivers that will allow them to better understand and act in ambiguous situations (e.g., Fonagy et al., 2017). These experiments have highlighted that infants exhibit *epistemic vigilance*, that is, they display scepticism and distrust towards information shared by adults (Fonagy et al., 2017). Such vigilance can therefore be understood as having a protective function, as excessive epistemic trust would not otherwise allow infants to discern whether the communicator and information are reliable.

Theoretical arguments suggest that secure attachment relationships provide the ideal setting to first acquire epistemic trust because here the child can safely relax their epistemic vigilance and invest in what the caregiver is sharing (Fonagy & Allison, 2014). Fonagy and colleagues (2014; 2015) argue that high levels of mentalizing normally observed within this secure context (Zeegers et al., 2017) play a key role in this regard. Mentalization refers to the imaginative ability to make sense of one's own and others' behaviour in terms of intentional mental states, such as needs, wishes, and desires (Fonagy et al., 2002). Caregivers with good mentalizing abilities show curiosity in their child's mind and are contingently 'marking' and 'mirroring', as well as attending to their child's inner world (Fonagy & Luyten, 2016; Slade, 2005).

Repeated experiences of being accurately held in mind this way provides the growing child with a feeling of being understood and that the caregiver is interested in them as a person; it is this feeling which is thought to stimulate trust that was is being shared with them is relevant and generalizable. In other words, "if you can know and understand me like *that* what *else* have you got for me?" (Gergely, 2013). Theory and research propose that relational safety and caregivers' mentalization supports the child to make sense of their own mind and encourages them to explore that of their caregivers, therefore giving rise to mentalizing abilities of their own (Fonagy et al., 2014). It is suggested that these abilities, in turn, support the growing capacity to discern when to relax one's epistemic vigilance, as the growing child will be better placed to make sense of other peoples' opaque states of mind that would then justify depending on their knowledge (Luyten et al., 2020).

It is unlikely that abusive parents with childhood histories of maltreatment grew up in environments as described above. Instead, their homelife was more likely characterised by experiences of "being unseen and unheard, and mis-seen and misheard" (Byrne, 2020, p. 207). There is increasing evidence suggesting that attachment trauma and adverse childhood experiences are associated with disruptions in developing the capacity to establish healthy attachment relationships and epistemic trust, as well as the capacity to mentalize

(Luyten et al., 2020). This is supported by robust evidence showing that mentalization based interventions which directly work on issues of epistemic trust are effective at engaging ‘hard-to-reach’ adults who have complex histories of attachment trauma or neglect, poor emotion regulation, and difficulties building stable trusting relationships (Bateman & Fonagy, 2016; Bateman et al., 2016; Byrne et al., 2019). Consequently, for parents with childhood histories of victimization, projections and simulations from their own past may obscure an accurate view of the child (Byrne, 2020; Fonagy et al., 2018). This may be reflected in the development of ‘parental mental representations’ or ‘internal working models’ of themselves as parents, of their children and the relationship to their child, which are distorted, giving rise to atypical parenting behaviours that are seen in parents of infants with disorganised attachment (Sleed, Isosävi et al., 2021). Such representations may result in a mis-use of mentalization that manifests in bizarre, hostile, or inappropriate attributions of mental states or through an absence, or defensive denial of the emotional world of the infant and/or themselves (Sleed, 2014).

Further, disruptions to their epistemic trust means that these parents may operate in a state of chronic uncertainty and vigilance (Bateman & Fonagy, 2016). Consequently, they may reject or avoid any new information, which impedes their ability to learn from social experiences (Fonagy et al., 2017). This may underpin the rigidity and often described ‘hard-to-reach’ nature of high-risk families known to Child Protection Services (CPS). Disruptions in the acquisition of ordinary mentalizing and epistemic trust may also lead to the individual not exercising sufficient vigilance, instead occupying a state of epistemic credulity; information may, thus, be received with insufficient discrimination, leaving the recipient vulnerable to misinformation and exploitation. Recently, confirmatory, and exploratory factor analyses have confirmed the three epistemic stances described above: trust, mistrust, and credulity (Campbell et al., 2021); importantly, the study also found cross-sectional links between epistemic mistrust and credulity, and childhood experiences of adversity, providing preliminary support to the idea that adverse childhood experiences disrupt epistemic trust.

Drawing on the developmental pathology framework outlined above, abusive behaviour observed in parents with childhood histories of victimization can be conceptualized as arising in the context of chronic epistemic mistrust (Byrne et al., 2019); not only do these individuals lack trust in information from adults, but they also distrust their child's communications. For instance, they may feel as the victim of their children's behaviours and perceive them as 'deliberately manipulating them'. In these circumstances, children can often be described as "controlling", "knowing the impact they are having on the parent", and being "selfish", etc. In more extreme cases, real clinical examples highlight how in the face of challenging behaviour, such parents may describe their traumatised child as being "the devil" or their five-year-old as possessing a "criminal mind" (Byrne, 2020; Slead, Isosävi et al., 2021). Conceivably, inaccurate mentalizing of the child as reflected in distorted parental representations of this kind, may more likely result in instances of physical chastisement and/or emotional abuse.

Furthermore, this framework also considers the effects of parenting stress. Parenting is naturally stressful and understandably, even for the average parent, leads to lapses in mentalizing (Slead, Fearon et al., 2021). As such, the strong and overwhelming emotions, such as worry, guilt, anger, joy etc., that it involves, can impinge on a parent's ability to mentalize and to stay curious in their child's behaviour (Byrne et al., 2019). During these moments, parents may make snap judgements about their child's intentions. Accordingly, for high-risk parents who already struggle to mentalize and emotionally regulate, and who often have on going life stressors, parenting may even more likely evoke lapses in mentalizing (Asen & Fonagy, 2017). The resulting bizarre, or hostile misattributions made about the child's behaviour may also more likely lead to instances of physical chastisement and emotional/physical abuse (Richey et al., 2016). Furthermore, high-risk parents who go on to be referred to CPS, can often feel undermined in their confidence to raise their child, thereby exacerbating their already heightened levels of stress (Byrne et al., 2019). This is important to the extent that parenting stress has also been shown to mediate the link between parental

history of maltreatment and parental sensitivity (Pereira et al., 2012) and can impinge on the capacity to mentalize (Nolte et al., 2013). Accordingly, parental stress may be more strongly associated with risk of CM in this parent population.

Currently, there is a treatment model under trial, the Lighthouse Parenting Programme (Byrne et al., 2019; Slead, Fearon et al., 2021), which aims to reduce CM in high-risk and maltreating parents by targeting epistemic trust, mentalization, high-risk parental representations, and parental stress. Though, as described, theoretically these factors represent valid therapeutic targets to decrease CM, no study to date has examined whether they are associated with CM in high-risk parenting groups. Therefore, it is important to test the theoretical link between these concepts and CM empirically. Notably, this has recently become possible owing to the development of reliable and valid empirical measures of epistemic trust (Campbell et al., 2021) and low mentalizing, as captured by the assessment of representational risk (Slead, Isosävi et al., 2021).

Objectives

The present empirical project was a cross-sectional study of a high-risk clinical sample of caregivers known to CPS, with their own childhood histories of abuse. The study used the Child Abuse Potential (CAP; Milner, 1986) Inventory to explore the risk of parental child abuse according to the developmental pathology framework outlined above. Therefore, the present study had the following aims: first, to evaluate the degree to which high-risk parents' CAP scores were associated with parents' self-reported parental stress, epistemic trust, mistrust, and credulity, as well as their parental representations; second, to assess whether parents' epistemic stance and parental representations in combination with parental stress and parents' adverse childhood experiences would independently or additively increase the amount of CAP variance explained. Given the paucity of prior research on these constructs, the current study aims were considered exploratory.

It is important to note, that the different domains of epistemic trust, as well as parental representations and parental stress are predicted to be *correlates* of CAP. There is no assumption of causality for any of the factors. Furthermore, it is readily acknowledged that this group of factors is not exhaustive. Risk factors that were included were ones that have strong theoretical and empirical support as valid correlates of CAP in a sample of high-risk parents among which intergenerational abuse is present.

To summarise, based on the literature reviewed above it was hypothesized that:

- 1) Parental epistemic trust, and parental epistemic mistrust/credulity would be negatively and positively associated with increased child abuse potential, respectively.
- 2) Parental representational risk would be positively associated with increased risk of parental child abuse.
- 3) Parental stress would be positively associated with increased risk of parental child abuse.
- 4) Parental epistemic trust, mistrust and credulity, and parental representational risk, would lead to significantly increased levels of child abuse potential after controlling for parental stress.
- 5) Parental epistemic trust, mistrust and credulity, and parental representational risk, would lead to significantly increased levels of child abuse potential after controlling for parental stress and parents' adverse childhood experiences.

Methods

Design

This study was a secondary analysis of data collected as part of the Supporting Parents Project (SPP); the SPP protocol is described in Sled, Fearon et al., (2021). The SPP was a two-arm partially clustered Randomised Controlled Trial (RCT) investigating the effectiveness and feasibility of the Lighthouse Parenting Programme (LPP) in Children's social care compared to social care as usual. The LPP is a mentalization-based intervention for parents with children known to Child Protection Services (CPS). The LPP aims to reduce parental CM through targeting parental mentalizing, high-risk parental representations, epistemic mistrust, and parental stress. The current study used a cross-sectional, correlational design on pre-treatment, baseline data that were collected as part of the SPP prior to the LPP intervention.

Participants

Study participants ($N=110$) were recruited according to the following criteria: Participants had to 1) have at least one (target) child/ren aged 0-12 years old; 2) have been identified as having caregiving difficulties that warranted the child being on a (CP) plan, (CiN) plan, or in pre-proceedings. Participants were excluded if 1) the target child was in care proceedings at the time of recruitment; 2) the referring professional considered the family likely to proceed to care proceedings in the six months that followed; 3) the referring professional considered the parent to be unsuitable for group-based interventions in the event that they would compromise the safety of others in a group setting (e.g., if the parent had a diagnosis of Anti-social personality disorder); 4) the parent had been a perpetrator of sexual abuse or had a history of sexual predatory behaviour; 5) the parent had been a perpetrator of sadistic abuse of children (i.e., perpetrating deliberate physical harm/torture); 6) the parent had severe learning disabilities; and last, if 7) the parent had acute psychosis.

Five local authority sites across England were involved in the SPP recruitment pathway, which took place during August 2021. Consecutive referrals were made from each site until enough parents consented to participate in the site cluster ($N=20-24$). The SPP evaluation team then contacted the potential participant to provide them with further study information. A total of 190 parents were initially referred to the SPP, of which 80 did not participate for several reasons including: the parents could not be contacted, they declined to take part, or the social worker withdrew the referral due to a change in the family's circumstances. Therefore, of the initial possible sample, 61% participated in the SPP.

As a result, the final sample consisted of 110 participants who consented to participate in the SPP project. As per inclusion requirements, participants were parents at high risk of mistreating their child/ren aged 0-12 years old, who had been identified as having caregiving difficulties by CPS, thus resulting in the target child being on a CP, CiN or a pre-proceedings plan. Demographics data are included in Table 1. The final sample mostly comprised birth mothers and fathers (91%). Almost all participants were born in the United Kingdom ($N=107$, 97%) and only three parents were born in another country (South Africa, China, and India). All parents spoke English fluently, with the majority (94%) describing themselves as being of white ethnicity. Most families belonged to low-income brackets, with only 10% receiving an annual household income of £30 000 or above and only a quarter of participants in paid employment. Families mostly lived in council, housing association or rented private accommodation. The number of children living in the household ranged from 1-8, with 75% of the sample having two or more children living in the household. Children of participating parents ranged in age from unborn babies to 18 years old. Participants chose to focus on one key child in the data collection. The average age of the key child was 7 years old. Comparing the final sample ethnicity with that of the wider population in the five local authority areas from which participants were recruited showed that the latter also mostly comprises families of white ethnic background (UK Census Area Profile, 2021). Last, a comparison of the final study sample ($N=110$) with the total number of participants who

consented to be approached by an SPP research officer ($N=190$) could not be conducted due to lack of consent.

Table 1

Demographic statistics for the sample

Characteristics	Total (%) N=110
Parent	
Gender	
Female	87 (79%)
Male	22 (20%)
Non-binary	1 (1%)
Age (years): mean (range)	32.98 (20-63)
Ethnicity	
White	103 (94%)
Other than White	7 (6%)
Single parent household	73 (66%)
Work status	
Employed	29 (27%)
Not employed looking for work	68 (64%)
Not employed and not looking for work	10 (9%)
Relationship status	
Divorced	8 (.07%)
Registered civil partnership	1 (.01%)
Married	17 (15.5%)
Never married and never registered	75 (68%)
Separated but still legally married	7 (.06%)
Widowed	1 (.01%)
Household yearly income category	
Prefer not to say	3 (3%)
Under £10 000	28 (26%)
£10 000 - £20 000	56 (52%)
£20 000 - £30 000	11 (10%)
Over £30 000	11 (10%)
Number of children living in the household	
1	27 (24.5%)
2	41 (37.3%)
3	21 (19.1%)
4	16 (14.5%)
5	3 (2.7%)
6	1 (.9%)
8	1 (.9%)
Secondary school qualification	
AS/A level or equivalent	14 (13%)
GCSE or equivalent	69 (64.5%)
NVQ or equivalent	8 (7.5%)

No qualification	16 (15%)
Source of household income	
State benefits only	73 (66%)
State benefits and earnings	24 (22%)
Earnings only	13 (12%)
Target Child	
Child gender	
Female	48 (44%)
Male	60 (56%)
Child age (years): mean (range)	6.9 (unborn – 15)
Social care status (at randomization)	
Closed to social care	8 (7%)
Targeted team around the child	1 (1%)
Child in Need plan	46 (42%)
Child Protection plan	43 (39%)
PLO proceedings	5 (5%)
Care proceedings	1 (1%)
Child Looked After	4 (4%)

Note. PLO: Public Law Outline; AS/A Level: Advanced Subsidiary/Advanced level; GCSE: General Certificate of Secondary Education; NVQ: National Vocational Qualification

Ethics

NHS ethics was not sought, and University College London (UCL) ethics was gained instead (Project ID number: 9593/002); an amendment was made to the original UCL ethics SPP protocol, which was approved in April 2022 (see Appendix F). It was then extended to Royal Holloway University of London (RHUL) through self-certification for a third-party project (see Appendix F).

Power analysis

The current study conducted secondary data analysis, which precluded control over the sample size. Therefore, the effect size was constrained by the overall number of participants obtained in the SPP ($N=110$). The sample size for the SPP study had been calculated consistent with its aim of being able to estimate the effectiveness of the LPP intervention, based on the SPP study design: a two-arm, partially clustered randomized controlled trial of LPP versus Children's social care as usual.

For the current study, post-recruitment sensitivity calculations for the linear hierarchical regression models (hypotheses four and five) indicated that the smallest ES that the secondary data analysis would be able to detect with the final sample ($N=110$) was $f^2=.106$, which corresponds to a $R^2 = .095$ (G*Power: sensitivity analysis, linear hierarchical multiple regression, $\alpha=.05$, power=.80, numerator $df=5$).

A post-hoc power analysis was also conducted to determine the number of participants that would have been needed for the multivariate analyses in this study. The post-recruitment power calculations using standardized ES conventions for this thesis indicated that for a linear multiple regression (fixed model, R^2 increase) with three tested predictors and five total predictors, a minimal sample size of 71 would have been required (G*Power: linear multiple hierarchical regression, $ES = .15$, $\alpha=.05$, power=.80, numerator $df=5$). This suggests that the current sample of 110 was sufficient to test the study hypotheses. However, although previous studies have found large ES between parental characteristics (e.g., parent anger hyperactivity, and parent self-esteem) and child abuse potential as measured by the CAPI Abuse scale (Stith et al., 2009), given that several of the measures proposed in this study had not yet been investigated in similar samples, power calculations for a small ES would have been preferable (Schafer & Schwarz, 2019). As a result, the current study was likely to be somewhat underpowered to detect small associations in the sample.

Measures

The measures and semi-structured interview are presented below in the order in which they were administered. The first measure was a non-standardized demographics questionnaire devised for the original study. This questionnaire included items on the following: parent age, parent gender, ethnicity, relationship status, number of parents in the household, employment status, parent yearly income category, source of the household income, target child age, target child gender, and social care status.

Epistemic Trust, Mistrust and Credulity Questionnaire (ETMCQ; Campbell et al., 2021)

The ETMCQ was used to evaluate participants' epistemic trust, i.e., their level of trust in communicated knowledge (see Appendix G). The ETMCQ is a self-report questionnaire consisting of 15 items to measure the three independent subscales of the epistemic trust construct: epistemic trust, mistrust, and credulity. Items are rated across a seven-point Likert-scale ranging from “strongly disagree” (=1) to “strongly agree” (=7). Subscale scores range from 5 to 35, with higher scores indicating higher Trust, Mistrust, and Credulity, respectively.

Campbell and Colleagues (2021) found acceptable internal consistency for the full ETMCQ scale ranging from Cronbach's $\alpha=.71$ to $\alpha=.78$; in the present sample, good internal consistency was observed for the Trust ($\alpha=.841$), Mistrust ($\alpha=.754$) and Credulity ($\alpha=.807$) subscales. Campbell and Colleagues (2021) also found good test-retest reliability and intra-class correlation coefficients in representative UK samples. Further, in their study, the subscales demonstrated construct validity by behaving in line with theoretical predictions, such that Mistrust and Credulity scores were associated with childhood adversity and higher scores on the global psychopathology severity index.

Assessment of Representational Risk (ARR; Slead, Isosävi et al., 2021)

The ARR was used to measure features of parental representations that are theoretically and empirically linked to attachment disorganisation, and to detect misuses in caregivers' mentalization (see Appendix G). The ARR is a coding system that is applied to transcripts from the short version of the Parent Developmental Interview (PDI-S; Slade et al., 2020); the PDI-S is a 30-item, semi-structured, clinical interview that asks parents to reflect on themselves as parents (e.g., “What gives you the most pain in being a parent?”), on a specific child (e.g., “What do you like most about your child?”) and their relationship with that child (e.g., “How do you think your relationship with your child is affecting his/her

development of their personality?”). In the current study, parents were asked to focus on the child they were ‘most worried’ about, the majority of which were receiving CiN, CP plans or were under PLO orders.

PDI interviews were audio-recorded and transcribed verbatim. Verbatim transcripts were then coded by four raters who had undergone a three-day training with the ARR developer and had passed a reliability test by obtaining good inter-rater reliability scores (see Appendix G). The coders were the author of the present study, two SPP research officers, and a psychotherapy doctoral student. The raters were all blind to the interviewee data. The interrater reliability for the Total ARR score is not yet available as a second rater is currently in the process of double coding 20% of the baseline interview transcripts.

The first eight items of the ARR reflect high-risk features of representations that have been empirically associated with problematic relational and child outcomes (i.e., hostile parental experience, hostile parental behaviour, fearful affect, helplessness, emotional distress, idealisation, enmeshment/role reversal, and incoherence) (e.g., George & Solomon, 2008; Lyons-Ruth et al., 2006; Main & Hesse, 2006). The last two scale items (i.e., supportive presence, mutual enjoyment) reflect positive protective factors that have been previously related to secure attachment, the lack of which thus indicates risk (e.g., Burns et al., 1997; van IJzendoorn et al., 1999). The ratings were based on the frequency and intensity with which the dimension emerged in the narrative, ranging from one (lowest frequency and/or intensity of the dimension) to five (highest frequency and/or intensity of the dimension); no evidence of the construct in the narrative would be scored at the lowest end of the scale. An ARR Total Risk score was computed by summing all 10 ARR items, with reverse scores for supportive presence and mutual enjoyment items.

The Total ARR risk score has shown adequate internal consistency with a Cronbach’s alpha of $\alpha=.74$ (Sleed, Isosävi et al., 2021), which was also obtained for the current study sample ($\alpha=.77$). Sleed and colleagues (2021) also found good criterion validity for the Total ARR risk score, as shown by its ability to discriminate between normative,

clinical and prison samples of mothers, as well as construct validity, wherein it was significantly associated with poorer reflective functioning, poorer parent-infant interactions, and higher maternal psychopathology.

Parental Stress Index Standard Form - Fourth Edition (PSI-4; Abidin, 1995)

The PSI-4 is a 120 item self-report measure that assesses the level of stress experienced in parenting by considering a parent's relationship with one of their children between the ages of one month and 12 years (see Appendix G). The PSI provides a Total Stress score based on the sum of a child domain which comprises six subscales (distractibility, adaptability, reinforces parent, demandingness, mood, acceptability) and a parent domain with seven subscales (competence, isolation, attachment, health, role restriction, depression, spouse parenting partner relationship). Ninety-one of the items on the Total Stress scale are in Likert-type format ('strongly agree' to 'strongly disagree'); the remaining ten items are response specific, with five options to choose from. In the current sample, the Total Stress scale showed a very good internal consistency with a Cronbach's alpha of $\alpha=.92$.

PSI-4 has displayed very good internal consistency and test-retest reliabilities (Pereira et al., 2012; Rios et al., 2022) and has shown significant associations with several theoretically related child and parent social, emotional, and behavioural health outcomes, indicating validity (Abidin, 2012). The PSI-4 has also been found to have predictive power in associating with behaviours, such as parental negativity in interactions (Mills-Koonce et al., 2007).

Child Abuse Potential Inventory (CAPI; Milner, 1986)

To estimate the risk of parents abusing their children the CAPI was used (see Appendix G). The CAPI is a 160-item self-report questionnaire, which specifically assesses the potential of parents neglecting and physically abusing their children, termed Child Abuse Potential (CAP). It does this by screening characteristics and attitudes which have been theoretically and empirically associated with parental physical child abuse (Milner, 1984; 1986). Items are answered in a dichotomous “agree” or “disagree” format and have a year four reading level. The measure yields a primary clinical scale, the 77-item Abuse scale, that provides the CAP estimate used in the current study. The Abuse scale has six empirically supported factor subscales: distress, unhappiness, problems with self, problems with family and problems from others. An example item that relates to the parental rigidity subscale includes: “a child needs very strict rules”, whilst one that taps into interpersonal difficulties with others is: “other people have made my life hard”. Abuse scale scores can range from 0 to 486; scores above 215 suggest that respondents have personal characteristics similar to parents who are known to be active child abusers, thus indicating an elevated risk profile.

Additionally, the CAPI contains three validity scales: lie, random response and inconsistency. If any of the validity scales are elevated, response distortion indexes (faking good, faking bad, and random response) are computed to determine if the possible profile is invalid. Generally, invalid profiles should not be used, though Milner (1986) suggests that for respondents with a 12th grade education or less, a higher threshold should be applied given the robust relationship between low education and high lie scale scores. Second, if both the abuse scale and lie scale score are elevated, then the former may still be used, based on the assumption that the CAP score may have been even higher had the respondent not attempted to answer items in a socially desirable manner. In the present study, both validity recommendations were applied. The data from CAPI was scored by PariConnect, which generated the CAPI Abuse scale, subscales, and validity scale scores.

A critical review by Walker and Davies (2010) indicated that the CAPI-Abuse scale has good internal consistency, and test-retest reliability estimates across a variety of parent samples along with consistent factorial structures. The systematic review included in the first part of this thesis obtained evidence in support of CAPI Abuse scale construct validity, as shown by its link with a range of theoretically related constructs (e.g., observed negative parenting behaviour, parental empathy, and attitude; Haskett et al., 1995; Kilpatrick, 2005). The systematic review also identified evidence supporting its criterion validity, as shown by its ability to discriminate between maltreating and normative parent samples (e.g., Haapsalo & Altonen, 1999). The CAPI Abuse scale in the current sample showed a good internal consistency with a Cronbach's alpha ($\alpha=.897$).

Adverse Childhood Experiences Questionnaire (ACEs; Felitti et al., 1998)

To estimate parents' childhood trauma the ACEs questionnaire was used (see Appendix G). This is a self-report tool to retrospectively evaluate numerous childhood adversities. It comprises 10 items regarding abuse (emotional, physical, and sexual), neglect (emotional and physical), separation of a parent, violence against the mother, as well as problems of a household member (substance abuse, mental disorder, and prison stay). Each item is answered with either yes (1) or no (0), resulting in a sum score between 0 and 10. The ACE has shown acceptable reliability (Anda et al., 2010), with a good internal consistency observed in the present sample (Cronbach's $\alpha=.820$).

Several studies have examined the psychometric properties of the ACEs scale. The items on this scale have shown good internal consistency (Anda et al., 2010). In line with theory, higher levels of ACEs have been found to be associated with higher levels of perceived stress and mental health problems measured concurrently (Anda et al. 2004) and prospectively (Schilling et al., 2007).

Procedure

As detailed in the SPP protocol (Sleed, Fearon et al., 2021), upon being referred to the SPP by social workers, eligible participants were contacted by the author of the present study and SPP officers by phone to arrange a pre-treatment assessment where the information sheet and consent form were administered (see Appendix H). The former contained information about the research aims and the confidentiality procedures. During this call, participants were informed that they could withdraw at any time, and without any consequences or changes to the children's social care they were currently receiving. Participants were also informed that the questionnaires and interviews contained content that they could find distressing, such as the ACE items or thinking about times that they had been separated from their child for the PDI-S interview.

After recording informed consent, baseline data were then collected remotely via an MS Teams video call. Following the completion of the demographics form, ETMCQ, and PDI-S interview, participants were then given the option to complete the remaining measures alone via an emailed link. Finally, participants could briefly discuss any questions or feedback they had concerning the study, following which an online shopping voucher was emailed or posted for their participation.

Analytic plan

All statistical analyses were performed on baseline assessment data using the Statistical Package for the Social Sciences (SPSS; version 25). By convention, the alpha level was set to 0.05. Data were screened for accuracy of data entry and any missing values. To assess internal consistency and reliability Cronbach's alpha was calculated for each measure. Descriptive statistics were used to summarise the characteristics of the total study sample, with means and standard deviations calculated for continuous variables and proportions for categorical variables.

Pearson's bivariate correlations were used for the first aim of the study, which was to explore the relationship between child abuse potential (CAP) and epistemic trust, mistrust, credulity, parental representations, and parental stress. Correlation coefficients and p -values were reported to attest the strength and significance of the relationship, respectively.

The second aim was to evaluate whether parents' epistemic stance and representational risk would significantly contribute to a model of CAP when accounting for other well established salient factors, namely parental stress, and parents' ACEs. For this aim, a set of linear hierarchical regressions were run in which CAP was entered as the outcome variable and parents' epistemic stance and representational risk were entered in a second and third step, respectively, after controlling for parental stress. Subsequent hierarchical models were run inverting the order of representational risk and epistemic stance to see which of these variables would explain incremental variance in CAP over and above the other. A second set of hierarchical models were run, but this time also controlling for ACEs (see Appendix J). Conducting the analysis in this way also enabled multicollinearity between regression variables to be tested. Indeed, it was deemed that this analysis would help to further knowledge about which of the variables (epistemic mistrust and credulity or representational risk) are most important by establishing that they are, first, distinct. If a variable is just a composite of other variables, then provided that the measure is valid, the construct would not result as incrementally predictive.

Prior to conducting the regression analyses for the main hypothesis testing, exploratory associations were tested between the Total ACE score and its subscales (see Appendix J) with CAP. Consequently, the risk of type I error increased, and should be considered in the interpretation of the results. These exploratory analyses were used to inform model building but were not considered in the interpretation of the findings. Though decreasing the p -value is a suggested strategy to decrease the risk of type I errors (e.g., Rose et al., 2019), this was not adopted here on the basis that, first, the adjusted p -value is defined variably, and second that lowering the p -value increases risk of incurring in type II

errors (Feise, 2002). Instead, reporting the magnitude and direction of effect sizes, was deemed a reasonable measure for balancing the risk of type I and II errors. Key assumptions were tested before conducting correlation and hierarchical multiple regression analyses following guidance for application to the behavioural sciences (Cohen et al., 2003). Assumptions tested included checks regarding multicollinearity between regression variables. As no assumptions were violated (see Appendix I), it was deemed appropriate to continue with the planned analyses.

Service user involvement

The primary study, conducted as part of the SPP, involved Experts by Experience (EbE) in several ways. First, consultations were carried out with EbE to prepare the study materials (i.e., information sheets and consent forms). Second, at the stage of writing the proposal, focus groups were run with graduates from the LPP. Last, EbE were invited to pilot interviews in which the study protocol was tested, and the feedback was used to design the study.

The results of the current study were also presented to an appropriate service-user group, to gather feedback that could help inform the interpretation of the results. The service-user group was formed by two Parent Champion volunteers at the Anna Freud Centre. Two one-hour focus groups were held in which, first, the findings were discussed and interpreted, and their feedback on what the results could mean was gathered. Second, dissemination audiences were discussed, along with how to best produce infographics in which the findings could be sensitively presented and shared with other stakeholders (e.g., high-risk parents, social workers, educational representatives, and service providers).

Results

Descriptive statistics

Table 2 presents the study sample descriptive statistics for all baseline scores used in this study. The N (total number of responses) changed per assessment type due to varied rates of missing data. Furthermore, the table includes cut off scores for the CAPI.

Table 2 shows that the parents were experiencing difficulties in some, though not all, the domains being assessed. The CAPI Abuse scale scores were highly elevated in this sample, with 60% of participants scoring above the cut off. The scores on each of the six factors were elevated for about half of the sample. Distortion indices from the three validity scales indicate that 25 participants were potentially invalid due to elevated lie, random response, and inconsistency scale scores. Distortion indices were therefore calculated revealing potential 'faking good' ($n=22$), 'faking bad' ($n=1$), and 'random responders' ($n=2$). Given the low education background and the fact that the validity scales were not particularly elevated, following CAPI manual recommendations (Milner, 1986) these participants were not removed from the dataset.

Table 2

Descriptive statistics for the sample

Scale	Total Mean (SD)	Range	Cut-off ^a	N (%) above cut-off
CAPI	N=107			
Abuse	244 (98)	22-429	215	64 (60%)
Distress	155 (74)	0-261	152	58 (54%)
Rigidity	15 (12)	0-55	30	15 (14%)
Unhappiness	28 (17)	0-69	23	58 (54%)
Problems with child and self	10 (8)	0-30	11	50 (47%)
Problems with family	18 (14)	0-38	18	54 (50%)
Problems with others	19 (6)	0-24	20	70 (65%)
CAPI (Valid only ^b) N=79	N=79			
	278 (83)	53-376	215	62 (79%)

PSI (T Score)	N=97	
Total stress score	60 (10.54)	35-95
Child domain	73.2 (23.4)	36-92
Parent domain	74.6 (21.8)	35-90
ETMCQ	N=107	
Trust	26 (5.64)	5-35
Mistrust	24 (4.98)	9-34
Credulity	21 (6.37)	7-35
ARR	N=101	
Total ARR score	26.77 (6.7)	13-43
ACEs	N=71	
ACE Total score	4.45 (2.92)	0-9

Note. ^a Cut-off scores as seen in CAPI manual (Milner, 1986), above which participant scores are considered clinically meaningful. ^b Includes only participants with valid CAPI profiles as indicated by validity scales and distortion indices.

Data screening and missing data

Overall, missing value analysis revealed that 6.1% of the data were missing. Conducting complete case analysis would have yielded a loss of 18 (16.36%) participants, due to missing values on one or more items. A missing values analysis indicated that Little's (1988) test for Missing Completely at Random (MCAR) was not significant, $\chi^2 = 61.182$, $df=67$, $p=.677$. These results suggested that the missing data had a random pattern, such that missing data on one variable could not be explained by any other variable in the data set (Tabernach & Fidell, 2001). As such a multiple imputation was run using 10 imputation frames. The pooled value was computed for the variables that were missing data and statistical analyses were then run on each imputed data set. In line with recommendations, the estimates of interest from all imputed data sets were combined into a single estimate by using the average of the estimates obtained from each of the imputed data sets (Asch et al., 2015; Li et al., 2015).

The distributions for each variable were then checked for normality. Two variables were found to be potentially skewed: the epistemic trust score ($z=-3.71$) and the total parental stress score ($z=3.09$). However, as suggested by Piovesana and Senior (2016), it

was reasoned that the relatively large sample size would generate stable means and standard deviations regardless of the skewness present, which here was small. As such, transformations were not performed. All variables were assessed for the presence of outliers using boxplots as described in Field (2009). Two potential univariate outliers were identified for the epistemic trust and parental stress variables. Though these two cases were found to lie just beyond the recommended three standard deviations from the mean, they were retained in the analysis on the basis that they were assumed to represent true participant responses, thus belonging to the sample.

Associations between Child Abuse Potential and Parental stress, Representational risk, and Epistemic trust, mistrust, and credulity

Bivariate Pearson's correlations were used to test the first three hypotheses that epistemic trust would negatively correlate with Child Abuse Potential (CAP) scores, whilst epistemic mistrust, credulity, risky parental representations, and parental stress would positively correlate with CAP. For all correlations, the effect size was interpreted as follows: r values equal or above .1 is small, .3 is medium and .5 is large (Cohen, 1988).

Parental stress was most highly correlated with CAP, showing a significant large, positive relationship between parental stress and CAP ($r(105) = .60, p < .001$), such that parents with higher levels of stress related to the parenting role, tended to have higher abuse potential. Parental risk representations were also significantly positively correlated with CAP ($r(105) = .39, p < .001$), whereby parents with higher levels of risk in their parental representations tended to be at higher risk of abusing their child. A moderate positive correlation was also obtained between CAP and epistemic mistrust and credulity, respectively ($r(105) = .40, p < .001$; $r(105) = .30, p = .002$); however, contrary to predictions, epistemic trust was not significantly correlated with CAP, showing a small negative association ($r(105) = -.18, p = .183$).

Predictors of Child Abuse Potential

Exploratory associations between CAP and the ACE Total score and ACE subscale scores were tested (Appendix J). Among these, only the scores on the ACE childhood sexual abuse (CSA) item were significantly associated with CAP, such that parents' self-reported CSA was related to higher CAP ($r(69) = .26, p = .026$). Thus, CSA was entered into the regression model as a potential confounding variable.

To test hypotheses four and five, exploratory linear hierarchical regressions were computed; CAP was entered as the outcome variable and epistemic mistrust and credulity, and representational risk were entered in block two and three, respectively, to assess whether they would each incrementally improve the prediction of CAP after controlling for parental stress and CSA. Separate models were run in which each of the three predictor variables (i.e., mistrust, credulity, and representational risk) were entered as the target variable in step three to assess which among these was the most important in explaining incremental variance in CAP. As epistemic trust was not significantly correlated to CAP, it was not included in these multivariate analyses.

First, hierarchical analyses that included only parental stress as the covariate were performed. Across all three models, predictor variables at step two made small, significant contributions to CAP variance. As shown in Table 3, only representational risk, however, approached a trend toward significance in predicting incremental variance in CAP whilst controlling for all other variables ($F(1,102) = 3.31, p = .072$); representational risk accounted for almost 2% of the variance in CAP over and above epistemic mistrust and credulity (5.7%), and parental stress (38%). Across all three models, in the final step the only significant predictor was parental stress, which was positively associated with CAP. However as shown in Table 3, in model one, which had representational risk as the target variable, it appears that all three predictors showed a trend towards significance: epistemic mistrust ($\beta = .153, p = .082$), epistemic credulity ($\beta = .143, p = .078$), and representational risk ($\beta = .148, p = .072$).

Table 3*Hierarchical regression analysis to predict CAP*

<i>Variables and steps in the equation</i>	β	ΔR^2
<i>Model 1</i>		
Block 1: Covariate		.378***
Parental stress	.615***	
Block 2: Predictor variables		.057**
Epistemic mistrust	.153 [†]	
Epistemic credulity	.143 [†]	
Block 3: Target variable		.018 [†]
Representational risk	.148 [†]	
<i>Model 2</i>		
Block 2: Predictor variables		.061**
Epistemic mistrust	.194	
Representational risk	.159	
Block 3: Target variable		.014
Epistemic credulity	.131	
<i>Model 3</i>		
Block 2: Predictor variables		.060**
Epistemic credulity	.183*	
Representational risk	.159 [†]	
Block 3: Target variable		.015
Epistemic mistrust	.141	

Note. * $p < .05$; ** $p < .01$; *** $p < .001$; [†] $p < .08$; NB. Block 1 was the same across models 1, 2 and 3

When including CSA as a covariate, once again across all three models, predictor variables that were entered in all possible combinations at step 2, accounted for significant increases in variance over and above parental stress and CSA (see Appendix J): epistemic mistrust and credulity (5.6%), epistemic mistrust and representational risk (5.9%), epistemic credulity and representational risk (5.9%). However, this time, none of the predictor variables individually accounted for incremental increases in CAP variance at step three. In the final step, parental stress was the only variable to significantly predict CAP ($\beta = .601$, $p < .001$).

Discussion

This study aimed to explore the relationship between child abuse potential (CAP) and parents' epistemic trust domains, high-risk parental representations (a measure of low parental mentalizing), and parental stress. It also aimed to determine which of these variables would be most important when it comes to predicting CAP among high-risk parents under Children's social care.

The findings support the hypothesis that parental representations, epistemic mistrust and credulity and parental stress are related to CAP, whilst contrary to predictions, epistemic trust was not significantly associated with CAP. Multivariate analyses suggest that parental stress plays a major role in predicting CAP; although epistemic mistrust, credulity and representational risk, when combined in all possible sets of pairs, made a significant contribution in explaining CAP over and above parental stress and parents' self-reported childhood sexual abuse (CSA), only representational risk approached a non-significant trend in making further additional contributions to CAP above and beyond epistemic mistrust and credulity when parental stress was controlled for. Across all models, when all variables were analysed together, only parental stress significantly predicted CAP, though in one analysis mistrust, credulity and representational risk showed non-significant trends towards predicting CAP alongside parental stress. When CSA was accounted for, parental stress was the only variable to significantly predict CAP when all variables were entered together.

Relationships between Child Abuse Potential (CAP) and Parental stress, Representational risk, and Epistemic trust, mistrust, and credulity

Parental stress

The finding that parental stress and CAP were strongly related is not surprising given that Milner (1986) asserts that child abuse is influenced by numerous factors, which include parents' psychological distress. This finding is consistent with previous research carried out using the CAPI among Child Protective Service (CPS) involved parents (Budd et al., 2006; Holden et al., 1989). Relative to these studies, the present investigation obtained an even larger effect size. One possible explanation is that Budd et al. (2006) measured parental stress at time two, by which point adolescent mothers living in foster care accommodations had received substantial support, which may have attenuated the relationship between their parental stress levels and abuse risk. Furthermore, in Holden et al. (1989) not all participants were child abusers known to CPS, thus in the current study, the larger association may be because participants' daily parenting responsibilities were combined with stress arising from being under the scrutiny of CPS agencies. Additionally, as pointed out by Experts by Experience (EbE) who informed the interpretation of the present findings, data were collected during the Covid-19 Pandemic. This constituted unprecedented times for many families and children, who underwent sudden and enduring changes to their daily routine owing to work and school closures as well as uncertainty regarding the future. Conceivably, this context may have contributed to exacerbating parental stress within this parent population and thus child abuse risk; this explanation is supported by a growing body of research suggesting that the pandemic was associated with increases in parental stress (Adams et al., 2021; Bjørknes et al., 2022).

According to the developmental pathology framework proposed by Asen and Fonagy (2017), it is possible that the large effect size seen here between parental stress and risk of child abuse, may speak to the mentalizing difficulties observed in this parent population,

which may therefore impact their ability to sensitively respond to their child (Fishburn et al., 2017). Such difficulties, in turn, may result in behavioural difficulties in their children, further increasing stress levels in the parental system (Abidin, 2012). Overall, these findings suggest that parental stress is a salient factor when considering child abuse risk among CPS involved parents with their own maltreatment histories.

Parental representations

The significant relationship between representational risk and CAP provides tentative evidence in support of the mentalizing framework (Asen & Fonagy, 2017), which asserts that parents who are at higher risk of maltreating their child are more likely to have mentalizing deficits or impairments owing to their own childhood histories of abuse. Furthermore, these results extend previous research indicating that CPS-involved parents score lower on mentalizing (Fishburn et al., 2017). Though Fishburn and colleagues (2017) used a different operationalisation of mentalization called ‘mind-mindedness’, which examines parents’ ability to be attuned to their infants’ mental states, the fact that in the present study parental risk representations were shown to be associated with child abuse risk, provides further evidence that mentalizing and its various operationalisations may be relevant when considering child abuse risk in CPS-involved populations. This idea is also supported by a relatively recent study by Terry et al. (2020), which found that helpless-hostile representations in caregivers during pregnancy, as measured by the hostile-helpless coding system (Terry et al., 2020) applied to the Pregnancy Interview (Slade, 2011), predicted child removal two years later. The current study extends these findings by suggesting that parental representations involving other types of distortions such as idealizing, incoherence and fearfulness, and among a wider age range of children that goes beyond infancy, relate to child abuse risk. Overall, the relationship between parental representations and child maltreatment risk is, therefore, worthy of further investigation among high-risk parent populations.

Epistemic trust, mistrust, and credulity

It was predicted that higher levels of epistemic trust would allow parents to be open to relevant knowledge about their child thus supporting them to accurately represent and mentalize their child; this in turn, would lead to appropriate parental responses, reflected in lower levels of CAP (Asen & Fonagy, 2017). Contrary to these predictions, though epistemic trust was found to have a negative correlation to CAP, this was small and not statistically significant.

One possible explanation for this result rests on the fact that the smallest effect sizes the study sample could detect was medium. Therefore, this study was underpowered to detect small effect sizes, which are preferable when using relatively new outcome measures (Schafer & Schwarz, 2019), such is the ETMCQ. Thus, the non-significant effect may indicate an absence of a relationship between epistemic trust and CAP, or it may indicate that the strength of the relationship is smaller than the study was powered to detect.

Alternatively, it is possible that epistemic trust may act as a default mode of social functioning and may therefore represent somewhat more of a 'neutral' value. Indeed, Campbell and colleagues (2021) found that epistemic trust could not be regarded as a resilience factor as it did not act as a moderator in buffering the effects of childhood traumatic experiences on adult mental health symptoms. The authors therefore concluded that being open to newly acquired information over and above average may not yield additional resilience against psychopathology among individuals with childhood histories of abuse. This argument could be extended to the current context, such that having higher levels of epistemic trust may not ameliorate current child abuse risk among parents with childhood histories of abuse. Formal mediation and moderation analyses using a longitudinal study design are necessary however to further explore this hypothesis given that both the current study and the one by Campbell et al. (2021) are cross-sectional. Notwithstanding, these preliminary findings suggest that parents with higher levels of epistemic trust are not at less risk of abusing their child.

The finding that epistemic mistrust and credulity were positively associated with CAP lends tentative support to the theory that parents whose epistemic trust has been eroded or never fully developed due to attachment trauma and adverse experiences, may be at higher risk of perpetrating abuse (Sleed, Fearon et al., 2021). According to this theory, this link may arise, first, because vigilance towards their child's communications and intentions could more easily lead to unreflective and hostile assumptions about the child's behaviour which are associated with abuse (Richey et al., 2016). Second, by not being able to update their social knowledge, such parents may form part of that 'hard-to-reach' parent population, which is resistant to change. Consequently, such caregivers would unlikely benefit from therapists or other parents in parent programs or social settings for instance, thus incurring in higher risks of child abuse (Byrne et al., 2019).

The relationship between credulity and CAP may be understood by the fact that parents with trauma backgrounds may at times have excessive openness owing to strong wishes to be able to rely on a trusted other (Campbell et al., 2021). In the context of the parent-child relationship, excessive credulity could, for example, be associated with children crossing boundaries and ineffective management of child behaviour. The child's dysregulated behaviour may exacerbate parental stress (Abidin, 2012) and increase parents' epistemic vigilance and apprehension of their child's intentions. Epistemic vigilance could underpin distorted representations of their child, which could more likely result in instances of physical or emotional abuse (Sleed, Fearon et al., 2021). Overall, these findings provide sufficient preliminary evidence to suggest that epistemic mistrust and credulity may be relevant in the context of child abuse risk in CPS-involved caregivers, thus warranting further explorations.

Predictors of CAP

When considering the type of predictors that would be most important in explaining CAP, hierarchical regressions indicated that the most reliable predictor of CAP, which also

explained the largest amount of variance in this caregiving population, is parental stress. Though, undoubtedly, parental stress is relevant in the context of CAP, it's important to note that the content overlap between the PSI-4 and CAPI Abuse scale, particularly with reference to the distress items in the CAPI Abuse scale, may have artificially inflated the relationship obtained here, thus as argued elsewhere, leading to a potential overestimation of the variance explained (Miragoli et al., 2018).

The strong predictive power of parental stress, over and above the other variables tested here, may also speak to a potential mechanism in which representational risk and epistemic mistrust and credulity domains link to CAP via their effect on parental stress. For instance, as argued by Fonagy et al. (2017), being vigilant to socially communicated knowledge hinders the ability to learn adaptive coping tools throughout life which may, in turn, exacerbate stress within the parent-child system. Further, having parental representations in which the child is viewed through a distorted lens could, for instance, increase the sense that the child is being 'demanding', as seen in high parental stress scenarios, which could then aggravate risk of child abuse through lapses in mentalizing (Nolte et al., 2013) and insufficient emotion regulation in parents (Asen & Fonagy, 2017; Sled, Fearon et al., 2021). Future research would benefit from formally testing this hypothesis through mediational models and longitudinal designs, which would allow these putative mechanisms to be formally assessed.

Epistemic trust and mistrust incrementally explained variance in CAP above parental stress and CSA, suggesting that these constructs do not explain variance in CAP solely through their covariance with parental stress; however, epistemic mistrust and credulity, individually, did not incrementally explain CAP over and above representational risk or each other, respectively. One already mentioned reason may be linked to insufficient power, such that that mistrust and credulity may contribute small increments in explaining CAP variance, however the study could not detect these adequately. It is unlikely that these results were due to unreliable measurement given the very good internal consistency estimates were

obtained in the current sample. The ETMCQ, however, has not been validated in a high-risk or maltreating sample of parents, which may have thus introduced measurement biases that cannot be accounted for here. Future examination of its psychometric properties among maltreating caregiver samples is thus recommended.

Alternatively, the lack of additional variance explained in CAP over and above the other factors may be due to mistrust and credulity being analysed separately as it was assumed that they represent independent factors of the epistemic construct. However, these factors may share a high level of covariance when it comes to CAP and may thus not be mutually exclusive in this context. Accordingly, parents may present with both outlooks on communicated knowledge; a parent may regularly adopt a credulous stance owing to a strong wish and need to trust information from others. This however may then make parents vulnerable to exploitation, in turn exacerbating their epistemic vigilance and vice versa (Campbell et al., 2021).

A theoretical explanation for the lack of additional variance in CAP provided by mistrust and credulity, respectively, relative to representational risk consists in the fact that they may relate to CAP based on their putative effects on mentalizing (Fonagy & Allison, 2014; Sled, Fearon et al., 2021). Parents who are epistemically vigilant are not able to take in new information; suspicious of what is being shared with them, they may be more likely to distort or use pre-mentalizing modes to understand their child and therefore jump to hostile or inaccurate conclusions about what they are seeing (Asen & Fonagy, 2017). Building on the current study findings, future research would be required however to fully assess the potential role of mistrust and credulity. Currently, however, the present findings suggest that epistemic credulity and mistrust may not be the most salient predictors of CAP relative to parental stress and possibly representational risk, among high-risk parents with childhood histories of maltreatment.

The fact that representational risk did not reach full significance in the incremental variance it explained over and above the other predictor variables tested here, may once

more be related to a power issue of the study. Alternatively, it could be related to the measurement properties of the ARR. Currently, inter-rater reliability estimates are not available for this variable as the double coding is yet to place in the primary SPP study. Notwithstanding, the internal consistency for this measure indicated that it was adequate, tentatively supporting its reliability. Furthermore, Slead, Isosävi et al. (2021) provided robust evidence for its construct and criterion validity in high-risk parent populations, which would support the validity of the measure in the current study sample.

The preliminary findings concerning representational risk suggest that first, the variance it explains in CAP is not owing to its overlap with parental stress, and second that it shows a trend approaching significance in contributing additional variance in CAP relative to Mistrust and Credulity. Though, non-significant, it is possible that representational risk may therefore be one of the more important variables when considering risk of child abuse. Its potential predictive power in this context is in line with previous study findings mentioned above (i.e., Terry et al., 2020), which showed that hostile-helpless maternal representations during pregnancy predicted subsequent child removal within two years from birth.

When parents' ACEs, specifically CSA were accounted for, mistrust and credulity and parental representations made a significant additional contribution in explaining CAP when entered in all possible combinations, however in the final step no variable individually made significant contributions in explaining CAP. This may be due to ACE data being collected at follow-up when attrition had occurred. In this way, further considerable reductions in power may have prevented significant small increments in variance from being detected when CSA was accounted for. The fact that CSA's variance in CAP was no longer significant when entered with parental stress, tentatively suggests that the relationship between CSA and CAP may be explained by the effects CSA has on parental stress. This is supported by a previous study which showed that parents' higher ACE scores contributed significant increments in explained variance in parental distress as measured by the Parenting Stress Index – Short Form, even after controlling for socio-economic status (Steele et al., 2016).

Strengths and limitations

One of the strengths of the present study consists in participants belonging to an at-risk clinical sample, who completed assessment measures in the pursuit of therapy. This participant demographic thus represented real world settings which bolster the external validity of the findings. Second, representational risk was measured using independent coders who were blind to participant data, therefore reducing risk of researcher bias. The use of a coding system may have also reduced the risk of social desirability bias in participants. Furthermore, the ARR coding system includes domains such as 'incoherence' and 'idealization' which are helpful for validly measuring representational risk among interviewees who wish to present well or are defensive. Similarly, CAPI contains the already mentioned validity scales to identify inaccurate responding, thus increasing validity of the findings; it also has a year four reading level and simple dichotomous format, which partly control for confounding effects related to differing education and intellectual abilities.

Notwithstanding, the conclusions of this study are limited on several accounts. First, the SPP study inclusion criteria limit the generalizability of findings to lower-income, mainly single-parent households with active and substantiated cases of child abuse. Furthermore, the current sample was predominantly of white ethnicity, so many ethnic and cultural factors remain unaccounted for. This study may also present with selection bias given the considerable loss to follow-up at baseline (42%). These potential participants had not yet consented to participate; thus, it was not possible to assess the extent of selection bias on any of the included variables and examine ways in which the sample was biased.

Furthermore, though this dataset was selected on the grounds that it would allow examination of the research question and testing of hypotheses, the secondary design introduced a number of constraints, namely study design, sampling, and measurements, which were not chosen to address the specific question of whether parental epistemic trust domains, parental stress and representations could predict CAP. Importantly, the cross-sectional design of the study precludes causal pathways and directions of interest from

being established. Therefore, though parental stress, for instance, was found to ‘predict’ CAP, given this consisted in baseline data collected during one time point, the predictors cannot speak to subsequent CAP scores over time. Though the causal model tested here has a strong theoretical rationale (Asen & Fonagy, 2017; Byrne, 2020; Fonagy et al., 2017; Sled, Fearon et al., 2021), alternative and reverse causation models of explanation may still be possible and therefore warrant future examination.

Furthermore, the study had a relatively small sample size. For the number of variables that were examined using hierarchical regressions, very large sample sizes would have been required to detect small effect sizes; it is therefore unlikely that this study was sufficiently powered to estimate the effect of all variables present in the multivariate models (Maxwell, 2000). This was exacerbated by using ACE data collected by the SPP at post-treatment when there had been considerable loss to follow-up (34.5%). Consequently, the decreased power precluded a strong examination of how the current study’s predictors may relate to the intergenerational transmission of abuse. It also prevented valid measurement of the other types of ACEs, which therefore went unaccounted for.

A full examination of the cycle of violence was further hindered by using reports on parental trauma exposure (i.e., ACEs) that were based on retrospective accounts and recollections. This introduced possible error in recall and reporting such that not all cases of prior childhood physical, emotional, and sexual abuse were effectively captured by the ACE questionnaire. As such, it was not possible to fully examine the extent to which parents’ childhood histories of victimization related to their current child abuse potential, nor whether credulity and mistrust, and parental representations could account for CAP over and above such histories of abuse, which would have allowed hypotheses concerning their potential role in maltreatment transmission to be tested.

Though a coding system and validity scales were used, the sole use of parent self-report measures may still have introduced participant response biases. This is particularly the case for CPS-involved caregivers taking part in research on child abuse (Tucker et al.,

2017), who either intentionally or unconsciously, may try to present themselves as parents, and their relationship with their child, in a favourable or falsified way.

Crucially, this study examined the relationship of constructs with *child abuse potential*, not actual child abuse. Although it has been well established that CAP is associated with neglect and emotional abuse (Milner, 1986; Walker & Davies, 2010) – future research should consider including other measures of abuse, such as parent-child observations, relevant informant reports, and behavioural analogues (e.g., Rodriguez et al., 2016) to triangulate with parents' evaluations. Thus, some caution is warranted when interpreting these findings due to these limitations, which are contingent on future research.

Future directions

Further research is required to confirm that epistemic trust domains, parental representations and parental stress play a role in the risk of child abuse. A reasonable starting point for a new study to pursue, would be a replication of a study like this, however, using non-memory recall parental trauma reports and multiple scales to measure each of the target constructs to prevent instrument sensitivity from biasing results. One method to overcome recall bias, could be to measure parents' maltreatment history by using official records supplemented by self-reports, as suggested by Jaffee (2017).

To mitigate power issues, a larger sample would be recommended together with a longitudinal study design to better assess the relationship between child abuse risk, and parents' epistemic stance and representational risk. A more highly powered sample would also allow further examination of whether epistemic trust truly acts as a 'neutral value'. Crucially, replications would be encouraged in ethnically more diverse samples to determine if the relationships found here still hold true for other populations, especially when cultural and ethnic attributes are taken into consideration. More diverse samples may also allow

factors such as socio-economic status to be considered, which here could not be controlled for given that there was little variability in parents' status.

These methodological considerations would help test the intergenerational transmission of abuse theory proposed by Fonagy and colleagues (2017) more extensively as it would be possible to test whether parental epistemic trust domains, parental representations and parental stress mediate the relationship between parents' childhood adversity and subsequent risk of child abuse, as hypothesised by the theory. Consistent with this theory, it would also be important for future studies to include measures of additional constructs, namely parental affect regulation, child behaviour adjustment, and parent psychopathology (e.g., borderline personality disorder traits) which are also conceptualised as contributing to the CM cycle of violence (Asen & Fonagy, 2017). This is consistent with the relatively small incremental amounts of variance explained by the predictors tested here, thus indicating the need to examine whether these additional factors may contribute to the unexplained variance in CAP.

Implications

The results of this study, overall, show that parental stress is a salient factor when considering child abuse potential and that parental representations and epistemic trust domains may also be important for parental risk of child abuse. The intent is not to suggest that parents scoring highly on these variables are child abusers, nor that placement decisions should be sought, but to highlight the presence of relationships and help ascertain which of these variables are most important.

Research from this study suggests that, in congruence with previous literature, parental stress is a strong parental risk factor. Arguably, however, most parents with high levels of parental stress do not abuse their children. Parental stress may thus be a helpful marker when assessed in conjunction with other well-established risk factors (e.g., parents' own history of childhood abuse, socio-economic status, and intimate partner violence). In this way, identification of parental stress among other variables could form the initial steps of a multiple stage screening process to identify multiple types of parenting risks.

In this regard, assessing the extent to which parental stress represents a risk to the family would also involve exploring how much it impinges on parents' cognitive resources, and therefore their mentalizing abilities. As already mentioned, the role of a parent, even in normative non-abusive cases, is ordinarily associated with frequent lapses in mentalizing. However, parental stress that incurs in frequent and excessively bizarre and hostile parental representations may signal higher levels of parental risk. The findings in the current study found that representational risk is likely to explain 2% of child abuse risk even above that accounted for by parental stress, which was measured using a robust and large outcome measure. Thus, among the risk factors to screen alongside parental stress, representational risk is also likely to be important in conducting a comprehensive and global parental risk assessment. When drawing from the domains of representational risk studied here as part of the ARR measure (Sleed, Isosävi et al., 2021), it would be worth understanding, for example, whether parents are able to appraise difficult and stressful parenting situations

using a balanced perspective that also takes into account contextual factors or whether difficult situations are subtly (or markedly) blamed on the child owing to undesirable ‘fixed’ traits that the parent attributes to that child (e.g., ‘jealous’, ‘selfish’, ‘moody’, ‘vicious’ etc.). Accordingly, clinical conversations could elicit parental representations by using prompts from the parent developmental interview that explore how the parent views themselves as caregivers, how they view their child and their relationship to that child. This would provide information not only on whether parents incur ‘lapses’ in mentalizing, which can form part of ordinary, stressful, and non-abusive daily parenting, but also on whether there are more enduring and chronic impairments in this parent ability, which would signal more risk. Comprehensively screening for the different domains in the ARR measure, such as the idealized, hostile, enmeshed, helpless, and fearful representations would enrich the clinical risk assessment and provide intervention targets that could be shared with the wider team and networks around the family.

Importantly, the findings on representational risk and epistemic trust domains also encourage further trialling of parenting interventions for high-risk parents, such as the Lighthouse parenting programme (LPP; Byrne, Slead, Fearon et al., 2021), which is already aimed at targeting parental stress, epistemic mistrust and mentalizing capacities in parents.

These parental factors could also be considered at the family level, for example using a psychological formulation to determine which influences on child abuse risk are most relevant for the family. This would help tailor the intervention to meet those individual needs. Additionally, the findings could inform social care approaches opening the prospect of staff training, to assess and address these characteristics in at-risk parents and in clinicians to learn how to formulate high-risk families using these concepts.

The salient relationship between parental stress and child abuse risk may also have wider implications for public funding and policy, supporting social care interventions for high-risk families. For example, at-risk families involved in Children’s Social Care experiencing high levels of stress could be referred for further family funding, respite care, early help

interventions, and home outreach services to manage and prevent further escalations in child abuse risk.

Conclusion

The current study provides an important first step in elucidating the relationship between CAP and epistemic trust domains, parental representations, and parental stress among high-risk parents with adverse childhood histories. All factors were positively related to CAP, except for epistemic trust which did not display significant associations. Further examination indicated that whilst epistemic mistrust, credulity and parental representations accounted for a significant amount of variance in CAP, irrespective of parental stress and parents' report of childhood sexual abuse, only representational risk showed trends approaching significance as a predictor of CAP above and beyond all factors (epistemic mistrust, credulity, and parental stress). Last, when considered together, only parental stress significantly predicted CAP.

Taking the strengths and limitations of study design into account, the findings highlight the salience of parental stress, and the potential relevance of parents' epistemic trust domains and representational risk when screening parental characteristics associated with child abuse potential. Further research is required, however, to fully assess the extent to which epistemic trust domains are related to child abuse risk and the extent to which parental adverse childhood experiences may impact on child abuse risk through its effects on parental epistemic trust, parental representations, and parental stress.

Future research should address some of the limitations discussed, such as having a more diverse sample, increasing the number of measures used per construct, avoiding the use of recall-based measures of parents' history of abuse, and using longitudinal designs to examine potential mechanisms underpinning these constructs more appropriately. Currently, these results may be considered useful in informing parenting programs using a mentalization based approach. They may also inform social care approaches opening the prospect of staff training, to assess and address these characteristics in at-risk parents and in clinicians to formulate families using these concepts.

Integration, Impact and Dissemination

In this section of the thesis, I will discuss how the systematic review and empirical study, despite representing stand-alone projects, mutually inform one another. I also include potential real-world implications of the findings, considering relevant child maltreatment stakeholders. Last, I outline the dissemination strategy for the systematic review and empirical project findings.

Integration

The overall purpose of this thesis was to explore parental characteristics that are theoretically and empirically considered to be related to child abuse risk among high-risk parents who have their own histories of childhood victimization. It argues that parental stress, parental representations, and parents' epistemic stance may be relevant when examining the risk of child maltreatment among this type of parent population. The empirical project is thought to be one of the first of its kind to explore the relationship between the chosen parental risk factors and child abuse potential within a clinical sample of at-risk parents.

The systematic review critically evaluated the evidence base for the psychometric properties, namely construct and criterion validity and responsiveness (or treatment sensitivity), of the Child Abuse Potential Inventory (CAPI) Abuse scale; it specifically focussed on high-risk and maltreating parents known to children's social care as this is a population with high clinical need. Overall, though not without limitations, the review highlighted that there was evidence to show that the measure is validly assessing child abuse potential, as evidenced by its relationships with multiple theoretically linked constructs. There was also evidence to suggest that this scale could distinguish between maltreating and normative samples of parents, as shown by significant differences on CAPI Abuse scores between these two groups. Further, the review identified moderate strength evidence in support of CAPI responsiveness to post-treatment change. The evidence

regarding how responsive CAPI is to post-treatment differences between treatment conditions was of lower quality due to inconsistency in these results, thereby requiring further examination of this measurement property.

When reflecting on possible topics for the systematic review, my supervisor and I considered choosing one that would be closely related to, and compliment the empirical project, whilst also making a unique contribution to the knowledge base. We reflected that focussing on the psychometric properties of the CAPI Abuse scale, would fulfil this remit given that it constituted the main outcome measure of the empirical study, and that it could also have additional research and clinical implications in the domains of risk assessment, epidemiological and intervention science.

In terms of which psychometric properties to evaluate, construct validity and treatment sensitivity appeared to be a good starting place as suggested by a relatively recent systematic review, which had found that there was insufficient evidence to speak to these CAPI measurement properties (Georgieva et al., 2022). Further, evaluating CAPI construct validity offered the opportunity not only to support the empirical project findings by clarifying whether, indeed, the CAPI is successful at capturing child abuse potential, but in the process, it also provided an evidence base of studies that used CAPI among high-risk parent populations. These studies, in turn, could then also be used to compare and inform the interpretation of the empirical project findings.

Though the choice to also consider CAPI Abuse scale treatment sensitivity was not strictly related to the outcomes of the empirical project, it was deemed important since the primary study from which the data of the empirical project were drawn, i.e., the Supporting Parents Project (SPP), used the CAPI as one of its primary outcome measures to evaluate the intervention effectiveness of the Lighthouse Parenting Programme (LPP). Additionally, the treatment sensitivity findings were considered to be helpful to future intervention scientists in the field of child maltreatment, especially considering the low availability of tertiary prevention programs for the most high-risk families in the UK (Mulcahy et al., 2014;

Alink, 2020). As a result, both measurement properties were considered important enough to proceed with the systematic review.

Finally, I chose to focus on evaluating CAPI's psychometric properties exclusively in a high-risk parent population as this was the sample of the empirical project, and it was thus deemed important to maximise study comparability, as well as understand how CAPI specifically performed in this parent population. The participant inclusion criteria of the systematic review therefore emulated those of the primary study from which the empirical project data were drawn. Though it would have been important and interesting to also evaluate the CAPI's measurement properties in a community sample, and compare these to high-risk ones, this would have not been feasible and was beyond the scope of the thesis. The empirical project, in turn, to some extent could be deemed as expanding on the findings from the systematic review, however, it did not aim to validate the CAPI. Nonetheless, future reviews could include it as indirect evidence of the CAPI Abuse scale construct validity.

In terms of how the systematic review helped to support findings from the empirical project, first it confirmed that no other study had considered the relationship between the constructs used in the current empirical study and the CAPI, thus asserting the unique contribution that the current empirical investigation has to the CAPI literature. Second, the review provided helpful background research. For instance, it supported the notion of an intergenerational transmission pattern even when measuring the risk of maltreatment using the CAPI. Indeed, some studies showed that parents with histories of childhood maltreatment had higher CAPI Abuse scores relative to those without (Holden et al., 1989), and were more likely to reach CAPI clinical cut off than those who did not have such histories (Craig & Sprang, 2007). The findings from the review papers also helped to consider how high the CAPI scores were in the current sample, relative to other high-risk parent samples examined. This informed my decision to use a more conservative CAPI cut-off score of 215, as this was the most common practice among the review studies. The fact that much of the sample exceeded this conservative cut off score, was then helpful in

confirming the extent of child abuse risk in the sample. Finally, and most importantly, the moderately strong evidence in support of the sufficient construct and criterion validity of the CAPI Abuse scale confirmed that it likely comprises an appropriate measure of child abuse risk, and therefore provided support that the current study findings likely speak to this construct.

Research challenges and their implication for the project

For this thesis, I am occupying a positivist philosophical position which is premised on quantitative measurement, data collection and analysis. This philosophical standpoint asserts that as a researcher my identity and experiences should be as far removed from the research process as possible to remain objective (Anderson, 2020). I endorse this philosophy to the extent that I hope any researcher repeating my study, regardless of our differences, would obtain the same results. However, drawing on a critical realist perspective also helped me to critically evaluate the research I conducted and the challenges this posed, by considering how my identity, experience and position as a researcher may have influenced the process.

As the empirical study was based on secondary data analysis, I could not influence the design, measures nor the sample that was chosen and therefore aspects of my identity may not be relevant to consider here. However, as I was involved in developing the research question, parts of the data collection, and the participant interview coding, invariably my position as an 'other white', cis, hetero, female clinical psychologist trainee, without children, may be relevant to reflect on. For instance, it was interesting to manage the risk that came up when working with the parent participants within the bounds of being a researcher, rather than in a clinical capacity. As a part time clinician, I was very tempted to signpost and suggest strategies to participants who were struggling to help them feel better. As someone who has often supported adults who are unwell in my own private life, I felt further called into action. However, this conflicted with the research protocol which required me to simply make

my field supervisor and relevant social worker aware of participants' risk. It is therefore possible that my relationship building with participants may have influenced the data collection process in unique ways owing to these aspects of my identity.

Similarly, when conducting the parent developmental interviews, I often found myself using skills that I had learnt from delivering psychological therapies, such as 'reflecting back', 'summarising answers', and clarifying whether I had understood responses correctly to increase awareness in the interviewee. As a result, the use of these clinical skills may have stimulated the participants' mentalizing beyond what was recommended by the interview protocol, the main purpose of which, was indeed to observe and measure naturally occurring mentalizing in parents. At times, my position as a clinician was therefore a challenge to the research process, as I unwittingly interacted and intervened in the process more than would be advised when simply measuring and observing a phenomenon. Another example consists in my own personal and clinical experiences of intergenerational trauma, which may have influenced my choice to want to include the Adverse Childhood Experiences variable in the final model I tested, thus influencing the direction of the empirical study question.

When rating the parent interviews using the Assessment of Representational Risk (ARR) coding system, I found that my lack of experience in raising children posed a challenge. For instance, when rating the 'hostile experience and behaviour' domains of the ARR, I found it difficult to ascertain how truly 'hostile' the parent's experience of their child was. I noticed (as also pointed out by my field supervisor) that at times I was potentially being quite unforgiving towards the parents and rating this item too harshly. My lack of lived experience on what constitutes normal parenting frustrations, as well as 'good enough' parenting with its 'good enough' repairs, may have thus introduced some bias into my ratings. Understandably, I was grateful to have regular reflective sessions with the research team I had joined on an honorary basis to triangulate my thinking and rational for the ratings, this way also supporting inter-rater reliability. Finally, as a white female, it is possible that I am approaching this topic from a westernised cultural viewpoint of what is considered

sensitive and healthy caregiving. I found it helpful to be able to critically reflect on this aspect of my identity in the team meetings, in which I was helpfully reminded to reflect from a child's perspective rather than drawing on western norms.

Reflections on my interest in the topic area

Child maltreatment is a topic that is important to me both on a professional and personal level. As such, it is a field that I have always been interested in. One of my first psychology jobs consisted in working as an assistant in a specialist trauma team alongside foster carers to support their training and own emotional needs, so that in turn, they could better support the needs of other foster carers and the children they cared for who had experienced severe maltreatment. This professional pursuit was perhaps part of a wish to understand my families own past and present experiences of harm and a way to help me make sense of the resilience that I had found in myself.

Crucially, my personal and clinical experiences of maltreatment have allowed me to witness and appreciate that, for the most part, no parent in their healthy mind wants to harm their child. Speaking to almost any parent, they will invariably and genuinely profess infinite love for their child, and if ever required, would give up their life with the blink of an eye to protect them. Yet, when working with high-risk families, I have been struck by how such undying and unconditional love can coexist with harm.

My work in a homeless hostel with families on the edge of care as a parent-toddler play assistant (Sleed et al., 2013), helped me to appreciate some of the factors that can get in the way of such unconditional love. For the families we were working with, including refugee families, this involved structural and health inequalities, such as stress arising from not having indefinite leave to remain, or living in small studio rooms as a whole family. There is clear evidence that these elements can increase risk of parental harm to the child and had the present empirical project sample been varied enough on these variables (e.g., socio-

economic status), I would have been keen to include their effects in the adjusted analyses. Here, I also learnt that when working within parent-child relationships, you cannot formulate solely based on these two individuals but must draw on the influence and legacy that the generations before them have exerted. Understanding intergenerational experiences of parenting practices helped me to appreciate how the parents who had not received adequate care and protection from their own caregivers, often struggled to stay curious and see the child before them, due to intergenerational patterns that eclipsed their view. At the time, the concept of mentalization was very helpful to me in supporting parents to be curious and understand their child beyond haste judgements such as the child “being naughty”, “possessive”, “rough and inconsiderate to other children” and to support the parents in remaining present and connected with their child in the context of play.

As a result of these collective experiences and maturing these interests, I was therefore grateful to have had the opportunity to further enrich my learning by testing concepts out empirically in the current project and to also learn about the link between epistemic trust and mentalizing within an attachment setting.

Impact

Child maltreatment is a widespread societal issue (WHO, 2022), which unfortunately is associated with devastating, lifelong effects for children around the world (DePasquale et al., 2019). In addition to being morally unacceptable, it also burdens society at a financial level. The estimated lifetime cost of non-fatal child maltreatment by a primary caregiver is approximately £89,000, with the largest contributors to this expense being social and short-term health care-related costs, as well as the costs arising from lower employment probability (Conti et al., 2017). Increasing knowledge of child maltreatment and developing effective interventions for high-risk parents is therefore a priority.

Provided that future research further substantiates the notion that parental representations and parents' epistemic stance are important factors in the context of child abuse risk, the empirical project results and systematic review have the potential to increase understanding of child maltreatment risk and to underpin important adaptations made to interventions for high-risk and maltreating parents.

In terms of clinical practice at an individual and family level, allied health professionals working in a mental health capacity with at-risk or maltreating parents could use knowledge concerning the parental characteristics which here were found to relate to child abuse risk, to inform their practice. First, as part of an assessment, exploring the parents' representations, perceived parental stress and epistemic stance, could help to evaluate child abuse risk more comprehensively; clinicians could refer to domains considered in the ARR measure and use these in a formal or informal capacity to assess the presence of risky representations (e.g., hostile, helpless) in parents. In this vein, clinicians could formulate the parent's epistemic stance to understand whether they are open to communication from other adults/professionals, as well as from their own children or, rather, whether they appear more 'rigid' and suspicious towards information shared. Understanding how strongly these features present in parents could thus inform the formulation of the family, in turn helpfully guiding interventions for such families. On a practical level, detecting

high levels of parental stress could also be a helpful reminder to professionals to sign post parents to relevant services within the NHS, local authority or third sector agencies, which could helpfully support and alleviate such stress.

Clinicians working therapeutically with high-risk families who assess parents as having high levels of epistemic mistrust and low mentalizing abilities of their child, could adopt communication systems that foster more openness in the parent. As suggested by Fonagy and colleagues (2017) this would entail the use of mentalization-based treatment components to help the parent feel validated, understood, and mirrored in ways they mostly likely did not experience from their own families or networks. Accordingly, clinicians would be showing explicit curiosity in the mind of the parent and showing how the parent, in turn, impacts their mind, thoughts and feelings (Byrne et al., 2019). This would also help the parent to become aware of mental states and their ability to impact these. According to this approach, clinicians would also be encouraged to help the parent become aware of what it feels like to be 'held in mind'. This is deemed to support the parent to "internalize more benign models of how minds influence minds" (Byrne, 2020). Importantly, once further robust evidence is obtained this will support future commissioning of intervention programs that, among other known risk factors, could target representational risk, mentalizing and epistemic trust as important prevention ingredients.

In terms of systematic review findings, professionals working in child protection settings, such as clinicians ordered by family courts to draft expert witness statements, could be more confident in the validity of the CAPI to assess risk of abuse. Though further research is required to attest the criterion and predictive validity of the CAPI, there is scope to carry on using it to help discern levels of risk, particularly as a tool which informs rather than determines the outcome. As such, it would be reasonable for professionals to use the CAPI Abuse scale score to complement their clinical risk judgements when assessing and formulating risk of maltreatment.

Once future research ascertains the relationship between child maltreatment risk and parental representations and epistemic stance more clearly, an appropriate next step would be to implement the learning at a service level by delivering appropriate training for practitioners working in health, education, and social care settings. Training could support educational representatives and local authority stakeholders (e.g., social- and early help workers) to become aware of how these parental characteristics relate to risk and how to assess and formulate their presence in families. Further, professionals could be supported and given the necessary supervision to foster mentalizing and reflective capacities of their own. This could support the creation of a wider mentalizing system around the family in which caregivers can develop but are also helped in sustaining their mentalizing capacity regardless of the natural strains of parenting and their ongoing life stressors (Campbell & Allison, 2022).

Crucially, once further due diligence is performed on this research topic, policy and social care guidance manuals should also be updated to further support professionals in safeguarding to conduct comprehensive and good quality risk assessments. For example, statutory guidance such as “Working together to safeguard children” (DfE, 2022), which already includes alerting features of child maltreatment, could be extended to include meaningful practice points that emerge from this line of research.

Dissemination

Meaningful dissemination represents a good strategy for maximising the impact of research findings (Penfield et al., 2014). Appropriate dissemination involves identifying the relevant stakeholders who consume the research so that findings can have a wide-reaching and transformative impact (Wilson & Marshall, 2010). Below I will share the dissemination strategy of the thesis findings, which in part, I have already implemented.

In a first instance, I shared the results with two experts by experience (EbE) who are employed by the Anna Freud Centre as 'Parent Champions'. These are parents and carers who have experience of having cared for a child with mental health difficulties and/or have experienced mental health issues themselves. I considered their involvement in my project to be important not only to the extent that it has been shown to increase self-esteem and well-being in EbE, which ethically as researchers is something we should all strive for, but also because it is known to support professionals to create useful services (Fernandez et al., 2003). Though my findings are still at an early stage in terms of service development, by sharing these findings across the relevant networks, I hope that it will help to mature an interest in the topic and findings, from which service implementation can be considered in the future. In addition, wider broadcasting in this way, will hopefully provide meaningful input to treatment programs which are already drawing on the concepts I have studied here, such as the Lighthouse Parenting Programme (LPP; Slead, Fearon et al., 2021).

After hosting two focus group sessions, the Parent Champions helped me to interpret the findings from the empirical project and advised me on how to best disseminate the results, which in turn, has informed my dissemination strategy. First, I will make an infographic for parents in which I will carefully select out any language that could be interpreted as 'shaming' or 'blaming' by parents; the Parent Champions helpfully suggested to highlight to parents that as professionals we know how much they love and want what is best for their children, even when they may be harming the child. The Parent Champions also suggested to communicate the findings in such a way as to empower the parents to

recognise potential risks in themselves; for example, the info-graphic could highlight some of the signs that parents may notice in themselves, for instance when feeling stressed in their parent role and to inform them of how this could prevent them from getting the best out of their family life; therefore the info-graphic could warmly encourage the parents to reach out to professionals or other supportive people in their lives to help them address some of their parental stress. Accordingly, helpful signposting may be included on the infographic. Thus, the infographic will be a visual starting point to help parents and caregivers think about the parental risk factors studied here and how these may relate to them.

The Parent Champions also suggested to create a video in which all the relevant stakeholders (e.g., parent, child, social care, educational representatives) share the main points of the research, i.e., that there are certain parental risk factors which may be important when thinking about child abuse. I am currently in discussion with my field supervisor and the communications team at the Anna Freud Centre to plan the next steps. The Parent Champions highlighted the importance that the various stakeholders communicate the research findings as part of a 'united front' which aims to support the reduction of child maltreatment by working *alongside*, rather than *against* parents. Last, I have emailed the Parent Champions the power point slide presentation that I had shared during the initial focus group, along with the meeting minutes. Therefore, should they wish to, they can further reflect on the research findings and use the learning to inform future research discussions they participate in at the Anna Freud Centre.

Additionally, my external supervisor and I are planning for me to share the findings in relevant academic and service-related forums forums, such as the social workers who were involved in the recruitment and implementation of the Supporting Parents Project. My supervisor, Dr Slead, will also be presenting the findings at the International Mentalization Based Treatment with Children and Young People Conference this coming June 2023 in Barcelona, Spain.

The design, methodology and results from the empirical study have also been presented to staff and trainee Clinical Psychologists at Royal Holloway, in May 2023 via an interactive presentation. This presentation included allocated time for questions and reflections in which study findings and methodological choices were discussed, as well as the research strengths and limitations. The thesis will also be uploaded onto Royal Holloway University London online repository (Pure), which can be accessed by both staff and students.

Furthermore, the work will also be presented at the specialist CAMHS service where I am currently on placement at one of the monthly CPD meetings and circulated to the multidisciplinary team. The message I will be focussing on is that high levels of stress in the parent child system, distorted parental representations, as well as being less open to socially exchanged knowledge may be present in the context of risk. I will therefore interactively describe these concepts and findings and encourage clinicians to be curious when these factors emerge in their assessments and treatments.

Finally, to make my research available more broadly to an academic audience, I will submit both the systematic review and empirical paper to peer reviewed journals for publication. The Journal of Child Abuse and Neglect, the Child Abuse Review, the Journal of Child Maltreatment and the Journal of Child Psychology and Psychiatry have been identified as potential journals to submit to due to their publication record of research on Child Maltreatment and with studies using the CAPI to measure it (e.g., Miragoli et al., 2018).

References

- Abidin, R. R. (1995). Manual for the parenting stress index. *Odessa, FL: Psychological Assessment Resources.*
- Abidin, R. R. (2012). *Parenting Stress Index* (Fourth). Psychological Assessment Resources.
- Adams, E. L., Smith, D., Caccavale, L. J., & Bean, M. K. (2021). Parents Are Stressed! Patterns of Parent Stress Across COVID-19. *Frontiers in Psychiatry, 12*, 300.
<https://doi.org/10.3389/FPSYT.2021.626456/BIBTEX>
- Alink, L. (2020, January 29). *Unravelling child maltreatment: What are the prevention ingredients?* [Conference presentation]. Parent-Infant Foundation Conference, London, United Kingdom.
https://parentinfantfoundation.org.uk/wp-content/uploads/2020/10/Plenary1_Lenneke-Alink.pdf
- Anda, R. F., Butchart, A., Felitti, V. J., & Brown, D. W. (2010). Building a Framework for Global Surveillance of the Public Health Implications of Adverse Childhood Experiences. *American Journal of Preventive Medicine, 39*(1), 93–98. <https://doi.org/10.1016/j.amepre.2010.03.015>
- Anderson, F., (2020). *The Relationship Between Maternal Experiences of Abuse and Maternal Sensitivity in Early Mother-Infant Interactions* [Doctoral dissertation, UEL (University East London)]
- Asch, D. A., Troxel, A. B., Stewart, W. F., Sequist, T. D., Jones, J. B., Hirsch, A. M. G., Hoffer, K., Zhu, J., Wang, W., Hodlofski, A., Frasch, A. B., Weiner, M. G., Finnerty, D. D., Rosenthal, M. B., Gangemi, K., & Volpp, K. G. (2015). Effect of Financial Incentives to Physicians, Patients, or Both on Lipid Levels. *JAMA, 314*(18), 1926. <https://doi.org/10.1001/JAMA.2015.14850>
- Asen, E., & Fonagy, P. (2017). Mentalizing Family Violence Part 1: Conceptual Framework. *Family Process, 56*(1), 6–21. <https://doi.org/10.1111/FAMP.12261>

- Assink, M., Spruit, A., Schuts, M., Lindauer, R., van der Put, C. E., & Stams, G. J. J. M. (2018). The intergenerational transmission of child maltreatment: A three-level meta-analysis. *Child Abuse and Neglect*, *84*(July), 131–145. <https://doi.org/10.1016/j.chiabu.2018.07.037>
- Baldwin, H., Biehal, N., Allgar, V., Cusworth, L., & Pickett, K. (2020). Antenatal risk factors for child maltreatment: Linkage of data from a birth cohort study to child welfare records. *Child Abuse and Neglect*, *107*. <https://doi.org/10.1016/J.CHIABU.2020.104605>
- Barlow, J., Johnston, I., Kendrick, D., Polnay, L., & Stewart-Brown, S. (2006). Individual and group-based parenting programmes for the treatment of physical child abuse and neglect. *Cochrane Database of Systematic Reviews*. <https://doi.org/10.1002/14651858.cd005463.pub2>
- Bateman, A., & Fonagy, P. (2016). Mentalization-Based Treatment for Personality Disorders. *Mentalization-Based Treatment for Personality Disorders*. <https://doi.org/10.1093/MED:PSYCH/9780199680375.001.0001>
- Bateman, A., O'connell, J., Lorenzini, N., Gardner, T., & Fonagy, P. (2016). *A randomised controlled trial of mentalization-based treatment versus structured clinical management for patients with comorbid borderline personality disorder and antisocial personality disorder*. <https://doi.org/10.1186/s12888-016-1000-9>
- Behl, L. E., Conyngham, H. A., & May, P. F. (2003). Trends in child maltreatment literature. *Child Abuse and Neglect*, *27*(2), 215–229. [https://doi.org/10.1016/S0145-2134\(02\)00535-5](https://doi.org/10.1016/S0145-2134(02)00535-5)
- Belsky, J. (1993). Etiology of child maltreatment: A developmental-ecological analysis. *Psychological Bulletin*, *114*(3), 413–434. <https://doi.org/10.1037/0033-2909.114.3.413>
- Bert, S. C., Guner, B. M., Lanzi, R. G., & Centersfor Preventionof Child Neglect. (2009). The influence of maternal history of abuse on parenting knowledge and behavior. *Family relations*, *58*(2), 176-187.

- Bjørknes, R., Skogen, J. C., Nærde, A., Sandal, G. M., Haug, E., Mæland, S., Fadnes, L. T., & Lehmann, S. (2022). Parental stress during the COVID-19 pandemic: A one-year follow-up. *PloS One*, *17*(12). <https://doi.org/10.1371/JOURNAL.PONE.0276190>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97–111. <https://doi.org/10.1002/JRSM.12>
- Bradshaw, K. M., Donohue, B., Cross, C., Urgelles, J., & Allen, D. N. (2011). Examination of the Relationship Between Parental Satisfaction and Child Maltreatment Potential While Considering Social Desirability. *Journal of Family Violence*, *26*(7), 545–549. <https://doi.org/10.1007/s10896-011-9389-x>
- Browne, K. D., Kevin D., & Herbert, Martin. (1997). *Preventing family violence*. 382. <https://www.wiley.com/en-gb/Preventing+Family+Violence-p-9780471941408>
- Budd, K. S., Heilman, N. E., & Kane, D. (2000). Psychosocial correlates of child abuse potential in multiply disadvantaged adolescent mothers. *Child Abuse and Neglect*, *24*(5), 611–625. [https://doi.org/10.1016/S0145-2134\(00\)00122-8](https://doi.org/10.1016/S0145-2134(00)00122-8)
- Budd, K. S., Holdsworth, M. J. A., & HoganBruen, K. D. (2006). Antecedents and concomitants of parenting stress in adolescent mothers in foster care. *Child Abuse & Neglect*, *30*(5), 557–574. <https://doi.org/10.1016/j.chiabu.2005.11.006>
- Burns, K. A., Chethik, L., Burns, W. J., & Clark, R. (1997). The early relationship of drug abusing mothers and their infants: An assessment at eight to twelve months of age. *Journal of Clinical Psychology*, *53*(3), 279–287. [https://doi.org/10.1002/\(SICI\)1097-4679\(199704\)53:3<279::AID-JCLP11>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1097-4679(199704)53:3<279::AID-JCLP11>3.0.CO;2-F)
- Byrne, G. (2020). " And whatever you say, you say nothing" Establishing epistemic trust in The Lighthouse MBT-parenting programme: a case study. *Journal of Psychological Therapies*, *5*(2), 206-228.

- Byrne, G., Sleded, M., Midgley, N., Fearon, P., Mein, C., Bateman, A., & Fonagy, P. (2019). Lighthouse Parenting Programme: Description and pilot evaluation of mentalization-based treatment to address child maltreatment. *Clinical Child Psychology and Psychiatry*, 24(4), 680–693. <https://doi.org/10.1177/1359104518807741>
- Campbell, C., & Allison, E. (2022). Mentalizing the modern world. *Psychoanalytic Psychotherapy*, 36(3), 206-217.
- Caliso, J. A., & Milner, J. S. (1992). Childhood history of abuse and child abuse screening. *Child Abuse & Neglect*, 16(5), 647–659. [https://doi.org/10.1016/0145-2134\(92\)90103-X](https://doi.org/10.1016/0145-2134(92)90103-X)
- Caliso, J. A., & Milner, J. S. (1994). Childhood Physical Abuse, Childhood Social Support, and Child Abuse Potential. *Journal of Interpersonal Violence*, 9(1), 27–44.
- Campbell, C., Tanzer, M., Saunders, R., Booker, T., Allison, E., Li, E., O'Dowda, C., Luyten, P., & Fonagy, P. (2021). Development and validation of a self-report measure of epistemic trust. *PLoS ONE*, 16(4 April), 1–21. <https://doi.org/10.1371/journal.pone.0250264>
- Carothers Bert, S., Mironovna Guner, B., & Lanzi, R. G. (2009). The Influence of Maternal History of Abuse on Parenting Knowledge and Behavior. *Fam Relat*, 58(2), 176–187. <https://doi.org/10.1111/j.1741-3729.2008.00545.x>
- Centers for Disease Control and Prevention. (2022). *Preventing Child Abuse and Neglect*. National Centre for Injury Prevention and Control, Centres for Disease Control and Prevention. https://www.cdc.gov/violenceprevention/pdf/can/CAN-factsheet_2022.pdf
- Chaffin, M., Hecht, D., Bard, D., Silovsky, J. F., & Beasley, W. H. (2012). A statewide trial of the safecare home-based services model with parents in child protective services. *Pediatrics*, 129(3), 509–515. <https://doi.org/10.1542/peds.2011-1840>
- Chen, M., & Chan, K. L. (2016). Effects of Parenting Programs on Child Maltreatment Prevention: A Meta-Analysis. *Trauma, Violence & Abuse*, 17(1), 88–104. <https://doi.org/10.1177/1524838014566718>

- Chilamkurti, C., & Milner, J. S. (1993). Perceptions and Evaluations of Child Transgressions and Disciplinary Techniques in High- and Low-Risk Mothers and Their Children. *Child Development, 64*(6), 1801. <https://doi.org/10.2307/1131470>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. Lawrence Erlbaum Associates, Publishers.
- Cohen, J., Cohen, P., West S.G., & Aiken, L. S. (2003). Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. In *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences (3rd ed.)*. Routledge. <https://doi.org/10.4324/9780203774441>
- Columb, M. O., Ffcm, F., & Atkinson, M. S. (2016). Statistical analysis: sample size and power estimations. *British Journal of Anaesthesia Education, 16*, 159–161. <https://doi.org/10.1093/bjaed/mkv034>
- Conti, G., Morris, S., & Melnychuk, M. (2017). The economic costs of child maltreatment in the UK: a preliminary study. *London: NSPCC*.
- Craig, C. D., & Sprang, G. (2007). Trauma exposure and child abuse potential: Investigating the cycle of violence. *American Journal of Orthopsychiatry, 77*(2), 296–305. <https://doi.org/10.1037/0002-9432.77.2.296>
- Csibra, G., & Gergely, G. (2006). *Social learning and social cognition: The case for pedagogy*. 249–274.
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148–153. <https://doi.org/10.1016/j.tics.2009.01.005>
- Department for Education. (2022). *Working together to safeguard children*. (DFE-00195-2018) https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/942454/Working_together_to_safeguard_children_inter_agency_guidance.pdf

- DePasquale, C. E., Handley, E. D., & Cicchetti, D. (2019). Investigating multilevel pathways of developmental consequences of maltreatment. *Development and Psychopathology*, 31(4), 1227–1236. <https://doi.org/10.1017/S0954579418000834>
- Dixon, L., Browne, K., & Hamilton-Giachritsis, C. (2005). Risk factors of parents abused as children: a mediational analysis of the intergenerational continuity of child maltreatment (Part I). *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 46(1), 47–57. <https://doi.org/10.1111/J.1469-7610.2004.00339.X>
- Dong, M., Anda, R. F., Felitti, V. J., Dube, S. R., Williamson, D. F., Thompson, T. J., Loo, C. M., & Giles, W. H. (2004). The interrelatedness of multiple forms of childhood abuse, neglect, and household dysfunction. *Child Abuse & Neglect*, 28, 771–784. <https://doi.org/10.1016/j.chiabu.2004.01.008>
- Donohue, B., Azrin, N. H., Bradshaw, K., Van Hasselt, V. B., Cross, C. L., Urgelles, J., ... & Allen, D. N. (2014). A controlled evaluation of family behavior therapy in concurrent child neglect and drug abuse. *Journal of consulting and clinical psychology*, 82(4), 706. <http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/722/CN-00998722/frame.html>
- Donohue, B., Pitts, M., Chow, G. M., Benning, S. D., Soto-Nevarez, A., Plant, C. P., & Allen, D. N. (2016). Development and initial psychometric examination of the Home Safety and Beautification Assessment in mothers referred to treatment by child welfare agents. *Psychological Assessment*, 28(5), 523–538. <https://doi.org/10.1037/pas0000197>
- Donohue, B., Plant, C. P., Chow, G., Schubert, K., Bradshaw, K., Cappellano, J. U., & Allen, D. N. (2019). Contribution of illicit/non-prescribed marijuana and hard-drug use to child-abuse and neglect potential while considering social desirability. *British Journal of Social Work*, 49(1), 77–95. <https://doi.org/10.1093/bjsw/bcy027>
- Donohue, B., Plant, C. P., Loughran, T. A., & Torres, A. (2017). Family assisted contingency management within the context of evidence-supported treatment for child neglect and drug

abuse. *Journal of Child and Family Studies*, 26(8), 2224–2236.

<https://doi.org/10.1007/s10826-017-0739-z>

Fang, Y., Boelens, M., Windhorst, D. A., Raat, H., & van Grieken, A. (2021). Factors associated with parenting self-efficacy: A systematic review. *Journal of Advanced Nursing*, 77(6), 2641–2661. <https://doi.org/10.1111/JAN.14767>

Feise, R. J. (2002). Do multiple outcome measures require p-value adjustment? *BMC Medical Research Methodology*, 2(1), 1–4. <https://doi.org/10.1186/1471-2288-2-8/COMMENTS>

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(4), 245–258.

[https://doi.org/10.1016/s0749-3797\(98\)00017-8](https://doi.org/10.1016/s0749-3797(98)00017-8)

Fernandez, C. V., Kodish, E., & Weijer, C. (2003). Informing study participants of research results: an ethical imperative. *IRB: Ethics & Human Research*, 25(3), 12-19.

Field, A. (2009). *Discovering statistics using SPSS*. Sage Publications.

Finzi-Dottan, R., & Harel, G. (2014). Parents' Potential for Child Abuse: An Intergenerational Perspective. *Journal of Family Violence*, 29(4), 397–408. <https://doi.org/10.1007/S10896-014-9592-7/FIGURES/1>

Fishburn, S., Meins, E., Greenhow, S., Jones, C., Hackett, S., Biehal, N., Baldwin, H., Cusworth, L., & Wade, J. (2017). *Mind-Mindedness in Parents of Looked-After Children*.

<https://doi.org/10.1037/dev0000304>

Fonagy, P., & Allison, E. (2014). The role of mentalizing and epistemic trust in the therapeutic relationship. *Psychotherapy*, 51(3), 372–380. <https://doi.org/10.1037/A0036505>

Fonagy, P., Campbell, C., & Bateman, A. (2017). Mentalizing, attachment, and epistemic trust in group therapy. *International Journal of Group Psychotherapy*, 67(2), 176-201.

- Fonagy, P., Gergely, G., Jurist, E. L., & Target, M. (2002). *Affect regulation, mentalization, and the development of the self*. Other Press.
- Fonagy, P., Gergely, G., Jurist, E. L., & Target, M. (2018). Affect regulation, mentalization and the development of the self. *Affect Regulation, Mentalization and the Development of the Self*, 1–577. <https://doi.org/10.4324/9780429471643>
- Fonagy, P., & Luyten, P. (2016). A Multilevel Perspective on the Development of Borderline Personality Disorder. *Developmental Psychopathology*, 1–67. <https://doi.org/10.1002/9781119125556.DEVPSY317>
- Fonagy, P., Luyten, P., & Allison, E. (2015). Epistemic petrification and the restoration of epistemic trust: A new conceptualization of borderline personality disorder and its psychosocial treatment. *Journal of Personality Disorders*, 29(5), 575–609. <https://doi.org/10.1521/PEDI.2015.29.5.575>
- George, C., & Solomon, J. (2008a). The caregiving behavioral system: A behavioral system approach to parenting. In J. Cassidy & P. Shaver (Eds.), *Handbook of Attachment: Theory, Research, and Clinical Application* (2nd ed., pp. 833–856). New York: Guilford Press.
- Georgieva, S., Tomas, J. M., & Navarro-Pérez, J. J. (2021). Systematic review and critical appraisal of Childhood Trauma Questionnaire — Short Form (CTQ-SF). *Child Abuse & Neglect*, 120, 105223. <https://doi.org/10.1016/J.CHIABU.2021.105223>
- Georgieva, S., Tomás, J. M., Navarro-Pérez, J. J., & Samper-García, P. (2022). Systematic Review and Critical Appraisal of Five of the Most Recurrently Validated Child Maltreatment Assessment Instruments from 2010 to 2020. *Trauma, Violence, and Abuse*. https://doi.org/10.1177/15248380221097694/ASSET/IMAGES/10.1177_15248380221097694-IMG2.PNG
- Gergely, G., & Csibra, G. (2013). Natural Pedagogy. *Navigating the Social World*, 127–132. <https://doi.org/10.1093/ACPROF:OSO/9780199890712.003.0023>

- Gray, J., Jordanova Pesevska, D., Sethi, D., Ramiro González, M. D., & Yon, Y. (2016). *Handbook on developing national action plans to prevent child maltreatment. World Health Organization. Regional Office for Europe.*
- Haapasalo, J., & Aaltonen, T. (1999). Child Abuse Potential How Persistent? *Journal of Interpersonal Violence, 14*(6), 571–585.
- Harnett, P. H., & Dawe, S. (2008). Reducing child abuse potential in families identified by social services: Implications for assessment and treatment. *Brief Treatment and Crisis Intervention, 8*(3), 226–235. <https://doi.org/10.1093/brief-treatment/mhn010>
- Haskett, M. E., Scott, S. S., & Fann, K. D. (1995). Child Abuse Potential Inventory and parenting behaviour: Relationships with high-risk correlates. *Child Abuse & Neglect, 19*(12), 1483–1495. [https://doi.org/10.1016/0145-2134\(95\)00107-4](https://doi.org/10.1016/0145-2134(95)00107-4)
- Hesse, E., & Main, M. (2006). Frightened, threatening, and dissociative parental behaviour in low-risk samples: description, discussion, and interpretations. *Development and Psychopathology, 18*(2), 309–343. <https://doi.org/10.1017/S0954579406060172>
- Hien, D., Cohen, L. R., Caldeira, N. A., Flom, P., & Wasserman, G. (2010). Depression and anger as risk factors underlying the relationship between maternal substance involvement and child abuse potential. *Child Abuse & Neglect, 34*(2), 105–113. <https://doi.org/10.1016/j.chiabu.2009.05.006>
- Higgins, J. P. T., López-López, J. A., Becker, B. J., Davies, S. R., Dawson, S., Grimshaw, J. M., McGuinness, L. A., Moore, T. H. M., Rehfues, E. A., Thomas, J., & Caldwell, D. M. (2019). Synthesising quantitative evidence in systematic reviews of complex health interventions. *BMJ Glob Health, 4*, 858. <https://doi.org/10.1136/bmjgh-2018-000858>
- Holden, E. W., Willis, D. J., & Foltz, L. (1989). Child abuse potential and parenting stress: Relationships in maltreating parents. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1*(1), 64–67. <https://doi.org/10.1037/1040-3590.1.1.64>

- Hubel, G. S., Rostad, W. L., Self-Brown, S., & Moreland, A. D. (2018). Service needs of adolescent parents in child welfare: Is an evidence-based, structured, in-home behavioural parent training protocol effective? *Child Abuse and Neglect*, 79(January), 203–212.
<https://doi.org/10.1016/j.chiabu.2018.02.005>
- Jacobson, N. S., & Truax, P. (1991). Clinical Significance: A Statistical Approach to Defining Meaningful Change in Psychotherapy Research. *Journal of Consulting and Clinical Psychology*, 59(1), 12–19. <https://doi.org/10.1037/0022-006X.59.1.12>
- Jaffee, S. R. (2017). Child Maltreatment and Risk for Psychopathology in Childhood and Adulthood. *The Annual Review of Clinical Psychology Is Online At*, 13, 525–551.
<https://doi.org/10.1146/annurev-clinpsy-032816>
- Jaffee, S. R., Bowes, L., Ouellet-Morin, I., Fisher, H. L., Moffitt, T. E., Merrick, M. T., & Arseneault, L. (2013). Safe, stable, nurturing relationships break the intergenerational cycle of abuse: A prospective nationally representative Cohort of children in the United Kingdom. *Journal of Adolescent Health*, 53(4 SUPPL). <https://doi.org/10.1016/J.JADOHEALTH.2013.04.007>
- Karanicolas, P. J., Bhandari, M., Kreder, H., Moroni, A., Richardson, M., Walter, S. D., Norman, G. R., & Guyatt, G. H. (2009). Evaluating agreement: Conducting a reliability study. *Journal of Bone and Joint Surgery*, 91(SUPPL. 3), 99–106. <https://doi.org/10.2106/JBJS.H.01624>
- Kilpatrick, K. L. (2005). The Parental Empathy Measure: A New Approach to Assessing Child Maltreatment Risk. *American Journal of Orthopsychiatry*, 75(4), 608–620.
<https://doi.org/10.1037/0002-9432.75.4.608>
- Kolko, D. J. (1996). Individual Cognitive Behavioural Treatment and Family Therapy for Physically Abused Children and their Offending Parents: A Comparison of Clinical Outcomes.
<Http://Dx.Doi.Org/10.1177/1077559596001004004>, 1(4), 322–342.
<https://doi.org/10.1177/1077559596001004004>

- Landsmeer-Beker, E. A., Bouwmeester-Landweer, M. B. R., Korbee-Haverhoek, H. D., Pieter, N., Kousemaker, J., Baartman, H. E. M., Wit, J. M., & Dekker, F. W. (2006). *Early home visitation in families at risk for child maltreatment*. <https://hdl.handle.net/1887/4396>
- Larson, R., & Richards, M. H. (1991). Daily Companionship in Late Childhood and Early Adolescence: Changing Developmental Contexts. *Child Development*, *62*(2), 284–300. <https://doi.org/10.1111/J.1467-8624.1991.TB01531.X>
- Laulik, S., Allam, J., & Browne, K. (2015). The Use of the Child Abuse Potential Inventory in the Assessment of Parents involved in Care Proceedings. *Child Abuse Review*, *24*(5), 332–345. <https://doi.org/10.1002/CAR.2294>
- Li, P., Stuart, E. A., & Allison, D. B. (2015). Multiple Imputation: A Flexible Tool for Handling Missing Data HHS Public Access. *JAMA*, *314*(18), 1966–1967. <https://doi.org/10.1001/jama.2015.15281>
- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202. <https://doi.org/10.1080/01621459.1988.10478722>
- Luyten, P., Campbell, C., Allison, E., & Fonagy, P. (2020). The Mentalizing Approach to Psychopathology: State of the Art and Future Directions. <https://doi.org/10.1146/Annurev-Clinpsy-071919-015355>, *16*, 297–325. <https://doi.org/10.1146/ANNUREV-CLINPSY-071919-015355>
- Lyons-Ruth, K., Dutra, L., Schuder, M. R., & Bianchi, I. (2006). From Infant Attachment Disorganization to Adult Dissociation: Relational Adaptations or Traumatic Experiences? *The Psychiatric Clinics of North America*, *29*(1), 63. <https://doi.org/10.1016/J.PSC.2005.10.011>
- Marsh, S., Dobson, R., & Maddison, R. (2020). The relationship between household chaos and child, parent, and family outcomes: a systematic scoping review. *BMC Public Health* *20*:1, *20*(1), 1–27. <https://doi.org/10.1186/S12889-020-08587-8>

- Maxwell, S. E. (2000). Sample size and multiple regression analysis. *Psychological Methods*, 5(4), 434–458. <https://doi.org/10.1037/1082-989X.5.4.434>
- McKenna, S. P., & Heaney, A. (2021). Setting and maintaining standards for patient-reported outcome measures: can we rely on the COSMIN checklists? *Journal of Medical Economics*, 24(1), 502–511. <https://doi.org/10.1080/13696998.2021.1907092>
- Milner, J.S. (1980). *The Child Abuse Potential Inventory: Manual* ((1st ed)). Webster, N.C: Psytec Corporation
- Milner, J. S. (1986). *The Child Abuse Potential Inventory Manual* ((2nd ed.)). Webster: Psytec Inc.
- Milner, J. S. (1994). Assessing physical child abuse risk: The child abuse potential inventory. *Clinical Psychology Review*, 14(6), 547–583. [https://doi.org/10.1016/0272-7358\(94\)90017-5](https://doi.org/10.1016/0272-7358(94)90017-5)
- Milner, J. S., & Crouch, J. L. (2012). Psychometric characteristics of translated versions of the Child Abuse Potential Inventory. *Psychology of Violence*, 2(3), 239–259. <https://doi.org/10.1037/A0026957>
- Milner, J. S., Crouch, J. L., McCarthy, R. J., Ammar, J., Dominguez-Martinez, R., Thomas, C. L., & Jensen, A. P. (2022). Child physical abuse risk factors: A systematic review and a meta-analysis. *Aggression and Violent Behavior*, 66(November 2019), 101778. <https://doi.org/10.1016/j.avb.2022.101778>
- Miragoli, S., Balzarotti, S., Camisasca, E., & Di Blasio, P. (2018). Parents' perception of child behavior, parenting stress, and child abuse potential: Individual and partner influences. *Child Abuse & Neglect*, 84, 146–156. <https://doi.org/10.1016/J.CHIABU.2018.07.034>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>

- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & PRISMA-P Group (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic reviews*, 4(1), 1.
<https://doi.org/10.1186/2046-4053-4-1>
- Mokkink, L. B., De Vet, H C W, Prinsen, C A C, Patrick, D L, Alonso, J, Bouter, L M, & Terwee, C B. (2018). *COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures*. 27, 1171–1179. <https://doi.org/10.1007/s11136-017-1765-4>
- Monroe, L. D., & Schellenbach, C. J. (2008). Relationship of Child Abuse Potential Inventory Scores to Parental Responses: [Http://Dx.Doi.Org/10.1300/J019v11n03_03](http://Dx.Doi.Org/10.1300/J019v11n03_03), 11(3–4), 39–58.
https://doi.org/10.1300/J019V11N03_03
- Mulcahy, G., Badger, J., Wright, H., & Erskine, C. (2014). ‘What happened next’: a study of outcomes for maltreated children following care proceedings. *Adoption and Fostering*, 38(4), 314–330. https://doi.org/10.1177/0308575914553362/ASSET/IMAGES/LARGE/10.1177_0308575914553362-FIG2.JPEG
- Mulder, T. M., Kuiper, K. C., van der Put, C. E., Stams, G. J. J. M., & Assink, M. (2018). Risk factors for child neglect: A meta-analytic review. *Child Abuse and Neglect*, 77, 198–210.
<https://doi.org/10.1016/J.CHIABU.2018.01.006>
- Nobes, G., & Smith, M. (2000). The Relative Extent of Physical Punishment and Abuse by Mothers and Fathers. *Trauma, Violence, & Abuse*, 1(1), 47–66.
<https://doi.org/10.1177/1524838000001001004>
- Nolte, T., Bolling, D. Z., Hudac, C. M., Fonagy, P., Mayes, L., Pelphrey, K. A., Coan, J. A., & Berkman, E. (2013). *Brain mechanisms underlying the impact of attachment-related stress on social cognition*. <https://doi.org/10.3389/fnhum.2013.00816>

- Norman, R. E., Byambaa, M., De, R., Butchart, A., Scott, J., & Vos, T. (2012). The Long-Term Health Consequences of Child Physical Abuse, Emotional Abuse, and Neglect: A Systematic Review and Meta-Analysis. *PLoS Medicine*, 9(11).
<https://doi.org/10.1371/journal.pmed.1001349>
- Office for National Statistics. (2021). *Census Profile for areas in England and Wales*. UK Public Record Office
(TS007).https://www.nomisweb.co.uk/sources/census_2021/report?compare=E06000004
- Office for National Statistics (2021). *Families and households in the UK: 2014*. UK Public Record Office.
<https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/bulletins/familiesandhouseholds/2015-01-28#main-findings>
- Ondersma, S. J., Chaffin, M. J., Mullins, S. M., & Lebreton, J. M. (2005). A brief form of the child abuse potential inventory: Development and validation. *Journal of Clinical Child and Adolescent Psychology*, 34(2), 301–311. https://doi.org/10.1207/S15374424JCCP3402_9
- Orme, W., Bowersox, L., Vanwoerden, S., Fonagy, P., & Sharp, C. (2019). The relation between epistemic trust and borderline pathology in an adolescent inpatient sample. *Borderline Personality Disorder and Emotion Dysregulation*, 6(1), 1–9. <https://doi.org/10.1186/S40479-019-0110-7/TABLES/2>
- Penfield, T., Baker, M. J., Scoble, R., & Wykes, M. C. (2014). Assessment, evaluations, and definitions of research impact: A review. *Research evaluation*, 23(1), 21-32.
- Pereira, J., Vickers, K., Atkinson, L., Gonzalez, A., Wekerle, C., & Levitan, R. (2012). Parenting stress mediates between maternal maltreatment history and maternal sensitivity in a community sample. *Child Abuse & Neglect*, 36(5), 433–437.
<https://doi.org/10.1016/J.CHIABU.2012.01.006>
- Pinheiro, P. S. (2006). Violence against children. Geneva: ATAR Roto Presse SA.

- Plant, C. P., Donohue, B., & Holland, J. M. (2016). Examination of life satisfaction, child maltreatment potential and substance use in mothers referred for treatment by child protective services for child neglect and substance abuse: Implications for intervention planning. *Applied Research in Quality of Life*, *11*(3), 805–816.
<https://doi.org/10.1007/s11482-015-9398-7>
- Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, *27*(5), 1147–1157. <https://doi.org/10.1007/s11136-018-1798-3>
- Richey, A., Brown, S., Fite, P. J., & Bortolato, M. (2016). The Role of Hostile Attributions in the Associations between Child Maltreatment and Reactive and Proactive Aggression. *Journal of Aggression, Maltreatment & Trauma*, *25*(10), 1043.
<https://doi.org/10.1080/10926771.2016.1231148>
- Rinehart, D. J., Becker, M. A., Buckley, P. R., Dailey, K., Reichardt, C. S., Graeber, C., VanDeMark, N. R., & Brown, E. (2005). The Relationship Between Mothers' Child Abuse Potential and Current Mental Health Symptoms: Implications for Screening and Referral. *The Journal of Behavioral Health Services & Research*, *32*(2), 155–166.
<https://doi.org/10.1007/BF02287264>
- Ríos, M., Zekri, S., Alonso-Esteban, Y., & Navarro-Pardo, E. (2022). Parental Stress Assessment with the Parenting Stress Index (PSI): A Systematic Review of Its Psychometric Properties. *Children*, *9*(11), 1649.
- Rodriguez, C. M., & Silvia, P. J. (2022). Psychometric Evidence for Indirect Assessment of Child Abuse Risk in Child Welfare-Involved Mothers. *Children*, *9*(5), 1–12.
<https://doi.org/10.3390/children9050711>

- Rodriguez, C. M., Smith, T. L., & Silvia, P. J. (2016). Parent–Child Aggression Risk in Expectant Mothers and Fathers: A Multimethod Theoretical Approach. *Journal of Child and Family Studies*, 25(11), 3220. <https://doi.org/10.1007/S10826-016-0481-Y>
- Rose, J., Nice, L., Kroese, B. S., Powell, T., & Oyebode, J. (2019). The role of relationship reciprocity and self-efficacy on wellbeing and burnout in clinical psychology trainees. *Clinical Psychology Forum*, 2019(315), 38–48. <https://doi.org/10.53841/bpscpf.2019.1.315.38>
- Saini, S. M., Hoffmann, C. R., Pantelis, C., Everall, I. P., & Bousman, C. A. (2019). Systematic review and critical appraisal of child abuse measurement instruments. *Psychiatry Research*, 272, 106–113. <https://doi.org/10.1016/J.PSYCHRES.2018.12.068>
- 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(3), 513–535. [https://doi.org/10.1016/S0005-7894\(04\)80030-3](https://doi.org/10.1016/S0005-7894(04)80030-3)
- Schäfer, T., & Schwarz, M. A. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology*, 10(APR), 813. <https://doi.org/10.3389/FPSYG.2019.00813/BIBTEX>
- Schilling, E. A., Aseltine, R. H., & Gore, S. (2007). *Adverse childhood experiences and mental health in young adults: a longitudinal survey*. <https://doi.org/10.1186/1471-2458-7-30>
- Schofield, T. J., Lee, R. D., & Merrick, M. T. (2013). Safe, stable, nurturing relationships as a moderator of intergenerational continuity of Child Maltreatment: A meta-analysis. *Journal of Adolescent Health*, 53(4 SUPPL). <https://doi.org/10.1016/J.JADOHEALTH.2013.05.004>
- Sim, J., & Wright, C. C. (2005). The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Physical Therapy*, 85(3), 257–268. <https://doi.org/10.1093/PTJ/85.3.257>

- Slade, A. (2005). Parental reflective functioning: An introduction. *Attachment and Human Development*, 7(3), 269–281. <https://doi.org/10.1080/14616730500245906>
- Slade, A. (2011). *The Pregnancy Interview – Revised*. (Unpublished manuscript). New York, NY: The City University of New York.
- Slade, A., Aber, J. L., Bresgi, I., Berger, B., & Kaplan, M. (2004). *The Parent Development Interview – Revised*. Unpublished protocol. New York, NY: The City University of New York. The City University of New York.
- Slade, A., Aber, J. L., Berger, B., Bresgi, I., & Kaplan, M. (2020). *Parental Developmental Interview Short Version*. [Unpublished interview protocol]. Yale Child Study Centre.
- Sleed, M. (2014). *The assessment of relational risk in early parent-infant relationships* [Doctoral dissertation, UCL (University College London)].
- Sleed, M., Fearon, P., Midgley, N., Martin, P., Byrne, G., & Zywek, L. (2021). *The Supporting Parents Project (SPP): Randomised Controlled Trial and implementation process study of the Lighthouse Parenting Programme (LPP)*. <https://doi.org/https://doi.org/10.17605/OSF.IO/GXYS9>
- Sleed, M., Isosävi, S., & Fonagy, P. (2021). *The assessment of representational risk (ARR): Development and psychometric properties of a new coding system for assessing risk in the parent-infant relationship*. <https://doi.org/10.1002/imhj.21932>
- Sleed, M., James, J., Baradon, T., Newbery, J., & Fonagy, P. (2013). A psychotherapeutic baby clinic in a hostel for homeless families: Practice and evaluation. *Psychology and Psychotherapy: Theory, Research and Practice*, 86(1), 1-18. Statista Research Department (2023), *Child abuse in the U.S. – number of victims 2021, by perpetrator relationship*.

- Steele, H., Bate, J., Steele, M., Dube, S. R., Danskin, K., Knafo, H., Nikitiades, A., Bonuck, K., Meissner, P., & Murphy, A. (2016). Adverse childhood experiences, poverty, and parenting stress. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement*, 48(1), 32–38. <https://doi.org/10.1037/cbs0000034>
- Stith, S. M., Liu, T., Davies, L. C., Boykin, E. L., Alder, M. C., Harris, J. M., Som, A., McPherson, M., & Dees, J. E. M. E. G. (2009). Risk factors in child maltreatment: A meta-analytic review of the literature. *Aggression and Violent Behavior*, 14(1), 13–29. <https://doi.org/10.1016/J.AVB.2006.03.006>
- Stoltenborgh, M., Bakermans- Kranenburg, M. J., Alink, L. R. A., & Van IJzendoorn, M. H. (2015). The Prevalence of Child Maltreatment across the Globe: Review of a Series of Meta-Analyses. *Child Abuse Review*, 24(October 2014), 37–50. <https://doi.org/10.1002/car.2353>
- Straus, M. A. (1979). Measuring Intrafamily Conflict and Violence: The Conflict Tactics (CT) Scales. *Journal of Marriage and the Family*, 41(1), 75. <https://doi.org/10.2307/351733>
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Allyn and Bacon.
- Terry, M., Finger, B., Lyons-Ruth, K., Sadler, L. S., & Slade, A. (2020). Hostile/Helpless maternal representations in pregnancy and later child removal: A pilot study. *Infant Mental Health Journal*, 42, 60–73. <https://doi.org/10.1002/imhj.21887>
- Terwee, C. B., Prinsen, C A C, Chiarotto, A, Westerman, M J, Patrick, D L, Alonso, J, Bouter, L M, De Vet, H C W, & Mokkink, L B. (2018). *COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study*. 27, 1159–1170. <https://doi.org/10.1007/s11136-018-1829-0>
- Thomas, R., & Zimmer-Gembeck, M. J. (2011). Accumulating Evidence for Parent-Child Interaction Therapy in the Prevention of Child Maltreatment. *Child Development*, 82(1), 177–192. <https://doi.org/10.1111/j.1467-8624.2010.01548.x>

- Thornberry, T. P., & Henry, K. L. (2013). Intergenerational Continuity in Maltreatment. *Journal of Abnormal Child Psychology*, 41(4), 555. <https://doi.org/10.1007/S10802-012-9697-5>
- Thornberry, T. P., Knight, K. E., & Lovegrove, P. J. (2012). Does Maltreatment Beget Maltreatment? A Systematic Review of the Intergenerational Literature. <Http://Dx.Doi.Org/10.1177/1524838012447697>, 13(3), 135–152. <https://doi.org/10.1177/1524838012447697>
- Timmer, S. G., Urquiza, A. J., Zebell, N. M., & McGrath, J. M. (2005). Parent-Child Interaction Therapy: Application to maltreating parent-child dyads. *Child Abuse & Neglect*, 29(7), 825–842. <https://doi.org/10.1016/j.chiabu.2005.01.003>
- Tucker, M. C., Rodriguez, C. M., & Baker, L. R. (2017). Personal and couple level risk factors: Maternal and paternal parent-child aggression risk. *Child Abuse & Neglect*, 69, 213–222. <https://doi.org/10.1016/J.CHIABU.2017.04.030>
- Urgelles, J., Donohue, B., Wilks, C., Van Hasselt, V. B., & Azrin, N. H. (2012). A standardized method of preventing and managing emergencies within the context of evidence-based therapy implementation. *Behavior Modification*, 36(4), 558–579. <https://doi.org/10.1177/0145445512448192>
- US Department of Health and Human Services. (2001). Child maltreatment 1999. *Washington, DC: US Government Printing Office.* https://www.acf.hhs.gov/sites/default/files/documents/cm99_0.pdf
- US Department of Health and Human Services. (2006). Child maltreatment 2006. *Washington, DC: US Government Printing Office.* <https://www.acf.hhs.gov/sites/default/files/documents/cb/cm06.pdf>
- US Department of Health and Human Services. (2021). Child maltreatment 2023. *Washington, DC: US Government Printing Office.* <https://www.acf.hhs.gov/sites/default/files/documents/cb/cm2021.pdf>

- Vachon, D. D., Krueger, R. F., Rogosch, F. A., & Cicchetti, D. (2015). Assessment of the harmful psychiatric and behavioral effects of different forms of child maltreatment. *JAMA Psychiatry*, 72(11), 1135–1142. <https://doi.org/10.1001/jamapsychiatry.2015.1792>
- van IJzendoorn, M. H. (1992). Intergenerational transmission of parenting: A review of studies in nonclinical populations. *Developmental Review*, 12(1), 76–99. [https://doi.org/10.1016/0273-2297\(92\)90004-L](https://doi.org/10.1016/0273-2297(92)90004-L)
- van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., Coughlan, B., & Reijman, S. (2020). Annual Research Review: Umbrella synthesis of meta-analyses on child maltreatment antecedents and interventions: differential susceptibility perspective on risk and resilience. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 61(3), 272–290. <https://doi.org/10.1111/jcpp.13147>
- van IJzendoorn, M. H., Schuengel, C., & Bakermans-Kranenburg, M. J. (1999). Disorganized attachment in early childhood: meta-analysis of precursors, concomitants, and sequelae. *Development and Psychopathology*, 11(2), 225–249. <https://doi.org/10.1017/S0954579499002035>
- Vorhies, V., Glover, C. M., Davis, K., Hardin, T., Krzyzanowski, A., Harris, M., Fagan, M., & Wilkniss, S. (2009). Improving outcomes for pregnant and parenting foster care youth with severe mental illness: An evaluation of a transitional living program. *Psychiatric Rehabilitation Journal*, 33(2), 115–124. <https://doi.org/10.2975/33.2.2009.115.124>
- Walker, C. A., & Davies, J. (2010). A critical review of the psychometric evidence base of the Child Abuse Potential Inventory. *Journal of Family Violence*, 25(2), 215–227. <https://doi.org/10.1007/s10896-009-9285-9>
- Wang, C.-T., & Holton, J. (2007). Total Estimated Cost of Child Abuse and Neglect in the United States. *Economic Impact Study*, September, 1–5. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.192.2911&rep=rep1&type=pdf>

- WHO. (2022). *Child maltreatment*. <https://www.who.int/news-room/fact-sheets/detail/child-maltreatment>
- WHO. (2020). *Child Maltreatment*. <https://www.who.int/news-room/fact-sheets/detail/child-maltreatment>
- Widom, C. S. (1989). Does violence beget violence? A critical examination of the literature. *Psychological Bulletin*, *106*(1), 3–28. <https://doi.org/10.1037/0033-2909.106.1.3>
- Widom, C. S., Czaja, S. J., & DuMont, K. A. (2015). Intergenerational transmission of child abuse and neglect: Real or detection bias? *Science (New York, N.Y.)*, *347*(6229), 1480. <https://doi.org/10.1126/SCIENCE.1259917>
- Wilson, A., & Marshall, A. (2010). The support needs and experiences of suicidally bereaved family and friends. *Death studies*, *34*(7), 625-640.
- Winter, S. M., Dittrich, K., Dörr, P., Overfeld, J., Moebus, I., Murray, E., Karaboycheva, G., Zimmermann, C., Knop, A., Voelkle, M., Entringer, S., Buss, C., Haynes, J. D., Binder, E. B., & Heim, C. (2022). Immediate impact of child maltreatment on mental, developmental, and physical health trajectories. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *63*(9), 1027–1045. <https://doi.org/10.1111/JCPP.13550>
- Yoon, S., Speyer, R., Cordier, R., Aunio, P., & Hakkarainen, A. (2022). A Systematic Review on Evaluating Responsiveness of Parent- or Caregiver-Reported Child Maltreatment Measures for Interventions. *Trauma, violence & abuse*, 15248380221093690. Advance online publication. <https://doi.org/10.1177/15248380221093690>
- Youngstrom, E. A., Frazier, T. W., Prinstein, M. J., Ong, M.-L., & Youngstrom, J. K. (2017). *Evidence-Based Assessment as an Integrative Model for Applying Psychological Science to Guide the Voyage of Treatment*. <https://doi.org/10.1111/cpsp.12207>

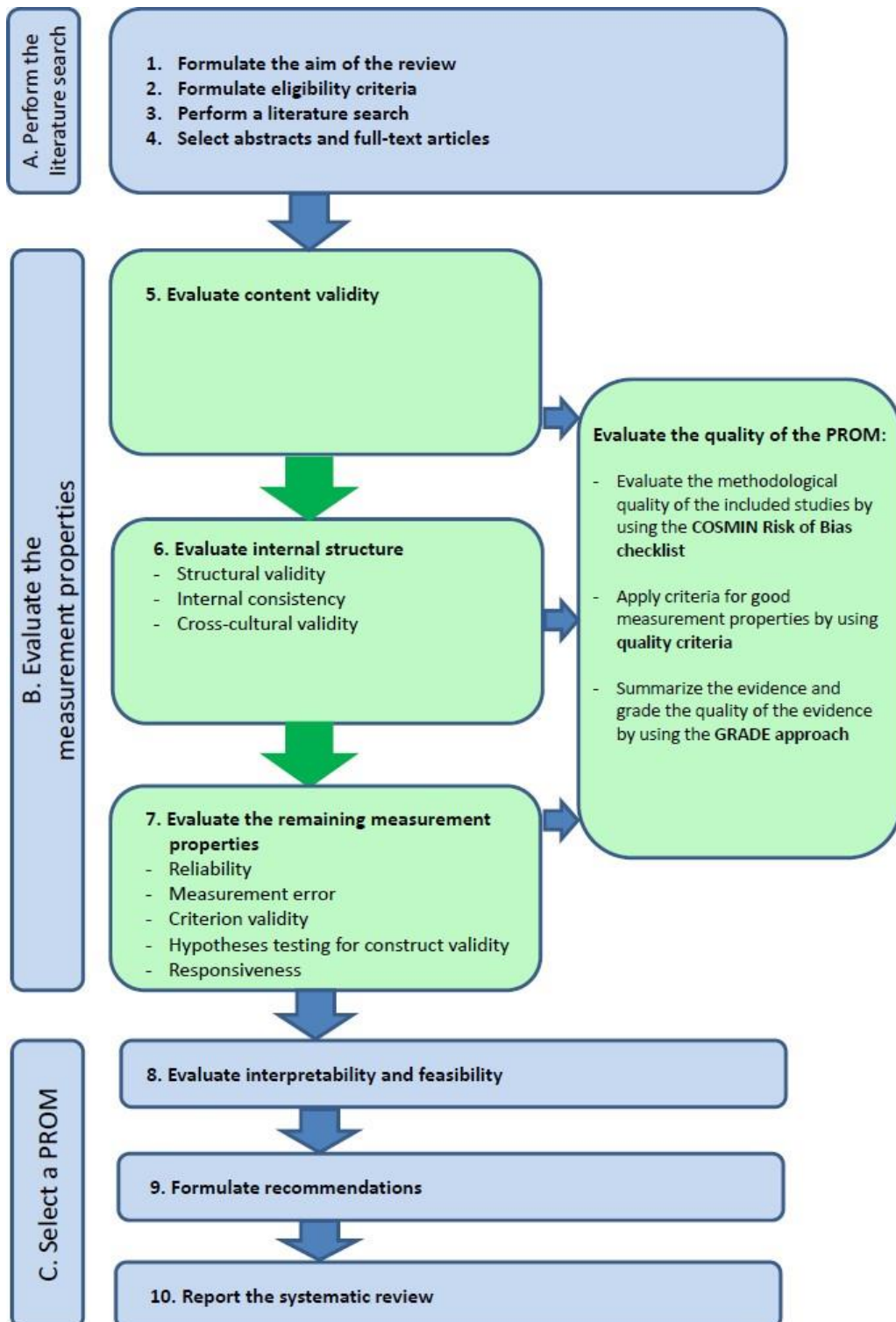
Zeegers, M. A. J., Colonnesi, C., Stams, G. J. J. M., & Meins, E. (2017). Mind matters: A meta-analysis on parental mentalization and sensitivity as predictors of infant-parent attachment. *Psychological Bulletin*, 143(12), 1245–1272. <https://doi.org/10.1037/BUL0000114>

Appendix A

Ten step procedure and outline of the COSMIN manual

Figure 1.

Ten steps for conducting a systematic review of PROMs



Appendix B

Step 1 – Cosmin Risk of Bias checklist

Table B1

Risk of Bias checklist used to assess study methodological quality adapted from the COSMIN manual for systematic reviews of measures (Mokkink et al., 2018).

Psychometric property	Aspect	Standard^a	Item description
Hypothesis testing for construct validity	<i>Box 9a:</i> Comparison with other outcome measurement (Convergent validity)	Design requirements	Is it clear what the comparator instrument(s) measure(s)? Were the measurement properties of the comparator instrument(s) sufficient?
		Statistical methods	Were design and statistical methods adequate for the hypotheses to be tested?
		Other flaws	Were there any other important flaws in the design or statistical methods of the study?
	<i>Box 9b:</i> Comparison between subgroups (Criterion and known-groups validity)	Design requirements	Was an adequate description provided of important characteristics of the subgroups?
		Statistical methods	Were design and statistical methods adequate for the hypothesis to be tested?
		Other flaws	Were there any other important flaws in the design or statistical methods of the study?
Responsiveness	<i>Box 10d:</i> Comparison before and after an intervention	Design requirements	Was an adequate description provided of the intervention given?
		Statistical methods	Was the statistical method appropriate for the hypotheses to be tested?

	Other flaws	Were there any other important flaws in the design or statistical methods of the study?
<i>Box 10c: Comparison between treatment conditions</i>	Design requirements	Was an adequate description provided of important characteristics of the subgroups?
	Statistical methods	Was the statistical method appropriate for the hypotheses to be tested?
	Other flaws	Were there any other important flaws in the design or statistical methods of the study?

Note. The Risk of Bias checklist was used for assessing the methodological quality of studies.

^a Each standard on methodological quality was rated using a four-point rating scale: inadequate, doubtful, adequate, and very good.

Appendix C

Step 2 – Quality criteria for good psychometric properties

Table C1

Criteria for good convergent-, criterion- and known-group validity and responsiveness adapted from the COSMIN manual for systematic reviews of measures (Mokkink et al., 2018).

Psychometric property	Aspect	Rating^a	Quality criteria^b
Hypothesis testing for construct validity	Convergent validity: how closely the instrument is related to other variables and other measures of the same construct	+	The result is in accordance with the hypothesis
		?	No hypothesis defined by the review team
		-	The result is not in accordance with the hypothesis
	Criterion/Known-groups validity: how well the tool can differentiate between groups where a difference is expected <i>a priori</i>	+	The result is in accordance with the hypothesis
		?	No hypothesis defined by the review team
		-	The result is not in accordance with the hypothesis
Responsiveness	Comparison before and after a parenting intervention	+	Meaningful changes in scores before and after intervention (e.g., Hedges' $g \geq 0.50$)

	?	Not all information for '+' reported (e.g., lack of information to calculate Hedges' g)
	-	Criteria for '+' not met (e.g., Hedges' $g < 0.50$)
Comparison between treatment conditions	+	Meaningful changes in scores between treatment conditions (e.g., Hedges' $g \geq 0.50$)
	?	Not all information for '+' reported (e.g., lack of information to calculate Hedges' g)
	-	Criteria for '+' not met (e.g., Hedges' $g < 0.50$)

Note. The criteria for good responsiveness was used for rating the results of single studies on responsiveness and rating the pooled results of all studies per measure.

^a + = Sufficient, - = Insufficient, ? = Indeterminate.

^b The quality criterion for good responsiveness on comparison of change scores before and after intervention and between intervention treatment conditions was determined as a medium effect size (Hedges' $g = 0.5$) using (Cohen, 1988) conventions to interpret effect size, which was decided by the review team for this current review as suggested by the COSMIN manual (Mokkink et al., 2018).

Appendix D

Step 2 - Applying criteria for good psychometric properties: convergent-, criterion, known-group validity, and responsiveness

Table D1

Additional Analysis: CAPI Abuse scale convergent validity by study

Study	Convergent validity measure	Relationship between measures/constructs	Results (r)	Expected correlation	Expected correlation met?	Quality rating
Bradshaw et al., 2011 Parent Satisfaction with Youth Scale (PSYS); parents' content with their children across various domains in the parent-child relationship	PSYS – communication	Dissimilar but related	(-.24*)	0.1 < r <.5	Yes	+
	PSYS – relationship	Dissimilar but related	(-.30**)	0.1 < r <.5	Yes	+
	PSYS – reaction to my praise	Dissimilar but related	(-.24*)	0.1 < r <.5	Yes	+
	PSYS – overall happiness	Dissimilar but related	(-.29**)	0.1 < r <.5	Yes	+
	PSYS – safety skills	Dissimilar but related	(-.25)	0.1 < r <.5	Yes	+
	PSYS – family involvement	Dissimilar but related	(-.18)	0.1 < r <.5	Yes	+
	PSYS – following household rules	Dissimilar but related	(-.29)	0.1 < r <.5	Yes	+
	PSYS – compliance	Dissimilar but related	(-.24)	0.1 < r <.5	Yes	+
	PSYS – household chores	Dissimilar but related	(-.18)	0.1 < r <.5	Yes	+

	PSYS – school and education activities	Dissimilar but related	(-.35*)	0.1 < r <.5	Yes	+
Budd et al., 2000 (valid CAPI profiles only)	WRAT-R reading standard	Dissimilar but related	(-.41**)	0.1 < r <.5	Yes	+
Wide Range Achievement Test (WRAT); standardised screening measure of educational skills in reading, arithmetic, spelling	WRAT-R arithmetic standard	Dissimilar but related	(-.42**)	0.1 < r <.5	Yes	+
	WRAT-R Combined	Dissimilar but related	(-.30*)	0.1 < r <.5	Yes	+
Symptom Checklist 90-Revised (SCL-90 R); Global Severity Index (GSI) T score (number of psychological symptoms and magnitude of perceived disturbance)	GSI T score	Dissimilar but related	(.44**)	0.1 < r <.5	Yes	+
Parent Opinion Questionnaire (POQ); measure of unrealistic parental expectations regarding appropriate child behaviour	POQ	Dissimilar but related	(.27)	0.1 < r <.5	Yes	+
Home Observation for the Measurement of the Environment (HOME) Inventory; procedure that measures quant/qual of social, emotional, cognitive support available to child in the home	Home inventory	Dissimilar but related	(-.29)	0.1 < r <.5	Yes	+
Arizona Social Support Interview Schedule	ASSIS Total support network	Dissimilar but related	(-.18)	0.1 < r <.5	Yes	+

(ASSIS); quantity and quality of an individual's social contacts in various functional domains	Average support satisfaction	Dissimilar but related	(-.43**; valid & invalid CAPI profiles)	0.1 < r <.5	Yes	+
	Years of education	Dissimilar but related	(-.27)	0.1 < r <.5	Yes	+
	Mother age	Dissimilar but related	(-.18)	0.1 < r <.5	Yes	+
	Multiple regression	Model to be significant; Exploratory, no H concerning individual predictors	WRAT-R Combined ($\beta=-.30^*$); GSI T ($\beta=.40^{***}$); Assis Average satisfaction ($\beta=-.22^*$); POQ total ($\beta=-.07$); HOME total ($\beta=-.10$); School grade completed ($\beta=-.11$)	$R^2=.47$, $F(6,53) = 7.843^{***}$? ^a
Revised CAPI Abuse Scale (without Distress factor items)	Multiple regression	Model to be significant; Exploratory, no H concerning individual predictors	WRAT-R Combined ($\beta=-.33^*$); GSI T ($\beta=.16$); ASSIS Average satisfaction ($\beta=-.29^*$); POQ Total ($\beta=.05$); HOME Total ($\beta=-.10$); School grade completed ($\beta=-.09$)	$R^2= .43$, $F(6,53) = 6.593^{***}$? ^a
Budd et al., 2006 (valid profiles only)	PSI-SF at T2	Dissimilar but related	(.44**)	0.1 < r <.5	Yes	+
	GSI	Dissimilar but related	(.37*)	0.1 < r <.5	Yes	+
	POQ Total score	Dissimilar but related	(.32*)	0.1 < r <.5	Yes	+
	HOME Total score	Dissimilar but related	(-.39*)	0.1 < r <.5	Yes	+

	ASSIS total positive support network	Dissimilar but related	(-.42*)	0.1 < r <.5	Yes	+
	ASSIS support satisfaction	Dissimilar but related	(-.41*)	0.1 < r <.5	Yes	+
Caliso & Milner, 1992	CTS - reasoning subscale	Dissimilar but related	(.25**)	0.1 < r <.5	Yes	+
Conflicts Tactics Scale (CTS); modified CTS designed to assess family conflict resolution techniques as a measure of parents' childhood history of abuse	CTS - verbal abuse subscale	Dissimilar but related	(.25**)	0.1 < r <.5	Yes	+
	CTS - violence subscale	Dissimilar but related	(.48***)	0.1 < r <.5	Yes	+
Caliso & Milner, 1994	Childhood Social Network Questionnaire (CSNQ); measures childhood social support, defined as the presence of childhood relationships before the age of 12 that were perceived as providing some degree of cognitive and emotional support.	Model to be significant; Exploratory, no H concerning individual predictors	In order: father support (12.4%; β =negative*); sibling support (9.4%; β = negative*); mother support (8.6%; β = negative*); sibling punishment (6.2%; positive*); outside support (5.9%; β = negative*); mother punishment (4.5%; β = positive*); friend support (n.s.); father punishment (not significant)		CSNQ factors accounted for 47% CAPI Abuse score variance	? ^a
Craig & Sprang, 2007	Age	Dissimilar but related	(-.21)	0.1 < r <.5	Yes	+
	Race	Dissimilar but related	(.11)	0.1 < r <.5	Yes	+
	Gender	Dissimilar but related	(.13)	0.1 < r <.5	Yes	+

Childhood physical abuse (CPA)	Dissimilar but related	(.18)	0.1 < r <.5	Yes	+
Childhood sexual abuse (CSA)	Dissimilar but related	(.23)	0.1 < r <.5	Yes	+
Adult physical abuse (APA)	Dissimilar but related	(.21)	0.1 < r <.5	Yes	+
Adult sexual abuse (ASA)	Dissimilar but related	(.27)	0.1 < r <.5	Yes	+
Domestic violence (DV)	Dissimilar but related	(.24)	0.1 < r <.5	Yes	+
Disaster	Dissimilar but related	(.08)	0.1 < r <.5	Yes	+
Death of a loved one (DOL)	Dissimilar but related	(.13)	0.1 < r <.5	Yes	+
Motor vehicle accident (MVA)	Dissimilar but related	(.11)	0.1 < r <.5	Yes	+
Hierarchical regression of eight types of traumas on square root of CAPI Abuse scale scores	H: parents' (1) CSA and (2) CPA history will be stronger predictors of caregiver child abuse potential than other trauma variables	Age, Race and Gender (R ²) = .06* ; Age, Race, Gender, CPA (R ²) = .02* ; Age, Race, Gender, CPA, CSA (R ²) = .03* ; Age, Race, Gender, CPA, CSA, APA (R ²) = .02* ; Age, Race, Gender, CPA, CSA, APA, ASA (R ²) = .02* ; Age, Race, Gender, CPA, CSA, APA, ASA, DV (R ²) = .01* ; Age, Race, Gender, CPA, CSA, APA, ASA, DV, Disaster, Motor vehicle	R ² =.16, F(11, 1047) = 19.05***		+ (1) - (2)

			accident, death of loved one (R^2) = .00				
	Logistic regression on likelihood of belonging to 'high-risk' classification (CAPI Abuse scale score ≥ 215)	H: the three trauma groups will have significantly higher odds of predicting child abuse potential scores in the clinical range than will the group without trauma exposure	Child only trauma > no trauma (*) (2,26x); Adult only trauma > no trauma (*) (3.03x); Child-adult trauma > no trauma (*) (4.23x) non-Caucasian group > Caucasian group (*) (1.86x); women > men (*) (2.1x); younger (≤ 32) > older (*) 53%	$R^2 = .15$	Yes	+ ³	
Donohue et al., 2016	Home appearance factor	Dissimilar but related	(.21)	$0.1 < r < .5$	Yes	+	
	Home safety factor	Dissimilar but related	(.16)	$0.1 < r < .5$	Yes	+	
Donohue et al., 2017							
	Relationship between 'child neglect incompatible parenting actions' (i.e., positive, supportive non-abusive parenting actions) that parents self-reported as occurring infrequently at baseline, with CAPI Abuse at baseline	Contingency management	Dissimilar but related	(.63 ^{**})	$0.1 < r < .5$	Yes	+
		H: infrequently occurring positive	Neglect incompatible actions (i.e., positive parenting actions) self-	$F(1,16) = 43.95^{**}$	Yes	+ ²	

	Multiple regression	parenting actions at baseline would not predict CAPI Abuse scale scores post-treatment; number of parenting actions set as goals at baseline would predict CAPI Abuse scale score post-treatment	reported as occurring infrequently at baseline did not predict CAPI Abuse posttreatment ($\beta=.284$) after controlling for baseline CAPI Abuse	$< .001$, $R^2 = .733$, $adjR^2 = .716$.		
			Number of positive parenting goals set at treatment initiation significantly predicted child abuse potential following treatment ($\beta=-.529^*$)			
Donohue et al., 2019	PR hard drug abuse	Dissimilar but related	(.23*)	$0.1 < r < .5$	Yes	+
Parent-reported drug abuse/marijuana use	PR marijuana	Dissimilar but related	(.15)	$0.1 < r < .5$	Yes	+
	Hierarchical regression	H: Participants' self-reported hard drug use would show greater utility in the prediction of CAPI Abuse than participant self-reported marijuana use			Yes	+
			Control variables (change in $R^2=.467^{**}$); PR hard-drug use + marijuana (change in $R^2=.074^{**}$); PR hard drug-use ($\beta=.230^{**}$); PR Marijuana use ($\beta=.193^*$)			
Significant others' (SO) report of mothers' hard drug/marijuana abuse	SO report of mothers' hard drug abuse	Dissimilar but related	(.19)	$0.1 < r < .5$	Yes	+

	SO' report of mothers' marijuana	Dissimilar but related	(.01)	0.1 < r <.5	Yes	+
	Hierarchical regression	Model to be significant; Exploratory, no H concerning individual predictors	Control variables (change in R ² =.047*); SO-reported hard drug abuse + SO-reported marijuana use (change in R ² =.016; SO-reported hard-drug use: β=.131; SO-reported marijuana use: β=-.007)			?
Haapsalo & Altonen, 1999	Childhood physical abuse	Dissimilar but related	(.29)	0.1 < r <.5	Yes	+
	Childhood psychological abuse	Dissimilar but related	(.16)	0.1 < r <.5	Yes	+
	SES variable based on education, occupational status, employment, and income	SES	Dissimilar but related	(.35)	0.1 < r <.5	Yes
Sum of five dichotomous variables: alcohol/drug abuse, institutionalization for mental health problems, medication for mental health problems, convictions for crimes and having been in jail	Social problems	Dissimilar but related	(.48*)	0.1 < r <.5	Yes	+
	Hierarchical regression	H: Both mothers' childhood experiences and their life	Step 3: Group (CPS; control) (β = .13); Childhood physical abuse (β=.10) Childhood psychological abuse			- +

		stresses, reflected in social problems and socio-economic difficulties, would be likely to contribute to CAPI Abuse scale score prediction	($\beta=.19$); Social problems ($\beta=.69^{**}$); SES ($\beta=.00$)			
Haskett et al., 1995	GSI	Dissimilar but related	(.79 ^{***})	0.1 < r <.5	No (r>.5)	-
Adult-Adolescent Parenting Inventory (AAPI); to measure parents' belief in the value of corporal punishment	AAPI; Physical punishment subscale	Dissimilar but related	(-.12)	0.3 < r <.5	No (r<.3)	-
	CBCL (internalizing problems)	Dissimilar but related	(.32 [*])	0.1 < r <.5	Yes	+
	Parent report form	CBCL (externalizing problems)	Dissimilar but related	(.45 ^{**})	0.1 < r <.5	Yes
The problems questionnaire developed specifically maltreating parents; 15 hypothetical child- and non-child related problem situations; in which parents report what could be done/they would do	CBCL (Total child problems)	Dissimilar but related	(.61 ^{***})	0.1 < r <.5	No	-
	Parental problem solving (mean effectiveness)	Dissimilar but related	(-.17)	0.1 < r <.5	Yes	+
	Parental problem solving (mean number of solutions)	Dissimilar but related	(-.04)	0.1 < r <.5	No (r<.1)	-

Child Adjustment - Teacher report form (TRF)	TRF internalizing	Dissimilar but related	(.30)	0.1 < r <.5	Yes	+
	TRF externalizing	Dissimilar but related	(.30)	0.1 < r <.5	Yes	+
	TRF Total problems	Dissimilar but related	(.29)	0.1 < r <.5	Yes	+
Parent-child dyads observed for 20-minutes in clinic to assess parenting during situations of high/low task demands. Negative behaviour defined as verbalizations indicating annoyance, criticisms, or physical reprimand (e.g., slapping, taking toys away); positive/neutral behaviours defined as hand-holding, responses to questions, compliments etc	Observed parenting behaviour; negative/positive/neutral behaviour, commands	Dissimilar but related	(.55***)	0.1 < r <.5	Yes	+
Haskett et al., 1995 (High CAP ≥ 215 vs Normal CAP < 215); observed parenting index; GSI; CBCL (externalizing, internalizing, Total problem behaviour); TRF (externalizing, internalizing, Total problem); AAPI; Parental problem solving	H: Higher CAP associated with more negative parent-child interactions, higher CBCL, GSI, AAPI, lower parental problem solving, higher TRF	Observed parenting index** GSI*** Ext. Beh.* Total problem Beh.** Remaining variables n.s.	p < .05	Yes	+8 -3	
Hien et al., 2010	Stepwise Hierarchical regression	H: anger arousal and	Block 1: (R ² = .11**): Demographics Age (β = -0.12); Divorced marital			+

Weschler Adult Intelligence scale (3 rd edition) (WAIS-III); vocab and information subtests only; Novaco Anger Inventory (NAI); measures of anger arousal and reactivity to operationalise emotion regulation	reactivity will significantly contribute to a model of child abuse potential that includes other salient factors, such as demographic variables, diagnostic histories of substance use and depressive disorders	status ($\beta=.18^*$); SES ($\beta=-.11$); WAIS-III ($\beta=-.17^*$) Block 2: ($R^2=.20^{**}$); History of substance abuse disorder (SCID) ($\beta=.02$); History of depressive disorder ($\beta=.13$); Current depressive disorder ($\beta=.29^{**}$) Step 3 ($R^2=.03^*$): NAI ($\beta=.22^{**}$);	13/16 PSI summary scores (medium-to-large, positive correlations, highly significant)	0.1 < r <.5	Yes	+
Holden et al., 1989	Parental Stress Index (PSI)	Dissimilar but related	13/16 PSI summary scores (medium-to-large, positive correlations, highly significant)	0.1 < r <.5	Yes	+
Kilpatrik, 2005	PEM - signals	Dissimilar but related	(-.22*)	0.1 < r <.5	Yes	+
Parental Empathy Measure (PEM); semi-structured interview, which via scenarios, aims to assess how parents would respond emotionally, behaviourally, explain why child is behaving in a certain way and their ability to detect child signals and their attributions about the child's behaviours	PEM - attributions	Dissimilar but related	(-.41**)	0.1 < r <.5	Yes	+
	PEM - Emotion	Dissimilar but related	(-.52**)	0.1 < r <.5	Yes	+
	PEM - Behaviour	Dissimilar but related	(-.41**)	0.1 < r <.5	Yes	+
	PEM Total	Dissimilar but related	(-.48**)	0.1 < r <.5	Yes	+
Plant et al., 2016	Social satisfaction	Dissimilar but related	(-.65***)	0.1 < r <.5	Yes	-

	Safety and control satisfaction	Dissimilar but related	(-.50 ^{***})	0.1 < r <.5	Yes	+
Rinehart et al., 2005	GSI	Dissimilar but related	(.56 ^{**})	0.1 < r <.5	No	-
	ASI, Alcohol composite score	Dissimilar but related	(.11)	0.1 < r <.5	Yes	+
	ASI, Drug composite score	Dissimilar but related	(.09)	0.1 < r <.5	Yes	+
Lifetime frequency of childhood abuse (FCA)	FCA	Dissimilar but related	(.17 ^{**})	0.1 < r <.5	Yes	+
Current exposure to interpersonal abuse (CEIA)	CEIA	Dissimilar but related	(.15 ^{**})	0.1 < r <.5	Yes	+
	Hierarchical multiple regressions (for each sample)	Model to be significant; Exploratory, no H concerning individual predictors	Mental health (GSI) only significant predictor of CAPI Abuse ($\beta=.51^{***}$) when analysed together with child and adult trauma (FCA: $\beta=-.03$; CEIA: $\beta = .02$), alcohol ($\beta=.06$) and drug scores ($\beta=-.02$) in both samples			?
Rodriguez & Silvia, 2022 (AIMS-P: high-risk sample; Triple-F: community sample)	Parent-CAAM - AIMS-P	Dissimilar but related	(.19)	0.1 < r <.5	Yes	+
Parent-child aggression acceptability movie task (Parent-CAAM); eight-90s clips designed to assess approval of parent child abuse; higher scores indicate greater parent child aggression approval	Parent-CAAM - Triple-F	Dissimilar but related	(.14)	0.1 < r <.5	Yes	+

Frustration Intolerance (FIT); analogue task designed to assess frustration tolerance in parent-relevant scenarios; number of seconds till quitting is captured with longer duration suggesting greater frustration tolerance	FIT - AIMS-P	Dissimilar but related	(-.36*)	0.1 < r <.5	Yes	+
	FIT - Triple-F	Dissimilar but related	(.01)	0.1 < r <.5	No	-
Noncompliance Implicit Association Test (N-IAT); implicit measure in which respondents sort terms describing child behaviour into good/bad, obey/disobey	N-IAT - AIMS-P	Dissimilar but related	(-.08)	0.1 < r <.5	No	-
	N-IAT - Triple-F	Dissimilar but related	(-.24*)	0.1 < r <.5	Yes	+
Response Analogue to Child Compliance Task (ReACCT); analogue task intended to simulate realistic parent-child interchange. Higher scores indicative of greater child abuse risk	ReACCT - AIMS-P	Dissimilar but related	(.27)	0.1 < r <.5	Yes	+
	ReACCT - Triple-F	Dissimilar but related	(.10)	0.1 < r <.5	Yes	+
Sprang et al., 2005						
Variable coded into extreme, severe, moderate and mild, according to nature of maltreatment, the act, extent of child's injury, the behaviour of the offending parents	Maltreatment severity	Similar	(.28**)	r>.5	No	-

according to US state CPS rating anchors						
Urgelles et al., 2012	Average number of emergencies reported during EPM with pre-treatment CAPI Abuse scores	Dissimilar but related	(.40*)	0.1 < r <.5	Yes	+

Note. * $p < .05$; ** $p < .01$; *** $p < .001$; r coefficients reported to two decimal places; n.s. = non-significant; H = included study authors' Hypothesis
^a? = indeterminate was assigned to multivariate findings in which the study authors did not provide hypotheses a priori but were conducting exploratory analyses to see which variable would be most important in explaining CAPI Abuse scale scores;
 β = standardised regression coefficient which allows direct comparison between variables entered in multivariate analysis to determine which has the most influence on the dependent variable (Field, 2009).

Table D2

Additional Analysis: CAPI Abuse scale criterion and known-group validity by study

Study	Characteristic	Result	Difference significant? ($p < .05$)	Effect size	Psychometric Quality rating
Caliso & Milner, 1992	Abusive parents vs comparison; Parent childhood history of abuse vs no childhood history of abuse	Abusers w history of abuse > comparison w history of abuse ($f = .36$); abusers w history of abuse > comparison w/o history of abuse ($f = .54$); comparison w history of abuse > comparison w/o history ($f = .18$)	yes	Medium; medium; small	+ (3)
Caliso & Milner, 1994	Abusive parents vs comparison; Parent childhood history of abuse vs no childhood history of abuse	Abusers w childhood history of abuse > non-abusers w childhood history of abuse; abusers w childhood history of abuse > non-abusers w/o childhood history of abuse; non-abusers w childhood history of abuse > non-abusers w/o childhood history of abuse	yes	Medium; medium; small	+ (3)
Craig & Sprang, 2007	Gender	Women > men (strength of relationship with gender weak; partial eta-squared = .01)	Yes	Small	?
	Trauma exposure and age of trauma exposure	Childhood trauma > no trauma (medium, $f = .23^*$); adult trauma > no trauma (small, $f = .19^*$); child and adult trauma > no trauma (medium, $f = .29^*$)	yes	Medium; small; medium	+ (3)

	Trauma exposure and age of trauma exposure	Adult trauma, child trauma, adult and child trauma	No	Small	-
Haapsalo & Altonen, 1999	CPS status	CPS group > control	yes	Large	+
Hien et al., 2010	Parental substance abuse, parental depression	Substance use Disorder (SUD) > comparison; Depressed > comparison;	Yes	Medium; medium	+ (2)
	Parental substance abuse, parental depression	Depressed vs SUD	No	Small	?
Holden et al., 1989	Maltreating status	Neglecting parents < abusing parents; neglecting parents < spouses of abusers and neglecters; neglecting parents < parents referred for reasons other than confirmed neglect	Yes	Medium; medium; large	+ (3)
		Abusers, spouses of abusers and parents referred for reasons other than neglect	No	Small	?
	Gender	Female > male	Yes	Small	?
Rodriguez & Silva, 2022	AIMS-P: CPS sample Triple-F: Community sample	AIMS-P > Triple-F	Yes	Large	+

Appendix E

Step 2 – Results for convergent validity

Table E1

Summary of convergent validity findings per study (summary of findings for step 1 & 2)

Study (Instrument)	n	Method quality	Rating
Bradshaw et al., 2011 (PSYS)	82	Adequate	+ ¹⁰
Budd et al., 2000 (WRAT-R arithmetic, reading standard, and combined score)	75	Adequate	+ ³
Budd et al., 2000 (SCL-90, GSI T score)	75	Very good	+
Budd et al., 2000 (POQ)	75	Very good	+
Budd et al., 2000 (HOME inventory)	75	Adequate	+
Budd et al., 2000 (ASSIS Total positive support network; support satisfaction)	75	Adequate	+ ²
Budd et al., 2000 Years of education	75	Very good	+
Budd et al., 2000 Mothers age	75	Very good	+
Budd et al., 2000 (High CAP ≥ 215 vs Normal CAP score < 215); completed school grade; reading level; emotional distress; average support satisfaction	75	Adequate	+ ⁴
Budd et al., 2000 WRAT-R; GSI T; POQ; HOME; years of education	75	Adequate	?
Budd et al., 2000 WRAT-R; GSI T; POQ; HOME; years of Education (controlled for distress items in CAPI)	75	Adequate	?
Budd et al., 2006 (PSI-SF at T2)	49	Adequate	+
Budd et al., 2006 (SCL-90, GSI T)	49	Very good	+
Budd et al., 2006 (POQ total score)	49	Very good	+
Budd et al., 2006 (Home inventory)	49	Adequate	+
Budd et al., 2006	49	Adequate	+ ²

(ASSIS Total positive support network; support satisfaction)			
Caliso & Milner, 1992 (CTS - Reasoning)	90	Inadequate	+
Caliso & Milner, 1992 (CTS - Verbal abuse)	90	Very good	+
Caliso & Milner, 1992 (CTS - Violence)	90	Very good	+
Caliso & Milner, 1994 (CSNQ: father-, mother-, sibling-, friend-outsider-support; father-, mother-, sibling-punishment) (SWR)	78	Doubtful	?
Craig & Sprang, 2007 Demographic variables (Age, Race, Gender)	1680	Adequate	+ ³
Craig & Sprang, 2007 (Child physical and sexual abuse)	1680	Adequate	+ ²
Craig & Sprang, 2007 (Adult physical and sexual abuse, Domestic violence)	1680	Adequate	+ ³
Craig & Sprang, 2007 (Disaster, Death of loved one, Motor vehicle accident)	1680	Adequate	+ ³
Craig & Sprang, 2007 (Age, Race, Gender, CPA, CSA, APA, ASA, DV, Disaster, MVA, DOL) (HMR)	1680	Adequate	+ -
Craig & Sprang, 2007 (Child only trauma, no trauma, Caucasian, non-Caucasian, Women, Men, Age Group) (LR)	1680	Adequate	+ ³
Donohue et al., 2016 (Home appearance Factor)	18	Adequate	+
Donohue et al., 2016 (Home safety factor)		Adequate	+
Donohue et al., 2017 (Child neglect incompatible parenting actions)	18	Inadequate	+
Donohue et al., 2017 (Neglect incompatible actions occurring infrequently; Positive parenting goals set at treatment initiation) (MR)	18	Inadequate	+ ²

Donohue et al., 2019 (PR Drug use; PR Marijuana use)	80	Adequate	+ ²
Donohue et al., 2019 (SO Hard drug use; SO Marijuana use)		Adequate	+ ²
Donohue et al., 2019 (Control variables; PR Hard drug use; PR Marijuana use) (HMR)	80	Adequate	+
Donohue et al., 2019 (Control variables; SO reported hard drug use; SO reported Marijuana Use) (HMR)	80	Adequate	?
Haapsalo & Altonen, 1999 (CPA)	50	Adequate	+
Haapsalo & Altonen, 1999 (CPsyA)	50	Inadequate	+
Haapsalo & Altonen, 1999 (SES)	50	Adequate	+
Haapsalo & Altonen, 1999 (Social problems)	50	Inadequate	+
Haapsalo & Altonen, 1999 (CPA; CPsyA; Social Problems; SES) (MR)	50	Inadequate	+ -
Haskett et al., 1995 (SCL-90-R, GSI T score)	41	Very good	-
Haskett et al., 1995 (AAPI)	41	Adequate	-
Haskett et al., 1995 (CBCL Internalizing, Externalizing)	41	Adequate	+ ²
Haskett et al., 1995 (CBCL Total child problems)	41	Adequate	-
Haskett et al., 1995 (Child adjustment: Internalizing, Externalizing, Total problems - Teacher report form)	41	Adequate	+ ³
Haskett et al., 1995 (Observed parenting Behaviour)	41	Very good	+
Haskett et al., 1995 (Parental problem solving questionnaire – mean number of	41	Very good	-

solutions)			
Haskett et al., 1995 (Parental problem solving questionnaire – mean effectiveness)	41	Very good	+
Haskett et al., 1995 (High CAP ≥ 215 vs Normal CAP < 215); observed parenting index; GSI; CBCL (externalizing, internalizing, total problem behaviour); TRF (externalizing, internalizing, total problem); AAPI; Parental problem solving (mean effectiveness, mean number of solutions)	41	Adequate	+ ⁸ - ³
Hien et al., 2010 (Age, Divorced marital status, SES, WAIS-III, SCID, History of depressive disorder, history of SUD, Current depressive disorder, NAI) (MR)	152	Inadequate	+
Holden et al., 1989 (PSI)	87	Adequate	+
Kilpatrik, 2005 (PEM)	103	Very good	+ ⁵
Plant et al., 2016 (Social satisfaction)	72	Very good	-
Plant et al., 2016 (Safety and control satisfaction)	72	Very good	+
Rinehart et al., 2005 (GSI)	536	Very good	-
Rinehart et al., 2005 (ASI, Alcohol, and drug composite score)	536	Adequate	+ ²
Rinehart et al., 2005 (FCA)	536	Very good	+
Rinehart et al., 2005 (CEIA)	536	Very good	+
Rinehart et al., 2005 (GSI, FCA, CEIA, ASI Alcohol, ASI Drug) (MR)	536	Very good	?
Rodriguez & Silvia, 2022 (Parent-CAAM) AIMS-P	38	Very good	+
Rodriguez & Silvia, 2022 (Parent-CAAM) TRIPLE-F	76	Very good	+
Rodriguez & Silvia, 2022 FIT (AIMS-P)	38	Adequate	+

Rodriguez & Silvia, 2022 FIT (TRIPLE-F)	76	Adequate	-
Rodriguez & Silvia, 2022 N-IAT (AIMS-P)	38	Adequate	-
Rodriguez & Silvia, 2022 N-IAT (TRIPLE-F)	78	Adequate	+
Rodriguez & Silvia, 2022 (ReACCT) AIMS-P	38	Very good	+
Rodriguez & Silvia, 2022 (ReACCT) TRIPLE-F	76	Very good	+
Sprang et al., 2005 (Maltreatment severity)	208	Very good	-
Urgelles et al., 2012 (Average no. of emergencies reported during EPM)	26	Adequate	+

Note: +ⁿ: n refers to the number of the measure's subscales that were or were not in line with hypothesis; 'Study' in the title refers to single analyses that were reported in each article; **PSYS**: Parent Satisfaction with Youth Scale; **WRAT-R**: Wide Range Achievement Test; **SCL-90 R**; **GSI**: Symptom Checklist 90-Revised; Global Severity Index; **POQ**: Parent Opinion Questionnaire (unrealistic parenting beliefs); **HOME Inventory**: Home Observation for the Measurement of the Environment; **PSI-SF**: Parental Stress Index-Short Form; **ASSIS**: Arizona Social Support Interview Schedule; **CTS**: Conflict Tactics Scale; **CSNQ**: Childhood Social Network Questionnaire; **CPsA**: Childhood Psychological Abuse; **SES**: Socioeconomic Status; **CPA**: Child physical abuse; **CSA**: Child sexual abuse; **APA**: Adult physical abuse; **ASA**: Adult sexual abuse; **DV**: Domestic Violence; **MVA**: Motor vehicle accident; **DOL**: Death of loved one; **PR Hard drug use**: Parent Reported hard drug use; **SO Hard drug use**: Significant others' report of mothers hard drug use; **AAPI**: Adult-Adolescent Parenting Inventory (physical punishment subscale); **WAIS-III**: Weschler Adult Intelligence Scale (3rd ed); **NAI**: Novaco Anger Inventory; **SCID**: Structured Clinical Interview for DSM-IV, (**SUD**) Substance Abuse Disorder section; **CBCL**: Child Behaviour Checklist; **PSI**: Parental Stress Index; **PEM**: Parental Empathy Measure; **ASI**: Addiction severity Index; **FCA**: Lifetime Frequency of Child Abuse; **CEIA**: Current Exposure to Interpersonal Abuse; **Parent-CAAM**: Parent-child aggression acceptability movie task; **FIT**: Frustration Intolerance; N-IAT: Noncompliance implicit association test; **ReACCT**: Response Analogue to Child Compliance Task; **EPM**: Emergency prevention management; **HMR**: Hierarchical Multiple Regression; **LR**: Logistic Regression; **MR**: Multiple Regression; **SWR**: Step Wise Regression

Table E2

Summary of criterion validity findings per study (summary of findings for step 1 & 2)

Study (Subgroup characteristic)	n	Methodological quality	Effect size	Rating
Caliso & Milner, 1992 Abusive vs comparison parents; history of parent childhood abuse vs no history)	90	Very good	Small-to-Medium	+ ³
Caliso & Milner, 1994 (Abusive vs comparison parents; history of parent childhood abuse vs no history)	78	Adequate	Small-to-Medium	+ ³
Haapsalo & Altonen, 1999 (CPS vs control)	47	Very good	Large	+
Holden et al., 1989 (Maltreatment status)	87	Very good	Small-to-Large	+ ³ ?
Rodriguez & Silvia, 2022 (CPS status vs community sample)	114	Very good	Large	+

Table E3

Summary of known-group validity findings per study (summary of findings for step 1 & 2)

Study (Subgroup characteristic)	n	Methodological quality	Effect size	Rating
Craig & Sprang, 2007 Gender (Male vs Female)	1680	Very good	Small	?
Craig & Sprang, 2007 (Trauma exposure & Age of exposure)	1680	Very good	Small-to-Medium	+ ³ -
Hien et al., 2010 (Parental substance abuse vs parental depression)	152	Very good	Small-to-Medium	? + ²
Holden et al., 1989 (Gender; Male vs Female)	87	Very good	Small	?

Table E4

Summary of before and after responsiveness findings per study (summary of findings for step 1 & 2)

Reference	Methodological quality	Statistical method	Result size Hedges' g (CI)	Overall rating
Dononhue et al., 2014	Very good	Effect size	.41 (-.12 .93)	+
Harnett & Dawe, 2008	Doubtful	Effect size	.81 (.01 1.62)	+
Hubel et al., 2018	Very good	Effect size	.14 (.08 .36)	-
Kolko et al., 1998	Adequate	Effect size	.22 (-.33 .78)	-
Kolko et al., 1998	Adequate	Effect size	.34 (-.32 1.0)	-
Sanders et al., 2004	Very good	Effect size	.23 (.19 1.10)	+
Sanders et al., 2004	Very good	Effect size	1.16 (.65 1.67)	+
Thomas & Zimmer-Gembeck, 2011	Doubtful	Effect size	.02 (-.42 .45)	-
Timmer et al., 2005	Doubtful	p-value	NR	?
Vorhies et al., 2009	Doubtful	Effect size	.14 (-.42 .69)	-

NR: not reported

Table E5*Summary of between intervention and control group responsiveness findings per study**(summary of findings for step 1 & 2)*

Study	Methodological quality	Statistical method	Result size Hedges' g (CI)	Overall rating
Chaffin et al., 2012	Adequate	p-value	NR	?
Donohue et al., 2014	Very good	Effect size	-.08 (-.62 .45)	-
Dononhue et al., 2014	Very good	Effect size	-.04 (-.56 .48)	-
Hubel et al., 2018	Very good	Effect size	-.78 (-1.02 -.55)	+
Hubel et al., 2018	Very good	Effect size	-1.51 (-1.77 -1.25)	+
Kolko et al., 1998	Adequate	Effect size	-.94 (-1.66 -.21)	+
Kolko et al., 1998	Adequate	Effect size	-1.12 (-1.86 -.38)	+
Kolko et al., 1998	Adequate	Effect size	-.85 (-1.62 -.08)	+
Kolko et al., 1998	Adequate	Effect size	-.92 (-1.69 -.15)	+
Sanders et al., 2004	Very good	Effect size	-.11 (-.57 .34)	-
Sanders et al., 2004	Very good	Effect size	.08 (-.37 .54)	-
Timmer et al., 2005	Doubtful	p-value	NR	?
Thomas & Zimmer-Gembeck, 2011	Adequate	Effect size	.03 (-.42 .48)	-

Appendix F

Ethics amendment forms

Ethics

UCL RESEARCH ETHICS COMMITTEE



Amendment Request Form

Please complete this form to make any amendments to an already approved study. Carefully read the information below to check that your planned changes are covered by this form. Once completed, submit your application to ethics@ucl.ac.uk for consideration by the UCL REC.

1	Ethics ID Number: 9593/002
2	Project Title: The Supporting Parents Project (SPP): a RCT and implementation process study of the Lighthouse Parenting Programme
3	Name of PI: [REDACTED]
4	Name of Researcher(s) *for student projects:
5	Faculty and Department: Division of Psychology and Language Sciences
6	Type of Research: Undergraduate <input type="checkbox"/> Staff <input checked="" type="checkbox"/> Postgraduate Research <input type="checkbox"/> Postgraduate Taught <input type="checkbox"/>
7	Date of Original Ethics Approval: 28/04/2021
8	Amendment start date: 01/04/2022
9	Has this study been amended before: Yes <input type="checkbox"/> No <input checked="" type="checkbox"/> If yes, how many amendment requests have been submitted prior to this one? <i>(Please briefly describe all previous amendments and when they were approved).</i>

10	<p>Type of Amendment: <i>(Tick all that apply)</i></p> <p>Extension to approval (for 1 year) <input checked="" type="checkbox"/></p> <p>Data management/storage, retention and destruction <input type="checkbox"/></p> <p>Research method/protocol <input checked="" type="checkbox"/></p> <p>Location of research / research site / data source <input type="checkbox"/></p> <p>Participant group <input type="checkbox"/></p> <p>Sponsorship/Collaborators <input type="checkbox"/></p> <p>Information Sheet(s)/Consent Form(s) <input type="checkbox"/></p> <p>Consent method <input type="checkbox"/></p> <p>Data collection method <input type="checkbox"/></p> <p>Publication and sharing <input type="checkbox"/></p> <p>Recruitment Documents <input type="checkbox"/></p> <p>Principal Investigator* <input type="checkbox"/></p> <p>Update to research instruments/tools <input type="checkbox"/></p> <p>Other (Please specify in section 11) <input type="checkbox"/></p> <p><i>* To Note: Additions to the research team, other than the Principal Investigator, the Student Supervisor and the Medical Supervisor, do not need to be submitted as an Amendment. An updated list can be emailed to ethics@ucl.ac.uk to keep on record.</i></p>
11	<p>Details of Amendment(s):</p> <p><i>(Describe the amendment(s) to be made to the project, in accessible language. Include any changes to be made to the data management aspects of the study. Also, indicate which sections these amendments change in your updated Ethics Application form which must be included as part of your application).</i></p> <p>1) We would like to extend the ethics approval for one year to the 30th September 2023. This is because recruitment was slightly slower than anticipated. We have now completed the recruitment and received an extension to the project and additional funding to complete the project by this date.</p> <p>2) At the follow-up interviews, we would like to collect one additional questionnaire. This is the Adverse Childhood Experiences checklist (ACE's; attached).</p>

	<p>3) We would like to notify the committee of two student projects that will be linked with the study (see attachment for details).</p>
<p>12</p>	<p>Justification:</p> <p><i>(Provide a brief explanation of why these changes are required and why they are needed now).</i></p> <p>1) The extended timeline will enable us to complete all follow-up data collection with the participants who have now all been recruited.</p> <p>2) We expect that adverse childhood experiences could be an important mediator in understanding our results. This is something that has become apparent in our baseline interviews. We did not collect this at baseline, but as it is a retrospective checklist, we do not anticipate it to change and would like to add it to our battery of follow-up questionnaires.</p> <p>3) The students working with us will assist in data coding. In return, they will conduct meaningful secondary analyses of baseline data to explore relationships between different variables.</p>
<p>13</p>	<p>Ethical Considerations:</p> <p><i>(Explain all new ethical issues raised by the amendment and how these will be addressed. This section must NOT be left blank).</i></p> <p>1) We do not anticipate any ethical problems with the date extension. All participants have been randomized and expect to be contacted for follow-up about 6-8 months following the treatment group allocation. The extension will enable us to follow this protocol which is what participants expect.</p> <p>2) The ACE's questionnaire asks about personal and potentially very difficult past experiences. We expect these questions to resonate with many of the parents in our sample (parents with children known to child protection services, many of whom have experienced high levels of childhood trauma and abuse themselves). As with all questionnaires, the participants will be reassured that they do not need to answer anything they are not comfortable with. Furthermore, we are adding this measure to the follow-up interviews. Where possible, all data collection will be done by the same researcher who did the baseline consent and data collection with each participant. The baseline interviews, already enabled parents and the researcher allocated to them to build a rapport and trust, and we hope this will make participants feel comfortable answering these questions. They will also have the option to complete the checklist on their own if they do not want to talk about any of the experiences on the list. The researchers will have a consultation with the clinical lead of the project prior to</p>

	<p>follow-up data collection to discuss the approach to asking these questions sensitively and how to respond when parents are upset.</p> <p>3) The students will be part of the research team and participant consent has been given for the team to process data. The students will only access data from participants who have given additional consent for their pseudonomised data to be archived and used by other authorised researchers for further research (clause 8 of the consent form). The students will adhere to all existing procedures relating to consent, confidentiality and ethical conduct as reviewed and approved by the committee. Every student involved in the project will have an honorary contract with the Anna Freud Centre (AFC) and their access to any necessary data will be reviewed and approved by the AFC data protection officer following a thorough amendment of the DPIA.</p>
<p>14</p>	<p>Attachments:</p> <p><i>(List which attachments have been included. To Note: ALL Amendment Requests must be accompanied by an updated and highlighted version of your latest Ethics Application and supporting documentation that include all previously approved amendments, as appropriate, except for solely extension requests).</i></p> <p>ACE's.</p>
<p>15</p>	<p>Declaration:</p> <ul style="list-style-type: none"> ○ I confirm that the information in this form is accurate to the best of my knowledge and I take full responsibility for it. ○ I confirm that this amendment does not fundamentally change the study. ○ I confirm that all relevant data protection arrangements are still in place for the duration of this amendment. ○ I consider that it would be reasonable for the proposed amendments to be implemented. <p>Principal Investigator Name*: Dr Michelle Sleed</p> <p>Principal Investigator Signature:</p>

Date: 15/03/2022

** To Note: The named Principal Investigator must sign this form. Applications submitted without this section having been completed by the PI will be returned to the applicant.*

Last updated February 2021

RHUL self-certification

Ethics Review Details

You have chosen to self certify your project.	
Name:	Ponticelli, Sharim (2020)
Email:	[REDACTED]
Title of research project or grant:	Predictors of child abuse potential in a sample of high-risk parents: The role of epistemic trust, parenting stress and parental representations
Project type:	Royal Holloway postgraduate research project/grant
Department:	Psychology
Academic supervisor:	[REDACTED]
Email address of Academic Supervisor:	[REDACTED]
Funding Body Category:	No external funder
Funding Body:	
Start date:	30/05/2022
End date:	08/06/2023

Research question summary:

Though parental stress, parental representations of the child, and epistemic trust, theoretically, represent valid targets to reduce physical child maltreatment (Sleed et al., 2019), to date, empirical investigations on whether these factors represent pathways to child maltreatment in high-risk parent groups are lacking. This proposed study would thus test whether the following constructs individually predict an increased risk of parental child physical abuse as measured by the Child Abuse Potential Inventory (CAPI; Milner, 1994): 1) parental epistemic trust (i.e., treating others with deep suspicion and struggling to internalise new knowledge) as measured with Epistemic Trust, Mistrust and Credulity Questionnaire (ETMCQ; Campbell et al., 2021); 2) parents' hostile misattributions of their child, as measured by the Assessment of Representational Risk (ARR; Sleed et al., 2021); 3) parental stress as measured by the Parental Stress Index (PSI-4; Abidin, 1995). Second, this study would aim to test whether parental epistemic trust, parental representational risk, and parental stress will be associated with increased risk of parental child physical abuse, respectively, when demographic, socio-economic factors (i.e., parent age, gender, partnership status, income) and child-related factors (e.g., age, gender, adverse child childhood experiences as measured by the Adverse Childhood Experience (ACEs) Questionnaire and behaviour as measured by the Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001)) are controlled for.

Research method summary:

The proposed project aims to carry out secondary data analysis on the baseline data that has already been collected from a sample size of 110 participants as part of The Supporting Parents Project (SPP). The SPP is a Randomised Controlled Trial (RCT) which aims to evaluate the effectiveness and feasibility/acceptability of the Lighthouse Parenting Programme (LPP) in Children's Social Care. The LPP is a mentalization-based intervention for parents with children known to child protection services which aims to reduce parental child maltreatment through targeting parent mentalizing, epistemic trust and parental stress. The proposed study would be using a cross-sectional correlational survey design on baseline data that will be collected as part of the SPP prior to the LPP intervention. The proposed study would use

As part of this project, in addition to conducting secondary data analysis as an honorary research officer I would also be involved in collecting the baseline and outcome data of three SPP participants and assessing for features of 'representational risk' by using the ARR coding schedule on approximately 30 baseline and post-intervention interview transcripts, obtained using the Parent Developmental Interview (PDI; Slade et al., 2004). Access to a population of high-risk parents has already been agreed as part of the SPP ethics application, which was reviewed and approved by the University College London research ethics committee (Project ID Number: 9593/002).

My proposed study here has also already been approved by the UCL ethics research committee as as part of an amendment of the original SPP study. Please refer to form attached.

Working with participants that are 'at risk'

Will the research involve any of the following 'at risk' participants?

Children (under the age of 16),

No

Participants with cognitive or physical impairment that may render them unable to give informed consent,
No

Participants who may be at risk due to personal, emotional, psychological or other reasons,
No

Participants who may become at risk as a result of the conduct of the study (e.g. because it raises sensitive issues) or as a result of what is revealed in the study (e.g. criminal behaviour, or behaviour which is culturally or socially questionable),
No

Participants who are in unequal power relations (e.g. groups that you teach or work with, in which participants may feel coerced or unable to withdraw),
No

Participants who may potentially suffer negative consequences if identified (e.g. professional censure, exposure to stigma or abuse, damage to professional or social standing),
No

Details,

Access to a population of high-risk parents has already been agreed as part of the SPP ethics application, which was reviewed and approved by the University College London research ethics committee (Project ID Number: 9593/002).

My proposed study here has also already been approved by the UCL ethics research committee as part of an amendment of the original SPP study (9593/002). Please refer to email and ethics amendment forms attached.

Other considerations when working with people and their data

Does your study include any of the following?

Will it be necessary for the research that people take part in the study without their informed consent at the time?,
No

Will the research, however briefly, be managing identifiable or special category data as defined by GDPR? (Please see the Royal Holloway's research ethics intranet page for guidance),
No

Is pain or discomfort likely to result from the study?,
No

Could the study induce psychological stress or anxiety, or cause harm or negative consequences beyond the risks encountered in normal life?,
No

Does this research involve NHS patients, staff, premises, resources, data or tissue samples?,
No

If so what is the NHS Approval number,

Are drugs, placebos or other substances to be administered to the study participants, or will the study involve invasive, intrusive or

potentially harmful procedures of any kind?.

No

Will human tissue including blood, saliva, urine, faeces, sperm or eggs be collected or used in the project?.

No

Will the research work with a dataset that requires a data sharing agreement?.

No

Will financial inducements (other than reasonable expenses and compensation for time) be offered to participants?.

No

Is there a risk that any of the material, data, or outcomes to be used in this study has been derived from ethically-unsound procedures?.

No

Details,

Risks to the Environment / Society

Will the conduct of the research pose risks to the environment, site, society, or artifacts?.

No

Will the research be undertaken on private or government property without permission?.

No

Will geological or sedimentological samples be removed without permission?.

No

Will cultural or archaeological artifacts be removed without permission?.

No

Details,

Risks to Researchers, Research Collaborators, and Royal Holloway

Does your research present any of the following risks to researchers, research collaborators, or Royal Holloway?

Is there a possibility that researchers or research collaborators could be exposed to emotional or physical risks (e.g. by being alone with vulnerable, or potentially aggressive participants, by entering an unsafe environment, by working in countries in which there is unrest, accessing archives with troubling content, or by examining material that may cause secondary trauma)?.

No

Is the topic of the research sensitive or controversial such that the researcher could be ethically or legally compromised (e.g. as a result of disclosures made during the research)?.

No

Will the research involve the investigation or observation of, proximity to, or participation in illegal practices?.

No

Could any aspects of the research mean that Royal Holloway has failed in its duties of care?.

No

Is there any reputational risk concerning the source of your funding?.

No

Is there any other ethical issue that may arise during the conduct of this study that could bring the institution into disrepute?.

No

Details,

Declaration

By submitting this form, I declare that the questions above have been answered truthfully and to the best of my knowledge and belief, and that I take full responsibility for these responses. I undertake to observe ethical principles throughout the research project and to report any changes that affect the ethics of the project to the University Research Ethics Committee for review.

Certificate produced for user ID, NJJT030

Date:	29/04/2022 19:04
Signed by:	Ponticelli, Sharim (2020)
Digital Signature:	Sharim Ponticelli
Certificate dated:	4/29/2022 8:01:00 PM
Files uploaded:	amendment_request_form_SPP_March_2022_(002).docx SPP_high_risk_ethics_application_form_FINAL_amendment_1_March_22_(003).doc FW_APPROVED_Amendment_Request_9593002.msg

RHUL email confirmation of receipt of self-certification

Good Afternoon Sharim,

I can confirm a submission id 3230 entitled "Predictors of child abuse potential in a sample of high-risk parents: The role of epistemic trust, parenting stress and parental representations" was self-certified on the 28th of April 2022.

Kind Regards,



Appendix G

Study outcome measures

Epistemic Trust Mistrust Credulity Questionnaire (ETMCQ; Campbell et al., 2021)

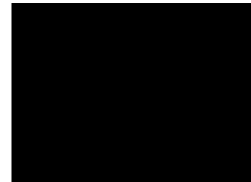
	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
1. I usually ask people for advice when I have a personal problem. (T)	1	2	3	4	5	6	7
2. I find information easier to trust and absorb when it comes from someone who knows me well. (T)	1	2	3	4	5	6	7
3. I'd prefer to find things out for myself on the internet rather than asking people for information. (M)	1	2	3	4	5	6	7
4. I often feel that people do not understand what I want and need. (M)	1	2	3	4	5	6	7
5. I am often considered naïve because I believe almost anything that people tell me. (C)	1	2	3	4	5	6	7
6. When I speak to different people, I find myself easily persuaded by what they say even if this is different from what I believed before. (C)	1	2	3	4	5	6	7
7. Sometimes, having a conversation with people who have known me for a long time helps me develop new perspectives about myself. (T)	1	2	3	4	5	6	7
8. I find it very useful to learn from what people tell me about their experiences. (T)	1	2	3	4	5	6	7
9. If you put too much faith in what people tell you, you are likely to get hurt. (M)	1	2	3	4	5	6	7
10. When someone tells me something, my immediate reaction is to wonder why they are telling me this. (M)	1	2	3	4	5	6	7
11. I have too often taken advice from the wrong people. (C)	1	2	3	4	5	6	7
12. People have told me that I am too easily influenced by others. (C)	1	2	3	4	5	6	7
13. If I don't know what to do, my first instinct is to ask someone whose opinion I value. (T)	1	2	3	4	5	6	7

14. I don't usually act on advice that I get from others even when I think it's probably sound. (M)	1	2	3	4	5	6	7
15. In the past, I have misjudged who to believe and been taken advantage of. (C)	1	2	3	4	5	6	7

Assessment of Representational Risk (ARR; Sled, Isosävi et al., 2021) Coding sheet

Interview number:		Child age:		Coder:	
	Page number	Notes	Score		
Hostility – parent's experience					
Hostile/frightening behavior					
Fearfulness					
Helplessness					
Emotional distress					
Idealization					
Enmeshment/Role reversal					
Incoherence					
Supportive presence					
Mutual enjoyment					

ARR reliability test certificate



Assessment of Representational Risk – reliability test

To Whom It May Concern:

This is to certify that

Sharim Ponticelli

Successfully completed the Reliability Test on **29th July 2022**

This certificate authorises **Sharim Ponticelli** to utilize the Assessment of Representational Risk coding system for clinical and research purposes. It also allows publication and presentation of research data obtained using the Assessment of Representational Risk coding system.

Signed:



.....
Michelle Slead
Course Leader

Our Patron: Her Royal Highness The Duchess of Cambridge The Anna Freud National Centre for Children and Families is a company limited by guarantee, company number 03819888, and a registered charity, number 1077106.

Copy right measures

Due to copy right reasons the measures listed here below are only accessible via a hyperlink (please click on measure name) to a shared access drive, which can only be opened by Royal Holloway, University London staff. These hyperlinks will be removed after the correction process to respect copy right.

Assessment of Representational Risk (Sleed, Isosävi et al., 2021)

Parental Developmental Interview, short version (PSI-S; Slade et al., 2020)

Child Abuse Potential Inventory (CAPI; Milner, 1986)

Parental Stress Index version 4 (PSI-4; Abidin, 1995)

Adverse Childhood Experiences (ACEs) questionnaire

ADVERSE CHILDHOOD EXPERINECES QUESTIONNAIRE

Our experiences in childhood can have a negative impact on our health and well-being as adults. We think it will be helpful for this project to know about what you experienced while growing up. We know that our experiences in childhood are only part of the story. Resilience - the ability to 'bounce back' - is just as important as adversity. Resilience can grow with positive experiences and relationships.

Below is a list of Adverse Childhood Experiences (ACEs).

Next to each question, please choose either YES or NO.

Please answer about things that happened before your 18th birthday.

1. Did a parent or adult in your home ever swear at you, insult you, or put you down?

- Yes
- No

2. Did a parent or adult in your home ever hit, beat, kick, or physically hurt you in any way?

- Yes
- No

3. Did you experience unwanted sexual contact (such as fondling or oral/anal/vaginal intercourse/penetration)?

- Yes
- No

4. Did you feel that no one in your family loved you or thought you were special?

- Yes
- No

5. Did you feel that you didn't have enough to eat, had to wear dirty clothes, or had no one to protect or take care of you?

- Yes
- No

6. Did you lose a parent through divorce, abandonment, death, or other reason?

- Yes
- No

7. Did your parents or adults in your home ever hit, punch, beat, or threaten to harm each other?

- Yes
- No

8. Did you live with anyone who had a problem with drinking or using drugs, including prescription drugs?

- Yes
- No

9. Did you live with anyone who was depressed, mentally ill, or attempted suicide?

- Yes
- No

10. Did you live with anyone who went to jail or prison?

- Yes
- No

11. Are there any comments you want to make about the answers you have given us above?

Signature:

Date:

Appendix H

SPP participant Information sheet and consent form

Supporting Parents Project

Information Sheet: Key Points

We would like to invite you to participate in this study. This first page gives you a summary of what the study is about. The pages that follow give you more detailed information.

- We are doing this study to try to find out more about the best ways that children's social care can support parents.
- This study will be testing how helpful different types of support are, including a new programme called The Lighthouse Parenting Programme.
- To do this, we will be looking at how things go for some parents who will join the Lighthouse programme and some parents who receive the usual support offered in children's social care services.
- Half the people taking part in the study will receive the Lighthouse programme and the other half will receive usual care. We do not choose who goes into which group- this will be decided at random by a computer.
- All parents will get the usual support offered by their children's social care service. The parents in the Lighthouse Programme will also attend a weekly Parents' Group and one-to-one sessions with an individual therapist every two weeks for 6-7 months. These meetings will be online.
- All parents who take part will have a few online meetings with our friendly researchers. There will be one or two meetings at the beginning (before we know which group they are in) and one or two meetings after 6-7 months (at the end of the Lighthouse/ Usual Support programme). During these meetings, you will go through some questionnaires with the researcher. You will have a chance to talk about your experiences of being supported.
- All parents who take part will be paid a £25 voucher for each of the two stages of the research as a thank-you for their time (so £50 in vouchers for taking part in the whole study).
- Your participation is completely voluntary, and you can exit the programme and the study whenever you want.

Now here is the more detailed explanation – it isn't just 'the fine print' – it contains some important information and if you read it you'll get a much clearer understanding of the study and what taking part would involve.

Participant Information Sheet

You are being invited to take part in the Supporting Parents Project (SPP). Before you decide, it is important for you to understand why the research is being done and what participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Participation is voluntary, please take time to decide whether or not you wish to take part. Thank you for reading this.

What is the purpose of the project?

The purpose of the project is to find out more about the best ways that children's social care services can support parents. There are many ways that parents can be supported and we need to learn more about what works best and for which families. One of the services we would like to learn more about is the Lighthouse Parenting Programme. We will look at how helpful the Lighthouse programme is compared with the usual services provided in children's social care for parents and their children.

This study is a randomised controlled trial. This means that parents will receive EITHER usual care alone OR usual care + the Lighthouse Parenting Programme. Overall, half of the parents will be in the first group and the other half in the second group. We do not choose who goes into which group- this is done by a computer and there is an equal chance of being in either group.

Why have I been chosen?

Your children's service is in one of several Local Authorities in England where the study will run. Up to 140 parents will take part in the study. Parents are being invited to take part if they have at least one child aged 0-12 years who is currently on either a Child in Need Plan, a Child Protection Plan, in pre-proceedings or a supervision order.

Do I have to take part?

No, it is up to you to decide whether or not to take part. This will not affect the support that you and your family would normally receive. If you decide to take part, you can withdraw at any time without giving a reason and without it affecting any benefits or support that you are entitled to. If you decide to withdraw you will be asked what you wish to happen to the information you have provided up to that point.

What will happen to me if I take part?

A researcher will contact you to discuss the study further and answer any of your questions. If you do decide to take part in the study, you will sign a consent form, a copy of which you can keep with this information sheet. You will have one or two online meetings with the researcher to go through some questionnaires, including questions about your family situation, and an interview about how things are going for you as a parent. You will then be placed in one of the groups (usual care OR usual care + Lighthouse programme) as noted above.

Over the following 6-9 months, no matter which group you are in, you will continue to be supported in the usual way. This will entail support from the children's social work team and may include referrals to other services that could be helpful. If you are in the Lighthouse group, you will also be

invited to attend a weekly parenting group and have fortnightly one-to-one sessions with an individual therapist.

You will then be invited for one or two follow up meetings with the researcher again. You will be asked to complete the same questionnaires and be interviewed again about how things are going.

The research team will also collect some information from the children's social care records about how your family is doing. This will include the outcomes of any review meetings and information about the amount of support you received from the children's social care service over the study period.

Will I be recorded and how will the recorded media be used?

Research interviews may be audio-recorded so that the researcher can listen to your story and not need to write everything down. You will be asked for your permission before anything is recorded. The recordings, as well as all other data collected from you, will be stored in secure files and will not be associated with your name or any other information that could identify you. As soon as the recorded interview has been written up and made anonymous, it will be deleted. All personal information, like your names and contact details, and will be stored separately from other information and safely deleted from the records 10 years after the study has finished. All research information is stored by identification number (not your name) and is strictly confidential - only the research team will have access to the information you provide.

The audio recordings made during this research will be used only for addressing the study aims. No other use will be made of them, and no one outside the research team will be allowed access to the recordings.

What are the possible disadvantages and risks of taking part?

Some people can find it upsetting talking about their thoughts and feelings – both with their social or family worker and/or the researcher. All information gathered by the researcher during your meeting with him/her is strictly confidential, but if there is anything that you do not wish to discuss, or if you want to interrupt the interview for any reason, your researcher will talk to you about this and pause or stop if you decide to do so.

It is important to know that if you take part in the study, you will not be able to choose whether you receive usual care or usual care + Lighthouse. This is because a computer will automatically allocate you at random. In either condition, the care that you will be receiving will be in accordance with current best practice and routine procedures and will be carried out by professionals working on-site at your local service.

The Children's Social Care team will need to share some of your family details with the research team, which may include sensitive information such as your child's social care status. This information will be treated as strictly confidential and will not be used in any way other than what you have agreed to.

What are the possible benefits of taking part?

Whichever care you receive, you will be offered regular support and you would be monitored closely throughout your time in the study. You might find the support you receive to be helpful in giving you different ways of thinking about and managing the demands of parenting. You might enjoy meeting

with the research team and having the opportunity to talk about the issues that are important for you and your family and the type of support that is and is not helpful.

This study will help us to learn more about the helpfulness of different parenting support services when provided by children's social care, which we hope in the future will lead to better service provision for other parents who are struggling with similar difficulties. You might feel good to know that you have been part of an important study and contributed to this.

You will also be offered vouchers for your time given to take part in the two research interviews (£25 voucher for each time point, up to a value of £50 altogether).

Has an Ethics Committee has checked the research project?

All research projects are looked at by an independent group of people, called a Research Ethics Committee, to protect your rights. This research has been reviewed and agreed by the UCL Research Ethics Committee (Project ID Number: 9593/002).

What if something goes wrong?

If you have any complaints about your care provided within the project, these can be addressed to your key worker in the social care team. Compensation for any injury caused by the management or conduct of treatment within this study will be in accordance with policy of the local authority from where you have been referred. You can ask for further information from a member of your social work team. If you have any complaints about the research part of the study, then please contact the lead researcher whose details are at the bottom of this letter. If you feel your complaint has not been handled to your satisfaction, you can contact the Chair of the UCL Research Ethics Committee –

██████████

Will my taking part in this project be kept confidential?

Your social care team who referred you to the study and will be notified if you agree to take part. They will also know which group you are allocated to. This is because the support being offered in the study will be provided by the social care team. All the information that we collect about you as part of the research will be kept strictly confidential and will only be available to the research team. This data will be stored for the duration of the trial and then for 10 years after, in accordance with usual policy. You will not be identifiable in any reports or publications about the study.

Data from this study will be archived with the ONS Secure Research Service. It will not use your name to identify you, rather an identification number. This personal data would be stored indefinitely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.

Limits to confidentiality

Please note that confidentiality will be maintained as far as it is possible, unless the researcher felt that you or another person were in danger of harm. In this case, they may need to inform relevant agencies of this. Wherever possible, this would be discussed with you beforehand.

What will happen to the results of the research project?

At the end of the study, reports will be written about the findings. The results within these reports will be presented in such a way that no one can identify you or know that you took part. Findings will be presented in terms of what we observe among the whole group rather than individuals. For example, the report might note that 60% of people in the study held a certain opinion.

It is expected that the results of the research will be available as published articles approximately 18 to 24 months after the start of the study. Please contact the principal researcher named at the bottom of this information sheet to obtain a copy.

All identifiable information held by the SPP team will be held safely at the Anna Freud Centre for 10 years. Data identified only by an ID number will be archived by the ONS Secure Research Service indefinitely and may be used for additional research.

Local Data Protection Privacy Notice

The Anna Freud Centre, also known as the Anna Freud National Centre for Children and Families, is the joint data Controller for the data processing in this research study, along with our funder, What Works in Children's Social Care. The Anna Freud Centre Data Protection Officer provides oversight of AFC activities involving the processing of personal data, and can be contacted at [REDACTED]

Further information is provided on the privacy notice you will be provided with along with this participant information sheet.

The categories of personal data used will be as follows: *name, address, phone number, email address, age, ethnicity, health data, family circumstances, social care record data.*

Our lawful bases for processing this data are specified in the General Data Protection Regulation: we rely upon Article 6(1)(e), we are conducting a task in the public interest, specifically research to evaluate the LPP intervention. As we collect Special Category data we also rely upon Article 9(2)(j) and Schedule 1(Part 1)(4) DPA (2018) and ensure our processing includes suitable and specific measures to safeguard your fundamental rights and interests while we process your data for research purposes.

Your data will be held securely by AFC for up to 10 years until the end of the research study, after which it will be anonymised and will endeavour to minimise the processing of personal data wherever possible. Then we review it before secure disposal.

If you are concerned about how your personal data is being processed, or if you would like to contact us about your rights, please contact AFC in the first instance at [REDACTED]

The full privacy notice is available here: [REDACTED]

Who is organising and funding the research?

The SPP is funded by What Works in Children's Social Care.

What happens next?

Please feel free to discuss the information above with others, if that would help you to decide whether to take part. You can keep this information sheet to look at whenever you need to. If you decide to take part, you will need to give consent and a member from the research team will contact you to arrange a time for the interview.

Contact for further information

For further information about this study, please contact the researcher or the principal investigator, Michelle Slead.

Email: [REDACTED]

Phone: [REDACTED]

Thank you for reading this information sheet and for considering taking part in this research study.

Supporting Parents Project
Consent Form for Parent/Caregiver

Name and Contact Details of the Principal Researcher: [REDACTED]
[REDACTED]

Name and Contact Details of the AFNCCF Data Protection Officer: [REDACTED]
[REDACTED]

Thank you for considering taking part in this research. The person organising the research must explain the project to you before you agree to take part. If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

I confirm that I understand that by ticking/initialling each box below I am consenting to this element of the study. I understand that it will be assumed that unticked/initialled boxes means that I DO NOT consent to that part of the study. I understand that by not giving consent for any one element that I may be deemed ineligible for the study.

		Tick Box
1.	I confirm that I have read and understood the Information Sheet for the above study. I have had an opportunity to consider the information and what will be expected of me. I have also had the opportunity to ask questions which have been answered to my satisfaction	
2.	I understand that my participation is voluntary and that I am free to withdraw at any time without giving a reason and without the care I receive or my legal rights being affected	
3.	I understand that I will be allocated to a group that receives <u>either</u> Usual Support <u>or</u> Usual Support + the Lighthouse Parenting Programme from my children’s social care service	
4.	I understand that information held by my local authority will be recorded by the research team for the purpose of this study (social care status of the children in my household, safeguarding information and service input)	
5.	I understand that my personal information will remain confidential and that all efforts will be made to ensure I cannot be identified, unless the researcher felt that I or another person were in danger of harm. In this case, they may need to inform relevant agencies of this.	
6.	I understand that if I decide to withdraw, any personal data I have provided up to that point will be deleted unless I agree otherwise	
7.	(Optional) I consent to my research interviews being audio recorded and understand that the recordings will be stored anonymously, using password-protected software and will only be used for research purposes.	
8.	(Optional) I understand that at the end of this study other authenticated researchers will have access to my pseudonymised data once it has been archived at the ONS	

	Secure Research Service. This data would be stored indefinitely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.	
--	--	--

Name of participant

Date

Signature

Researcher

Date

Signature

Appendix I

Statistical assumption testing for hierarchical linear regression

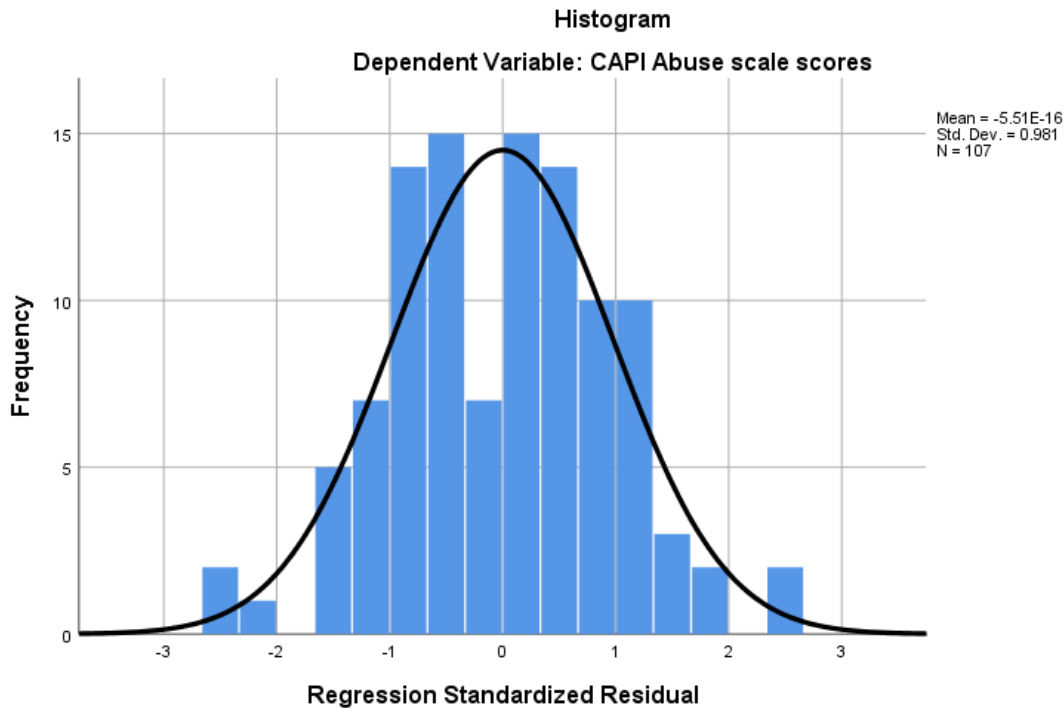
Scatter plots

CAPI Abuse scale scores were plotted against parental stress, epistemic mistrust and credulity and representational risk scores, respectively. Plots were then visually inspected to assess the spread of scores, the presence of bivariate outliers and potential relationships between the two variables. The scatter plots for each variable with CAPI Abuse scale scores revealed a good spread of scores on both variables and suggested a positive direction of effects between CAPI and each variable, respectively. The inspection revealed the presence of potential bivariate outliers, i.e., a score which may lie within normal limits for each of the variables individually, but which is highly unusual in combination. These data points were not removed since they did not lie too far away from most scores and their removal may have resulted in the loss of actual data points. Furthermore, analyses performed without the potential bivariate outliers did not reveal major differences in correlation coefficient sizes.

Normality of standardized residuals

Figure I1

Residual histogram plot for normality assumption testing



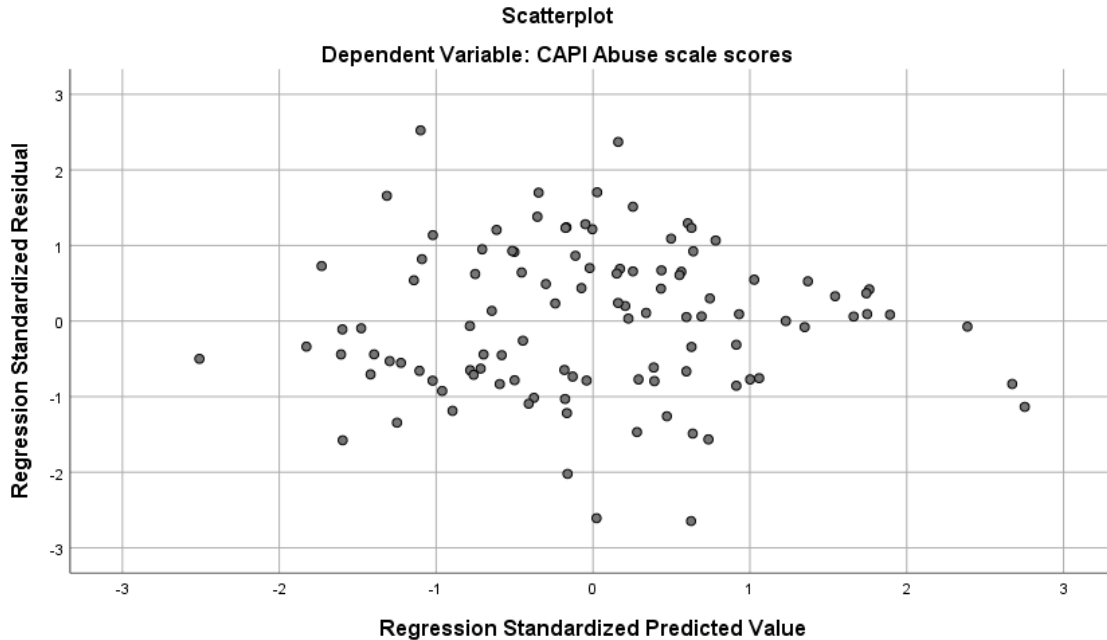
After running the primary regression model including only parental stress as the covariate variable a normality plot was calculated of the residuals for the regression of Child Abuse Potential (CAP) scores and all other variables in the model (parental stress, epistemic mistrust and credulity, and parental representations) (See Appendix figure I1). The histogram over the standardized residuals did not reveal positive or negative skewness given that neither of the tails were unequally stretched out. There may have been a small amount of kurtosis due to the middle bars being too high and piercing through the normal curve. Given that potential normality deviations were considered very small it was concluded that the *residuals were roughly normally distributed* and thus that it would be appropriate to proceed with the planned hierarchical linear regression analysis.

Homoscedasticity and linearity

Homoscedasticity and linearity were tested for by creating a scatter plot for the predicted values (x-axis) against the residuals (y-axis) (see Appendix figure I2). Visual inspection suggests that the dots seem to be slightly less dispersed vertically when moving from left to right. This indicates the presence of heteroscedasticity, suggesting a slight violation of the homoscedasticity assumption. The dots also seem to potentially follow a slight curved rather than linear pattern. Given that there were only slight violations to the homoscedasticity and linearity assumptions it was deemed appropriate to proceed with planned hierarchical analyses.

Figure I2

Residual scatter plot for homoscedasticity and linearity assumption testing



Outliers

Furthermore, potential outliers were screened for using Cook's distance which is the estimate of the influence of a data point. The highest cook's distance value was .155 which is below the recommended 1.0 cut off. Last, standardised residuals were also within the $-3 < x < +3$ minimum and maximum values.

Multicollinearity

Using SPSS, the variance inflation factor (VIF) for all variables was calculated for the full regression model (see Appendix Table I1). A convention is that a variable with a VIF greater than ten warrants further investigation for multicollinearity in the regression model, wherein two variables may be near perfect linear combinations of one another. The SPSS output table of VIF values for the full regression model included below shows that the VIF values did not exceed values greater than 1.367.

Table I1

Collinearity Statistics

Variable	Tolerance	VIF
Parental stress	1.000	1.000
Parental stress	.872	1.147
Epistemic mistrust	.738	1.355
Epistemic credulity	.872	1.209
Parental stress	.763	1.311
Epistemic mistrust	.731	1.367
Epistemic credulity	.824	1.213
Representational risk	.815	1.227

Appendix J

Additional statistical analyses

Table J1

Correlation analyses between ACEs and CAPI scores

ACE variable	<i>r</i> coefficient
Emotion abuse	.22
Physical abuse	.22
Sexual abuse	.26*
Emotional neglect	.17
Physical neglect	.02
Separated parents	.21
Family violence	.02
Alcohol and drug abuse	.14
Mental health	.06
Incarceration	-.04
ACE total	.17

Note. * $p < .05$. The only ACE variable to have a significant relationship to CAPI Abuse scale scores was Childhood sexual abuse, which showed a medium positive association.

Table J2

Hierarchical regression analysis to predict CAP including parents' childhood sexual abuse (CSA) as a covariate

<i>Variables and steps in the equation</i>	β	ΔR^2
<i>Model 1</i>		
Block 1: covariate		.379***
Parental stress	.601***	
CSA	.037	
Block 2: predictor variables		.056*
Epistemic mistrust	.152	
Epistemic credulity	.144	
Block 3: target variable		.017
Representational risk	.148	
<i>Model 2</i>		
Block 2: predictor variables		.059*
Epistemic mistrust	.194 [†]	
Representational risk	.159	
Block 3: target variable		.014
Epistemic credulity	.131	
<i>Model 3</i>		
Block 2: predictor variables		.059*
Epistemic credulity	.183 [†]	
Representational risk	.157	
Block 3: target variable		.015
Epistemic mistrust	.141	

Note. * $p < .05$; ** $p < .01$; *** $p < .001$; [†] $p < .06$; NB. Block 1 was the same across models

Assumption testing

Assumption testing was also carried out for the linear hierarchical regression model illustrated in Table J2 above which in addition to parental stress also included the ACE CSA variable as a covariate. Assumption testing showed that, by and large, the linear hierarchical regression assumptions were not violated, and thus it was reasonable to proceed with the intended statistical analysis.