

Case Study 1: An Evidence-Based Practice Review Report

How effective are parenting interventions for incarcerated parents at improving socio-emotional and behavioural outcomes of children?

Summary

The impact of parental incarceration on child outcomes is well documented, but less is known about the interventions to improve these outcomes. While studies have demonstrated some effectiveness of parental intervention on improving parent knowledge and skills, the research on the effectiveness of these interventions on direct child outcomes is less developed and clear.

A systematic review was conducted to investigate the effectiveness of parenting interventions for incarcerated parents at improving socio-emotional and behavioural outcomes of children. Five studies were identified and assessed for methodological quality, methodological relevance and relevance of study focus through a Weight of Evidence framework (Gough, 2007). Intervention programmes used in the studies focussed on parenting skills across a range of situations, including behaviour management and play. Effect sizes of the studies ranged from small to large, suggesting some evidence for these interventions. However, the review also highlighted some methodological issues, which take away from the evidence. Limitations and recommendations for further research are discussed to advance knowledge in this area.

Introduction

Children with Incarcerated Parents

More than four decades ago, Bakker, Morris and Janus (1978) referred to children with incarcerated parents as 'hidden victims' of crime because their rights were often neglected when the rights of inmates were championed. For some time after, the literature continued to use this term to describe how under-researched this population was (Jardine, 2017; Miller, 2006).

The impact of incarceration on inmates' children has since received a lot of attention and is well documented. These children have been found to be at higher risk of developing emotional and behavioural difficulties (Dallaire, 2007; Murray & Farrington, 2005). Farrington, Coid and Murray (2009) also found that children with incarcerated parents are themselves at greater risk of offending. Additionally, in a review, Purvis (2013) highlighted that this group of children have challenges with school performance and forming positive relationships.

While research on this group of children has increased over the years, so have rates of incarceration. In summarising available research, Armstrong, Egghs, Reid, Harnett and Dawe (2018) estimated that between 2-5% of the population under 18 years have a parent in prison. This highlights the growing need to provide effective support and intervention to improve the outcomes of this at-risk group.

Parenting Interventions for Incarcerated Parents

One strand of research in this area has focussed on the effectiveness of parenting programmes and reviews have reflected positive outcomes. A review of parenting programmes for incarcerated parents found that these programmes produced gains in parenting skills and knowledge (Newman, Fowler & Cashin, 2011).

More recently, Armstrong et al. (2018) conducted a meta-analysis of the effectiveness of parenting interventions for incarcerated parents. They arrived at a similar conclusion and reported a medium effect size of these interventions on parenting skills and knowledge. In addition, they found a small effect size for the interventions on the quality of the parent-child relationship. The authors looked at a third outcome – parent well-being – but found non-significant intervention effects. The reviews document the effectiveness of parenting programmes and interventions on parent outcomes for incarcerated parents.

There is a range of parenting interventions for incarcerated parents and there appears to be no programme favoured by most researchers (Armstrong et al., 2018; Troy, McPherson, Emslie & Gilchrist, 2018). Interventions are either developed specifically for the prison population (for an example, see Gonzalez, Romero & Cerbana, 2007) or are established programmes, such as the Incredible Years parent training, adapted for use for incarcerated parents (Menting, de Castro, Wijngaards-de Meij & Matthys, 2014). These programmes generally involve group-based sessions in prison settings, where participants learn about appropriate parenting skills guided by a trained facilitator. This is usually followed up by, or concurrently occur with, opportunities for more frequent interaction with their children. This is made possible by recruiting participants whose release date is close to the start of the study or negotiating with prison settings to accommodate extra or lengthier visits. The focus of these sessions are often tailored to the circumstances and needs of individual participants, and involve coaching or supervision (Landreth & Lobaugh, 1998; Menting et al., 2014).

Impact on Child Outcomes

Many parenting interventions involve increasing parental knowledge and skills within the context of improving parent-child relationships, better managing child behaviour and discipline, as well as communication. Incarcerated parents often have difficulties in these areas, possibly as a result of experiences in their years of development (Purvis, 2013). Zuckerman and Wright (2010) found that most fathers in incarceration grew up either without a positive relationship with their father or without a father altogether, and may then have greater difficulty with effective parenting subsequently. These interventions thus provide an opportunity to improve these skills.

Attachment theory, developed by John Bowlby and Mary Ainsworth, highlights the importance of developing a positive and stable relationship with a caregiver for a child's social and emotional development (Bretherton, 1992). Parenting interventions for incarcerated parents have the capacity to improve parenting knowledge and skills. Through improving the quality of the parent-child relationship, there is potential for good social and emotional development for the child. Given that social and emotional difficulties are precisely an area of need for children with incarcerated parents, these interventions represent an important opportunity to better address a critical area of development.

Children with incarcerated parents are, or have been, without a parent for a significant amount of time. It is thus more crucial that following this absence, parents are more equipped to build positive relationships with, and manage the needs of, their children. As a result of the absence, incarcerated parents are also more likely to experience helplessness and loss with regards to parenting (Arditti, Smock & Parkman, 2005). Intervention programmes help to specifically improve parenting skills and the quality of the parent-child relationship, which have been demonstrated to

moderate the association between incarceration and negative child outcomes (Armstrong et al., 2018).

There have been no reviews of the effectiveness of interventions for incarcerated parents on child outcomes. In a review of parenting programmes in prison, Purvis (2013) acknowledged the limited research on the effect of these programmes on children and suggested that some effects include better mental health and well-being as well as improved academic performance. As part of a systematic review of family support programmes in the criminal justice system, Troy et al. (2018) highlighted a few studies that measured child behaviour as an outcome. However, research related to the effectiveness of interventions for incarcerated parents at improving child outcomes is not yet conclusive and needs exploration.

Relevance to Educational Psychology

Home and school are the two settings that children spend the most time in. For an at-risk group like children with incarcerated parents, putting protective factors in place is even more important. Incarcerated parents who subsequently return to their families represent a crucial group who can support the development of these children. While there are many studies on the effectiveness of parenting programmes in prison, most of these focus on parent outcomes such as recidivism rates. In contrast, links to child outcomes in research are mostly indirect and limited to measures of parenting attitudes, knowledge and skills. Educational psychology research should work to extend this research to more draw direct links between these interventions and child outcomes.

This review thus aims to take a step in advancing the current state of knowledge. The review question asked is: How effective are parenting interventions

for incarcerated parents at improving socio-emotional and behavioural outcomes of children?

Critical Review of the Evidence Base

Literature Search

A systematic search of the literature was conducted on 1st February 2019 using the PsychINFO and ERIC databases. Table 1 shows the exact search terms that were used.

Table 1

<i>Search Terms</i>		
1		2
parent* train* OR parent* interven* OR parent* program* OR caregiver train* OR caregiver interven* OR caregiver program*	AND	jail* OR offen* OR prison* OR incarcera* OR crim* OR detain*

When trialling the search terms, it was found that including terms that specified child outcomes generated close to no results. These terms were thus removed from the initial search and studies were screened for the presence of these outcomes manually. The initial search yielded 702 studies in total. These studies were then screened using inclusion/exclusion criteria (see Table 2 for a full list of criteria) across various stages. Title screening left 80 studies. At this stage, 13 duplicates were identified and removed. The abstracts of the remaining 67 studies were then screened, leaving 11 studies to be screened at full text. A further 6 studies were excluded at this stage and 5 studies were identified to be eligible for the review. The

list of studies that were excluded following full text screening can be found in Appendix

A. Figure 1 illustrates the study selection steps.

Table 2

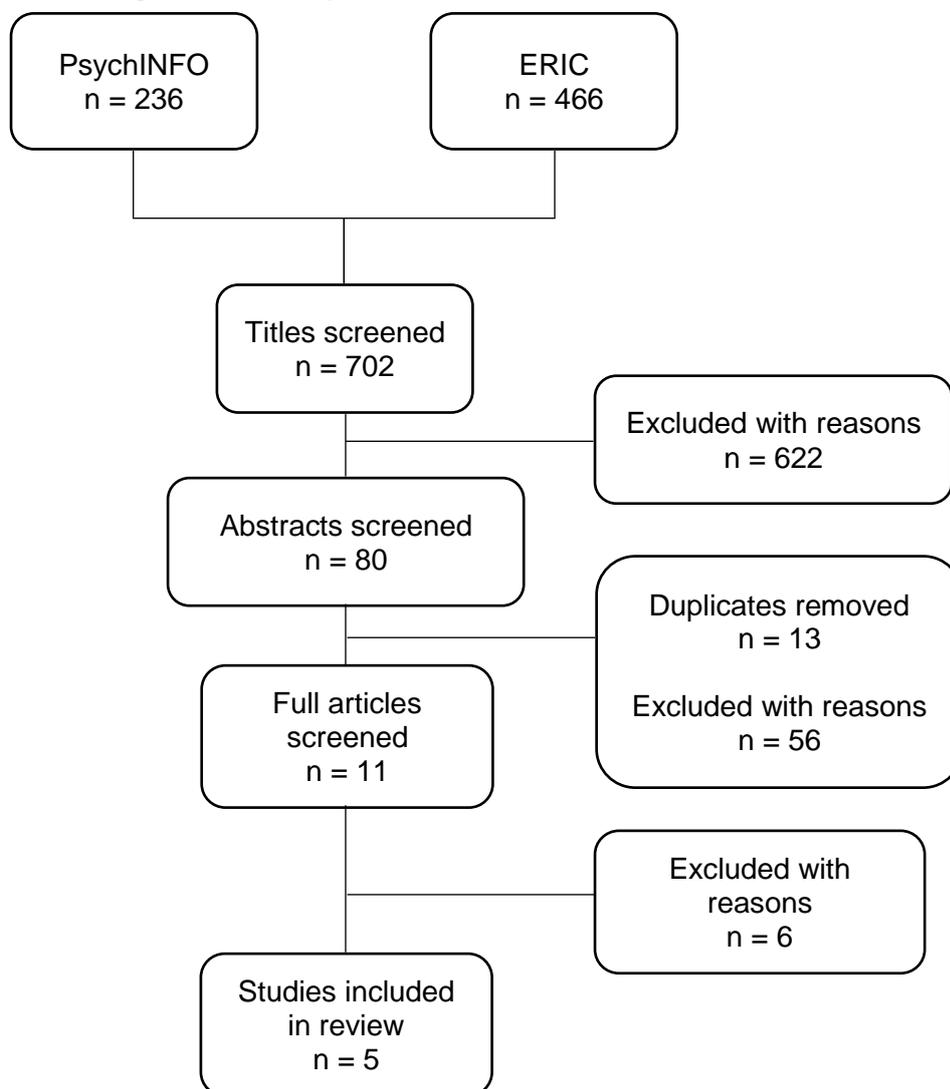
Inclusion/Exclusion Criteria

Criteria	Inclusion	Exclusion	Rationale
1. Intervention	<p>a. Study includes an active intervention programme</p> <p>b. Intervention is targeted at incarcerated parents</p>	<p>a. Study does not include an active intervention programme (e.g., explores only parental factors)</p> <p>b. Intervention is not targeted at incarcerated parents.</p>	<p>a. This review focuses on interventions/programmes.</p> <p>b. This review focuses on intervention for incarcerated parents</p>
2. Study design	<p>a. Study has pre-test and post-test measures.</p> <p>b. Study contains primary data.</p>	<p>a. Study does not have pre-test and post-test measures.</p> <p>b. Study does not contain primary data.</p>	<p>This facilitates the comparison of studies and provides a measure of the effectiveness of the intervention.</p>
3. Outcome measures	<p>Study includes a child-related social/emotional/behaviour outcome</p>	<p>Study does not include child-related outcome (e.g., outcomes include only parenting knowledge and attitudes)</p>	<p>This review focuses on the impact of parent intervention on child outcomes.</p>
4. Publication type	<p>Peer reviewed journals</p>	<p>Non-peer reviewed journals</p>	<p>There is greater confidence in the quality of studies and articles that have been through a peer review process.</p>

5. Publication language	Article published in English	Article published in a language other than English	The reviewer's first language is English
-------------------------	------------------------------	--	--

Figure 1

Flow Diagram for Study Selection



Searches for additional studies were conducted by reviewing meta-analyses and systematic reviews identified from the initial search terms. Studies identified by

these meta-analyses and systematic reviews were screened using the same inclusion/exclusion criteria and no additional relevant studies were found.

The study selection process generated five studies eligible for critical review, as seen in Table 3. A summary table reflecting the main characteristics of each study can be found in Appendix B.

Table 3

List of eligible studies

No.	Study
1	Frye, S., & Dawe, S. (2008). Interventions for women prisoners and their children in the post-release period. <i>Clinical Psychologist</i> , 12(3), 99–108.
2	Harris, Z. L., & Landreth, G. L. (1997). Filial therapy with incarcerated mothers: A five week model. <i>International Journal of Play Therapy</i> , 6(2), 53–73.
3	Harrison, K. (1997). Parental training for incarcerated fathers: Effects on attitudes, self-esteem, and children’s self-perceptions. <i>Journal of Social Psychology</i> , 137(5), 588-593.
4	Landreth, G., & Lobaugh, A. (1998). Filial therapy with incarcerated fathers: Effects on parental acceptance of child, parental stress, and child adjustment. <i>Journal of Counseling and Development</i> , 76(2), 157–165.
5	Menting, A. T. A., de Castro, B. O., Wijngaards-de Meij, L. D. N. V., & Matthys, W. (2014). A Trial of Parent Training for Mothers Being Released From Incarceration and Their Children. <i>Journal of Clinical Child and Adolescent Psychology</i> , 43(3), 381–396.

Weight of Evidence (WoE)

The five studies were evaluated based on the Weight of Evidence (WoE) framework (Gough, 2007). The WoE framework is a well-recognised framework and assesses studies on three dimensions. WoE A evaluated studies for methodological quality, WoE B evaluated the methodological relevance of the study and WoE C evaluated the relevance of the study focus to the review question. The scores of the studies, as assessed on each of the three WoE dimensions, were then averaged to

produce an overall score, WoE D. This is a measure of how much a study can contribute to answering the review question. Table 3 provides an overview of the WoE scores for the studies in this review. Detailed information about the WoE criteria and ratings can be found in Appendices C, D and E.

Table 3

Summary of Weight of Evidence ratings

Author(s)	WoE A: Methodological quality	WoE B: Methodological relevance	WoE C: Relevance of focus of study	WoE D: Overall weight of evidence
Frye & Dawe (2008)	1.5 Medium	1.5 Medium	2.33 Medium	1.78 Medium
Harris & Landreth (1997)	0.75 Low	1.5 Medium	1.33 Low	1.195 Low
Harrison (1997)	1 Low	2 Medium	2 Medium	1.67 Medium
Landreth & Lobaugh (1998)	1 Low	2.25 Medium	1.67 Medium	1.64 Medium
Menting, de Castro, Wijngaards- de Meij & Matthys (2014)	2 Medium	2.5 High	3 High	2.5 High

Note: WoE scores of 1.4 or less are classified as low, scores between 1.4 and 2.5 are classified as medium and scores or 2.5 or greater are classified as high.

Characteristics of included studies

Setting. Three studies were conducted in the United States (Landreth & Lobaugh, 1998; Harrison, 1997; Harris & Landreth, 1997), one was conducted in

Australia (Frye & Dawe, 2008) and the last study was conducted in the Netherlands (Menting et al., 2014).

Participants. Across the studies, 209 parents were involved at the start of the parenting programs. In one study (Menting et al., 2014), parents with multiple children provided ratings for up to three children. In the other four studies, parents were asked to pick a child of focus that he/she would rate. Only one study (Menting et al., 2014) had a sufficient sample size based on power calculations. As a result, it received a higher rating on WoE A, under 'Appropriate statistical analysis'.

All studies recruited participants from one gender exclusively – three studies involved only female participants (Frye & Dawe, 2008; Harris & Landreth, 1997; Menting et al., 2014), while the other two recruited only male participants (Harrison, 1997; Landreth & Lobaugh, 1998). Reasons for recruiting the specified genders differed across studies – some studies built on previous research on a specific gender, while others highlighted a need to investigate the effects of a parenting programme on an overlooked population. The reasons for gender exclusivity were assessed to be irrelevant in determining the quality of the studies.

In four of the studies (Frye & Dawe, 2008; Harris & Landreth, 1997; Landreth & Lobaugh, 1998; Menting et al., 2014), all children whose outcomes were assessed were aged between 2 and 10 years old. The average age of the children in these studies was similar too and ranged from between 5 and 7 years old. The last study, by Harrison (1997), assessed outcomes for an older group of children aged between 8 and 17 years old. This study used age-appropriate measures for children in different age groups – one was for children between 8 and 12 years old, and the other was for adolescents between 13 and 17 years old. Outcomes for adolescents were analysed separately.

Intervention. Intervention programs used across the studies varied. Two studies (Harris & Landreth, 1997; Landreth & Lobaugh, 1998) employed Filial Therapy. Through Filial Therapy training, parents learn child-centred play principles and skills to then use when interacting with their child (Landreth & Lobaugh, 1998). This aims to strengthen the parent-child relationship and enables parents to be change agents in their children's lives. Menting et al. (2014) used another well-researched programme, the Incredible Years parent training programme. Through this programme, participants in small groups learnt parenting skills, which was followed up by home visits involving individualised coaching adapted to the participants' challenges and context. In the study by Frye and Dawe (2008), the Parents Under Pressure programme was used. This is a programme designed for families with challenges such as alcohol or substance abuse, and focuses on parenting skills within that context. For example, one module focuses on helping parents regulate their emotion states without the dysfunctional use like substance use. Its effectiveness was demonstrated in a related population (Dawe & Harnett, 2007) and the programme was adapted for use for incarcerated mothers. The last study, by Harrison (1997), did not report using a manualised programme. Participants attended parent training classes in small groups conducted by the author, who was a certified parent training instructor. The classes involved learning about different developmental stages and behaviour management techniques. A variety of instruction media was used, including presentations, discussions, videos and role-playing.

Control group. Of the five studies, four involved a control group. The exception (Frye & Dawe, 2008) employed a single-group pre-post design. The lack of a control group represents a threat to validity (Barker, Pistrang & Elliot, 2016), as differences in pre-test and post-test measures may be due to maturation effects and

not the intervention. In the study conducted by Frye and Dawe (2008), a majority of the parents (9 of 12) began the intervention while still incarcerated; as such, at some point during the intervention, these parents became a much more active part of their children's everyday lives compared to when the intervention had just begun. The lack of a control group made it difficult to account for the effect of this change and resulted in lower WoE A and B ratings for this study.

All four studies that involved a control group initially assigned participants to the conditions randomly. However, in two of the studies, this was subsequently abandoned either partially (Menting et al., 2014) or completely (Harris & Landreth, 1997). In both cases, this was due to a difficulty in obtaining a sufficient number of participants in the intervention condition. Random assignment reduces statistical bias and increases confidence that any effects found can be attributed to the intervention (Flay et al., 2005). Studies that utilised random assignment received a higher rating under WoE B. One of the studies used random assignment for four of the six recruitment periods (Menting et al., 2014) and received partial credit for doing so.

Attrition. Attrition was a common issue across most studies. Harris and Landreth (1997) initially identified 51 participants for their study, but lost more than half of their participants early on in the study. As a result, the remaining participants had to be assigned to the experimental and control groups to keep them at a similar size. While the authors did not specify the reasons behind attrition, other studies experienced the same issue, albeit to a smaller extent. Of the 12 participants recruited in one study, only 8 completed the programme (Frye and Dawe, 2008). Menting et al. (2014) similarly reported a high attrition rate of just over 20% across the study. In the study by Landreth and Lobaugh (1998), very few children in the control condition completed the post-test self-concept scale; intervention effects on this measure was

thus analysed by comparing the pre-test and post-test scores of the experimental group, without comparison to the control group. While these issues were mostly out of the researchers' control, they affected the quality of the study. These studies hence received a lower rating under WoE A. To combat the attrition rate, Menting et al. (2014) conducted intention-to-treat analyses alongside per-protocol analyses. This helped to account for participants who were invited to take part but did not attend any sessions. The study hence received a higher WoE A rating.

Intervention accessibility. In one of the studies, one inclusion criterion stated that participants had to be able to read and write in English (Landreth & Lobaugh, 1998). While the necessary level of literacy was not specified, levels of literacy amongst the prison population is significantly below that of the general population, and a large group of inmates do not have functional literacy (Creese, 2016). Inclusion criteria such as these likely reduces the number of parents who can participate and benefit from intervention. There are also implications on generalisability of findings, as the participants in the study may be less representative of incarcerated parents as a whole. This study thus received a lower rating under WoE C.

Outcome measures. All studies measured outcomes related to child socio-emotional/mental health and behaviour, which was an inclusion criteria of this review. A total of six assessment tools were used to capture these outcomes, as seen in Table 4. All measures assessed were also standardised, which resulted in a high rating under WoE B.

Table 4

Measures used in studies

Study	Measure
Frye & Dawe (2008)	Strengths and Difficulties Questionnaire (SDQ)
Harris & Landreth (1997)	Filial Problem Checklist (FPC)
Harrison (1997)	Self-perception Profile for Children/Adolescents
Landreth & Lobaugh (1998)	Filial Problem Checklist (FPC) Joseph Preschool and Primary Self-concept scale (JSCS)
Menting, de Castro, Wijngaards-de Meij & Matthys (2014)	Eyberg Child Behavior Inventory (ECBI) Teacher's Report Form/Caregiver-Teacher Report Form (TRF/C-TRF)

The reliability and validity of these measures were reported in most studies and assessed to be good. An exception was the Filial Problem Checklist, which was used in two studies (Harris & Landreth, 1997; Landreth & Lobaugh, 1998), as information about reliability and validity of this measure was not available. As a result, the WoE B ratings for these two studies were lower.

As seen from Table 4, two studies used multiple sources of information. Landreth and Lobaugh (1998) assessed child outcomes from the perspectives of the parent and child, while Menting et al. (2014) assessed outcomes from the perspectives of the parent and a teacher. The use of an additional source of information increases reliability and aids in triangulation of data.

Reducing bias. All of the studies involved a measure that was completed by parents. Apart from the two studies that were highlighted in a previous section for assessing outcomes using multiple sources, only one study made attempts to reduce this bias (Harrison, 1997). Participants in this study were blind to the condition they were in and agreed to not discuss their experiences with other participants outside of

their group. The use of an alternative intervention control group also helped keep the participants blind to their condition – participants in the control group watched family-oriented videotapes and had discussions about them, but received no guidance about parenting. This study thus received a higher rating under WoE B. Notably, participants in the control groups in other studies were either on a waitlist for intervention or received support services as usual. As such, it was very likely that participants knew which group they belonged to.

Results

Table 5 summarises the effect sizes for child-related outcomes of all five studies. Only one study (Menting et al., 2014) reported effect sizes. Effect sizes were calculated for the other studies where possible. In the study by Frye and Dawe (2008), effect sizes were calculated by dividing the pre-post mean difference by the standard deviation of the pre-test scores. Two studies (Harris & Landreth, 1997; Landreth & Lobaugh, 1998) employed control groups; the effect sizes in these studies were calculated by dividing the difference in pre-post changes in the intervention and control groups by the pooled standard deviation of the pre-test scores. The exception was the JSCS measure in the Landreth and Lobaugh (1998) study; as highlighted earlier, there was an insufficient number of children in the control condition completing the post-test measure. As such, the effect size was calculated in a similar way to that done in the Frye and Dawe (2008) study. In the last study (Harrison, 1997), an effect size could not be calculated due to insufficient information. In this study, the author reported ANOVA results, which found that the intervention, time and interaction effects all did not achieve significance at the $p=.05$ level. For all other studies, the reported or calculated effect size was assigned a descriptor based on Cohen's guidance

(Cohen, 1988), where small = 0.2, medium = 0.5 and large = 0.8. It should also be noted that all effect sizes in Table 5 are presented as positive numbers to avoid confusion, though some were calculated to be negative. For example, SDQ scores were lower post-test than pre-test, reflecting an improvement in behaviour (Frye & Dawe, 2008). This resulted in a negative effect size, which was converted to a positive number.

All effect sizes measured ranged from small to large. This was true for all outcomes measured, which included intensity and frequency of problem behaviour and self-concept. The effect sizes thus reflected an improvement in child outcomes on these domains.

It was interesting to note that the study that received the highest WoE D rating (Menting et al., 2014) also produced the smallest effect sizes. Notably, the other three studies with effect sizes all presented relative methodological weaknesses, as reflected in the WoE A and WoE B ratings. These included the lack of a control group and a lack of measures to reduce bias. One of the largest effect sizes came from Harris and Landreth (1997), who received the lowest WoE D. Their study used outcome measures that did not have reliability or validity information; it also suffered from a high attrition rate. As such, the intervention effect may have been exaggerated in these studies.

Menting et al. (2014) conducted the only study that gathered information from teachers, in addition to parents. While effect sizes reflected small to medium effect sizes of intervention based on information from both sources, it should be noted that the authors reported that intervention effects for teacher rating only approached significance and were at best marginal, based on p-value ($p = .06$).

Lastly, Frye and Dawe (2008) was the only study to include a follow-up measure, conducted three months after the completion of intervention. The effect size was very similar to the one measured at post-test, which reflects that intervention effects persisted for three months without active intervention. This in turn suggests the potential for longer-term effects.

Table 5
Effect sizes

Author(s)	Number of children rated	Variable	Source of effect size	Effect size	Descriptor ^a	WoE D rating
Frye & Dawe (2008)	8	SDQ (Post)	Calculated	d = 0.69	Medium	Medium
		SDQ (3 months after invention)	Calculated	d = 0.71	Medium	
Harris & Landreth (1997)	22	FPC	Calculated	d = 1.04	Large	Low
Harrison (1997)	19	Self-perception profile	Could not be calculated (In study, author reported no significant effects for intervention, time and the interaction effect at the p=.05 level)		-	Medium
Landreth & Lobaugh (1998)	32	FPC	Calculated	d = 1.09	Large	Medium
		JSCS ^b	Calculated	d = 0.87	Large	
Menting, de Castro, Wijngaards-de Meij & Matthys (2014)	102	ECBI – intensity	Provided	d = 0.47	Small	High
ECBI – frequency	Provided	d = 0.41	Small			
(C)TRF	Provided	d = 0.62 ^c	Medium			

^a Effect size descriptors for d were taken from Cohen (1988). 0.2 = small, 0.5 = medium, 0.8 = large.

^b The JSCS was a self-concept rating by the child. In the study, some children in the control group could not be contacted. As such, the authors provided a t-statistic comparing pre-test and post-test scores in the experimental group. The effect size was calculated from this statistic.

^c While a medium effect size was reported, the authors also reflected that the p-value of the analysis was p = .06, indicating that the effect was marginal and approaching significance.

Conclusion and Recommendations

Conclusion

The purpose of this review was to assess the effectiveness of parenting interventions for incarcerated parents at improving socio-emotional and behavioural outcomes for children. This review evaluated five studies; of these, four reflected significant improvements in child outcomes. Taken together, the evidence seems to suggest that parenting interventions for incarcerated parents can be effective at improving socio-emotional and behavioural outcomes for their children. It should be noted that due to small sample sizes, all studies except Menting et al. (2014) were underpowered. As such, the results should be interpreted with caution.

Of the five studies, only one received a high WoE D rating (Menting et al., 2014). This suggests that the evidence in general lacked quality and relevance in answering this review question.

The parenting interventions used across the studies differed in content, frequency and duration of sessions as well as mode (group- or individual-based). This suggests that effective intervention for this population can take different forms. At the same time, it poses difficulties to the generalisability of findings. Specifically, the effectiveness of other parent interventions that were not part of this review in improving child outcomes is uncertain.

Recommendations

The small number of studies found in this review, as well as the range in publication years, reflect the scarcity of research in this area. A number of studies reported high attrition rates, which affected the ability to conduct statistical analyses or random assignment of participants (Harris & Landreth, 1997; Landreth & Lobaugh, 1998). The study by Menting et al. (2014) recruited participants over a three year

period partly due to the difficulty of doing so. These highlights some challenges in conducting research with this population, which may have impacted the number of studies in this area. Still, it is heartening to see that the most recent study in the review was the one that received a high WoE D rating. Going forward, more research is required in this area, and studies can use the Menting et al. (2014) study as a model.

Studies involving this population are likely to continue to have a relatively small sample. In addition to collecting quantitative data, qualitative data can serve to advance our knowledge in this area. For example, researchers can collect information on the process and facilitators of changes to child outcomes as they take place through interviews with parents, then group and analyse them thematically. Studies with a small number of participants lend themselves well to such research.

The only study that looked at outcomes for adolescents found no significant intervention effects (Harrison, 1997). This is in contrast to the results of the other studies in the review. It would be worthwhile to further explore the effectiveness (or lack of) of parenting intervention in changing adolescent outcomes. Findings can help guide consensus on whether interventions are useful for incarcerated parents of children across all ages, or if it is critical to target such interventions at parents with younger children.

Providing targeted interventions might be necessary because of the cost of these interventions. Information on cost was not readily available for any of the studies and were thus not examined in this review. However, it was noted that only one of the studies (Frye & Dawe, 2008) conducted individual intervention, while the other four involved group-based intervention. Despite the individualisation, the intervention produced only a medium effect size, which was similar or smaller than

the other effect sizes measured. This suggests that this particular intervention was not as cost-effective. This is, however, a very blunt assessment, and research on intervention cost-effectiveness might guide decision-making on the best intervention available. In the current state of research, where there is a myriad of intervention options (Troy et al., 2018), studies investigating the cost-effectiveness can prove to be useful.

Lastly, given that there is evidence of intergenerational links for crime and offending (Farrington et al., 2009), it would be worthwhile to investigate longer-term outcomes of interventions for incarcerated parents. The follow-up assessments conducted by Frye and Dawe (2008) suggest that these interventions could potentially have longer-term effects. Menting et al. (2014) also alluded to the importance of studying this and noted that repeated assessments for the participants in the study are in progress. Official data, such as arrest records, can also paint part of the picture. In time to come, this information can help shed light on the true utility of these interventions.

References

- Arditti, J. A., Smock, S. A., & Parkman, T. S. (2005). "It's been hard to be a father": A qualitative exploration of incarcerated fatherhood. *Fathering*, 3(3), 267-288.
- Armstrong, E., Eggins, E., Reid, N., Harnett, P., & Dawe, S. (2018). Parenting interventions for incarcerated parents to improve parenting knowledge and skills, parent well-being, and quality of the parent-child relationship: A systematic review and meta-analysis. *Journal of Experimental Criminology*, 14(3), 279-317.
- Bakker, L. J., Morris, B. A., & Janus, L. M. (1978). Hidden victims of crime. *Social Work*, 23(2), 143-148.
- Barker, C., Pistrang, N., & Elliott, R. (2016). Research methods in clinical psychology an introduction for students and practitioners (3rd ed.). Chichester: Wiley-Blackwell.
- Bretherton, I. (1992). The origins of attachment theory: John Bowlby and Mary Ainsworth. *Developmental psychology*, 28(5), 759.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: L. Erlbaum Associates.
- Creese, B. (2016). An assessment of the English and maths skills levels of prisoners in England. *London Review of Education*, 14(3), 13-30.
- Dallaire, D. H. (2007). Incarcerated mothers and fathers: A comparison of risks for children and families. *Family relations*, 56(5), 440-453.
- Dawe, S., & Harnett, P. (2007). Reducing potential for child abuse among methadone maintained parents: Results from a randomised control trial *Journal of Substance Abuse Treatment*, 32, 381-390.
- Farrington, D. P., Coid, J. W., & Murray, J. (2009). Family factors in the intergenerational transmission of

- offending. *Criminal Behaviour and Mental Health*, 19(2), 109-124.
- Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., ... & Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention science*, 6(3), 151-175.
- Frye, S., & Dawe, S. (2008). Interventions for women prisoners and their children in the post-release period. *Clinical Psychologist*, 12(3), 99–108.
- Gonzalez, P., Romero, T., & Cerbana, C. B. (2007). Parent education program for incarcerated mothers in Colorado. *Journal of Correctional Education*, 357-373.
- Gough, D. (2007). Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. *Research papers in education*, 22(2), 213-228.
- Guyatt, G. Sackett, M., Sinclair, J., Hayward, R., Cook, D., & Cook, R. (1995). User's guides to medical literature. A Method for Grading Health Care Recommendations. *Journal of the American Medical Association*, (2754), 1800–1804.
- Harris, Z. L., & Landreth, G. L. (1997). Filial therapy with incarcerated mothers: A five week model. *International Journal of Play Therapy*, 6(2), 53–73.
- Harrison, K. (1997). Parental training for incarcerated fathers: Effects on attitudes, self-esteem, and children's self-perceptions. *Journal of Social Psychology*, 137(5), 588-593.
- Higgins, J. P. T., Sterne, J. A. C., Savović, J., Page, M. J., Hróbjartsson, A., Boutron, I., Reeves, B., & Eldridge, S. (2016). A revised tool for assessing risk of bias in randomized trials In: Chandler J., McKenzie J., Boutron, I., Welch, V. (eds). *Cochrane Methods. Cochrane Database of Systematic Reviews* 2016, 10(1).
- Jardine, C. (2017). Constructing and maintaining family in the context of imprisonment. *The British Journal of Criminology*, 58(1), 114-131.

- Landreth, G., & Lobaugh, A. (1998). Filial therapy with incarcerated fathers: Effects on parental acceptance of child, parental stress, and child adjustment. *Journal of Counseling and Development, 76*(2), 157–165.
- Menting, A. T. A., de Castro, B. O., Wijngaards-de Meij, L. D. N. V., & Matthys, W. (2014). A Trial of Parent Training for Mothers Being Released From Incarceration and Their Children. *Journal of Clinical Child and Adolescent Psychology, 43*(3), 381–396.
- Miller, A. L., Perryman, J., Markovitz, L., Franzen, S., Cochran, S., & Brown, S. (2013). Strengthening incarcerated families: Evaluating a pilot program for children of incarcerated parents and their caregivers. *Family Relations, 62*(4), 584–596.
- Miller, K. M. (2006). The impact of parental incarceration on children: An emerging need for effective interventions. *Child and Adolescent Social Work Journal, 23*(4), 472-486.
- Murray, J., & Farrington, D. P. (2005). Parental imprisonment: effects on boys' antisocial behaviour and delinquency through the life-course. *Journal of Child Psychology and psychiatry, 46*(12), 1269-1278.
- Newman, C., Fowler, C., & Cashin, A. (2011). The development of a parenting program for incarcerated mothers in Australia: A review of prison-based parenting programs. *Contemporary Nurse, 39*(1), 2–11.
- Purvis, M. (2013). Paternal Incarceration and Parenting Programs in Prison: A Review Paper. *Psychiatry, Psychology and Law, 20*(1), 9–28.
- Troy, V., McPherson, K. E., Emslie, C., & Gilchrist, E. (2018). The Feasibility, Appropriateness, Meaningfulness, and Effectiveness of Parenting and Family Support Programs Delivered in the Criminal Justice System: A Systematic

Review. *Journal of Child and Family Studies*, 27(6), 1732–1747.

Zuckerman, S., & Wright, D. (2010). Learning how to father effectively through jail programming. *American Jails*, 26-32.

Appendix A: Studies excluded after full screening

Study	Reason for exclusion
Eddy, J. M., Martinez Jr, C. R., & Burraston, B. (2013). VI. A randomized controlled trial of a parent management training program for incarcerated parents: Proximal impacts. <i>Monographs of the Society for Research in Child Development</i> , 78(3), 75-93.	Criterion 3 – no child-related outcome.
Hansen, G. V., (2018). Does Fatherhood Training in Prison Improve Fathering Skills and Reduce Family Challenges? <i>Child Care in Practice</i> , 24(2), 198-211.	Criterion 3 – no child-related outcome.
Hayes, D., Butler, M., Devaney, J., & Percy, A. (2018). Allowing Imprisoned Fathers to Parent: Maximising the Potential Benefits of Prison based Parenting Programmes. <i>Child Care in Practice</i> , 24(2), 181-197.	Criterion 3 – no child-related outcome.
Kubiak, S. P., Kasiborski, N., & Schmittel, E. (2010). Assessing Long-Term Outcomes of an Intervention Designed for Pregnant Incarcerated Women. <i>Research on Social Work Practice</i> , 20(5), 528–535.	Criterion 3 – no child-related outcome.
Miller, A. L., Perryman, J., Markovitz, L., Franzen, S., Cochran, S., & Brown, S. (2013). Strengthening incarcerated families: Evaluating a pilot program for children of incarcerated parents and their caregivers. <i>Family Relations</i> , 62(4), 584–596.	Criterion 1b – intervention was conducted for children and caregivers who were not in prison.
Shortt, J. W., Eddy, J. M., Sheeber, L., & Davis, B. (2014). Project home: A pilot evaluation of an emotion-focused intervention for mothers reuniting with children after prison. <i>Psychological services</i> , 11(1), 1.	Criterion 3 – no child-related outcome.

Appendix B: Mapping the Field

Author(s)	Participants	Study design	Intervention	Measure(s)	Domain of child functioning assessed
Frye & Dawe (2008)	<p>12 mothers close to release or on community custody orders</p> <p>1 child of focus per mother</p> <p>Average age of child = 5.6 years (range 2-12)</p>	Single-group pre-test post-test design with no control	<p>Parenting Under Pressure program (individualised) - 10 modules about different aspects of parenting (e.g., mindful child management), contextualised to individual family situations</p> <p>No pre-determined number of intervention sessions</p> <p>Average number of treatment sessions = 20</p> <p>Mean duration of sessions = 86 minutes</p>	<p>Strengths and Difficulties Questionnaire – Total Problem Score</p> <p>Administered pre-test, post-test and 3 months after end of intervention</p>	Aggregated measure of social, emotional and behavioural functioning (rated by parent)
Harris & Landreth (1997)	<p>22 mothers in detention centre</p> <p>1 child of focus per mother</p> <p>Average age of child = 5 years (range 3-10)</p>	<p>Between-groups quasi-experimental design</p> <p>Non-randomised assignment to experimental and control groups</p> <p>Control were on a wait-list</p>	<p>Filial therapy (group-based) – participants learnt child-centred play therapy principles through didactic instruction and practice</p> <p>Participants had 2-hour sessions twice a week for 5 weeks</p>	<p>Filial Problem Checklist</p> <p>Administered pre-test and post-test</p>	Problem behaviours (rated by parent)

Author(s)	Participants	Study design	Intervention	Measure(s)	Domain of child functioning assessed
Harrison (1997)	30 fathers in prison 1 child of focus per father Children aged 8-17 years	Between-groups experimental design Random assignment to experimental and control groups, matched on age, number of years of education and number of children Control received alternative intervention with no professional guidance	Group sessions with a focus on - Knowledge of child developmental stages - Behaviour management techniques Participants had 2.5-hour sessions thrice a week for 6 weeks	Self-perception Profile for Children (for 8-12 years old) / Adolescents (for 13-17 years old) Administered pre-test and post-test	Self-concept (rated by child)
Landreth & Lobaugh (1998)	32 fathers in prison 1 child of focus per father Average age of child = 6.2 years (range 4-9)	Between-groups experimental design Random assignment to experimental and control groups, matched on education level, ethnic origin and age of child) Control group received typical services and were instructed to continue seeing their children on weekly visits	Filial therapy (group-based) – participants learnt child-centred play therapy principles through didactic instruction and practice Participants had 1.5-hour sessions once a week for 10 weeks	Filial Problem Checklist Joseph Preschool and Primary Self-concept scale Administered pre-test and post-test	Problem behaviours (rated by parent) Self-concept (rated by child)

Author(s)	Participants	Study design	Intervention	Measure(s)	Domain of child functioning assessed
Menting, de Castro, Wijngaards-de Meij & Matthys (2014)	113 mothers in prison or recently released Up to 3 children of focus per mother Average age of child = 7.4 years (range 2-10)	Between-groups quasi-experimental design In 4 of 6 recruitment periods, random assignment to experimental and control groups In 2 of 6 recruitment periods, all participants were assigned to experimental group due to low numbers.	Training sessions based on Incredible Years Parent Training (group-based) – watched videotapes of parent models, then practised parenting skills through role-play Participants had 2-hour group sessions once a week for 12 weeks This was followed by 4 1.5-hour home visits for each mother	Eyberg Child Behavior Inventory Teacher's Report Form (for 6-18 years old) OR Caregiver-Teacher Report Form (for 1.5-5 years old) Administered pre-test, after completion of group sessions, halfway through the home visits and at the end of all home visits.	Problem behaviours (rated by parent) Problem behaviours (rated by teacher)

Appendix C: WoE A – Methodological Quality

WoE A evaluates a study through its methodological quality and is not specific to the review question. The APA Task Force Review Coding Protocol for group designs (Kratochwill, 2003) was used, with adaptation, to evaluate the studies in this review. The adapted protocol assessed measurement (Section II A), comparison groups (Section II B), appropriate statistical analyses (Section II C) and follow-up assessment (Section II I). The changes to the protocol are detailed below, together with the rationale for the changes:

Section removed	Rationale
I B7 – B8	Studies all used quantitative methods.
II C2 – C5	Study outcomes were evaluated separately, and not as part of WoE A.
II D	Clinical significance was not of relevance to this review.
II E	This review was interested only in one specific component – parent training.
II F	Different parenting programs were used for different studies.
II G	The replication of studies was not of relevance to this review.
II H	Site of implementation was not of relevance to the review.

Studies were assigned a rating of 0-3 on each of the four assessment measures, guided by criteria provided by Kratochwill (2003). These criteria are reflected below. The ratings on all measures were then averaged to form the overall WoE A score for the study. An example of a completed coding protocol can be found in Appendix F.

Criteria for Measurement

Rating	Criteria
3 (Strong evidence)	<ul style="list-style-type: none">- All measures used have a reported reliability of .85 or higher- Multiple methods of data collection- Multiple sources of information
2 (Promising evidence)	<ul style="list-style-type: none">- Most measures used have a reported reliability of .70 or higher- Either multiple methods of data collection or multiple sources of information
1 (Weak evidence)	<ul style="list-style-type: none">- Most measures used have a reported reliability of .50 or higher
0 (No evidence)	<ul style="list-style-type: none">- Measures do not meet criteria for a rating of '1'

Criteria for Comparison Group

Rating	Criteria
3 (Strong evidence)	<ul style="list-style-type: none">- An active comparison group was used- Initial group equivalency was established- Change agents were counterbalanced- Equivalent mortality with low attrition
2 (Promising evidence)	<ul style="list-style-type: none">- A 'no intervention' comparison group was used- At least two of: initial group equivalency; counterbalancing of change agents; equivalent mortality with low attrition.
1 (Weak evidence)	<ul style="list-style-type: none">- Presence of a comparison group
0 (No evidence)	<ul style="list-style-type: none">- No comparison group

Criteria for Appropriate Statistical Analysis

Rating	Criteria
3 (Strong evidence)	<ul style="list-style-type: none">- Appropriate statistical analysis with appropriate units of analysis- Sufficiently large N- Familywise error rate controlled for, where appropriate- Significant outcomes for at least 75% of primary measures

- 2 (Promising evidence)
 - Appropriate statistical analysis with appropriate units of analysis
 - Sufficiently large N
 - Familywise error rate controlled for, where appropriate
 - Significant outcomes for at least 50% of primary measures
- 1 (Weak evidence)
 - Appropriate statistical analysis with appropriate units of analysis
 - Familywise error rate controlled for, where appropriate
 - Significant outcomes for at least 25% of primary measures
- 0 (No evidence)
 - Criteria for rating of '1' were not met.

Criteria for Follow-up Assessment

Rating	Criteria
3 (Strong evidence)	<ul style="list-style-type: none"> - Follow-up assessments conducted over multiple intervals with all participants in original sample - Follow-up assessment measures are similar to those used to analyse primary outcomes
2 (Promising evidence)	<ul style="list-style-type: none"> - Follow-up assessments conducted at least once with majority of participants in original sample - Follow-up assessment measures are similar to those used to analyse primary outcomes
1 (Weak evidence)	<ul style="list-style-type: none"> - Follow-up assessments conducted at least once with some participants in original sample
0 (No evidence)	<ul style="list-style-type: none"> - No follow-up assessment

WoE A ratings

Author(s)	Measure				Overall WoE A
	Measurement	Comparison group	Appropriate statistical analysis	Follow-up assessment	
Frye & Dawe (2008)	1	0	2	3	1.5 Medium
Harris & Landreth (1997)	0	1	2	0	0.75 Low
Harrison (1997)	1	2	1	0	1 Low
Landreth & Lobaugh (1998)	1	2	1	0	1 Low
Menting, de Castro, Wijngaards-de Meij & Matthys (2014)	2	2	3	1	2 Medium

Note: WoE scores of 1.4 or less are classified as low, scores between 1.4 and 2.5 are classified as medium and scores or 2.5 or greater are classified as high.

Appendix D: WoE B - Methodological Relevance

WoE B is a review-specific rating related to the suitability of the research design (Gough, 2007). Criteria for WoE B are based on evidence hierarchies (Guyatt et al., 1995). Specifically, studies with the best research designs are those with fewer threats to validity.

For this review, these factors included:

- Control group: A well-established control group increases the likelihood that significant outcomes can be attributed to the intervention and not other factors such as maturation effects.
- Measures taken to reduce bias: All studies used parents as a source of information about child outcomes. Parents were the subject of intervention and may not have been blind to the condition they belonged to. The Cochrane risk-of-bias tool (Higgins et al., 2016) highlight these as possible sources of bias. Studies that put measures in place to reduce these biases have greater validity.
- Use of standardised measures: Their use increase reliability and validity of findings.
- Multiple assessment time points: The use of multiple time points help increase validity of findings as changes are tracked more gradually.

WoE B weighting criteria

Criterion	3 (High)	2 (Medium)	1 (Low)
Control group	Random assignment	Non-random assignment	No control group
Measures taken to reduce bias	Parents were blind to condition <u>and</u> an additional source was used to rate child.	Parents were blind to condition <u>or</u> an additional source was used to rate child.	No measures reported
Use of standardised measures	Use of multiple standardised measures	At least one outcome is a standardised measure.	All outcomes are non-standardised measures
Multiple assessment time points	4 or more assessment time points	3 assessment time points	Pre-test and post-test

WoE B ratings

Author(s)	Measure				Overall WoE B
	Control group	Measures taken to reduce bias	Use of standardised measures	Multiple assessment time points	
Frye & Dawe (2008)	1	1	2	2	1.5 Medium
Harris & Landreth (1997)	2	1	2	1	1.5 Medium
Harrison (1997)	3	2	2	1	2 Medium
Landreth & Lobaugh (1998)	3	2	3	1	2.25 Medium
Menting, de Castro, Wijngaards-de Meij & Matthys (2014)	2.5*	2	3	2.5**	2.5 High

Note: WoE scores of 1.4 or less are classified as low, scores between 1.4 and 2.5 are classified as medium and scores or 2.5 or greater are classified as high.

* Random assignment took place for 4 of 6 recruitment periods. All participants in the other 2 recruitment periods were assigned to the intervention condition due to a lack of participants during these periods.

** Two measures of child outcomes were taken for this study. One measure was administered at 3 time points, and the other was administered at 4 time points.

Appendix E: WoE C - Relevance of Focus of Study

WoE C is a review-specific rating of how relevant the study is to answer the review question (Gough, 2007). For this review, the relevant factors include:

- Focus of study: A study that was primarily focused on the impact of intervention on child outcomes would be more relevant, given the review question. This is especially so given that much of the literature on parenting interventions for incarcerated parents focuses on parent outcomes (Purvis, 2013).
- Accessibility of intervention to parents: For the findings to be generalisable to as many incarcerated parents as possible, studies should not have additional inclusion criteria for the intervention used. These criteria exclude considerations for the child's safety.
- Year of publication: Recent publications are more up-to-date with current psychological theory and receive a higher score.

WoE C weighting criteria

Criterion	3 (High)	2 (Medium)	1 (Low)
Focus of study	Child outcome is the primary outcome measure	Child outcome is one of the key outcome measures (e.g., identified through one of the hypotheses)	Child outcome is one of the outcome measures
Accessibility of intervention to parents	No additional requirements	Maximum of one additional requirement (e.g., minimum literacy level)	More than one additional requirement
Year of Publication	2010 and after	2000 - 2009	1999 and before

WoE C ratings

Author(s)	Focus of study	Measure Accessibility of intervention to parents	Year of publication	Overall WoE C
Frye & Dawe (2008)	2	3	2	2.33 Medium
Harris & Landreth (1997)	1	2	1	1.33 Low
Harrison (1997)	2	3	1	2 Medium
Landreth & Lobaugh (1998)	2	2	1	1.67 Medium
Menting, de Castro, Wijngaards-de Meij & Matthys (2014)	3	3	3	3 High

Note: WoE scores of 1.4 or less are classified as low, scores between 1.4 and 2.5 are classified as medium and scores or 2.5 or greater are classified as high.

Appendix F: Example of Coding Protocol

Coding Protocol: Group-Based Design

Domain:

- School- and community-based intervention programs for social and behavioral problems
- Academic intervention programs
- Family and parent intervention programs
- School-wide and classroom-based programs
- Comprehensive and coordinated school health services

Name of Coder: _____

Date: _____

Full Study Reference in proper format: Menting, A. T. A., de Castro, B. O., Wijngaards-de Meij, L. D. N. V., & Matthys, W. (2014). A Trial of Parent Training for Mothers Being Released From Incarceration and Their Children. *Journal of Clinical Child and Adolescent Psychology*, 43(3), 381–396.

Intervention Name (description of study): Incredible Years Parent Training

Study ID Number: 5

Type of Publication:

- Book/Monograph
- Journal Article
- Book Chapter
- Other (specify):

1. General Characteristics

A. General Design Characteristics

A1. Random assignment designs (if random assignment design, select one of the following)

Completely randomized design – only for 4 of 6 recruitment periods. For the last 2, all participants were assigned (non-random) to intervention group because of small numbers

Randomized block design (between participants, e.g., matched classrooms)

Randomized block design (within participants)

Randomized hierarchical design (nested treatments)

A2. Nonrandomized designs (if non-random assignment design, select one of the following)

Nonrandomized design

Nonrandomized block design (between participants)

Nonrandomized block design (within participants)

Nonrandomized hierarchical design

Optional coding for Quasi-experimental designs

A3. Overall confidence of judgment on how participants were assigned (select one of the following)

Very low (little basis)

Low (guess)

Moderate (weak inference)

High (strong inference)

Very high (explicitly stated)

N/A

Unknown/unable to code

B. Participants

Total size of sample (start of study): 113__

Intervention group sample size: 86

Control group sample size: 27

C. Type of Program

- Universal prevention program
- Selective prevention program
- Targeted prevention program
- Intervention/Treatment
- Unknown

D. Stage of Program

- Model/demonstration programs
- Early stage programs
- Established/institutionalized programs
- Unknown

E. Concurrent or Historical Intervention Exposure

- Current exposure
- Prior exposure
- Unknown

2. Key Features for Coding Studies and Rating Level of Evidence/Support

(Rating Scale: 3= Strong Evidence, 2=Promising Evidence, 1=Weak Evidence, 0=No Evidence)

A. Measurement (Estimating the quality of the measures used to establish effects)

A1 The use of the outcome measures produce reliable scores for the majority of the primary outcomes

Yes

- Eyberg Child Behavior Inventory (ECBI) – Internal consistency of .89-.93
- Teacher’s Report Form/Caregiver-Teacher Report Form – Internal consistency of .86-.97

No

Unknown/unable to code

A2 Multi-method (at least two assessment methods used)

- Yes
- No
- N/A
- Unknown/unable to code

A3 Multi-source (at least two sources used self-reports, teachers etc.)

- Yes
- No
- N/A
- Unknown/unable to code

A4 Validity of measures reported (well-known or standardized or norm-referenced are considered good, consider any cultural considerations)

- Yes validated with specific target group
- In part, validated for general population only
- No
- Unknown/unable to code

Overall Rating for measurement_2__

3= Strong Evidence 2=Promising Evidence 1=Weak Evidence 0=No Evidence

B. Comparison Group

B1 Type of Comparison Group (Select one of the following)

- Typical intervention (typical intervention for that setting, without additions that make up the intervention being evaluated) – Usual services + offer to find adequate services when needed
- Attention placebo
- Intervention element placebo
- Alternative intervention

- Pharmacotherapy
- No intervention
- Wait list/delayed intervention
- Minimal contact
- Unable to identify type of comparison

B2 Overall confidence of judgment on type of comparison group

- Very low (little basis)
- Low (guess)
- Moderate (weak inference)
- High (strong inference)
- Very high (explicitly stated)
- Unable to identify comparison group

B3 Counterbalancing of change agent (participants who receive intervention from a single therapist/teacher etc were counter-balanced across intervention)

- By change agent
- Statistical (analyse includes a test for intervention)
- Other
- Not reported/None

B4 Group equivalence established (select one of the following)

- Random assignment
- Posthoc matched set
- Statistical matching
- Post hoc test for group equivalence

B5 Equivalent mortality

- Low attrition (less than 20 % for post)
- Low attrition (less than 30% for follow-up)

Intent to intervene analysis carried out

Overall rating for Comparison group 2

3= Strong Evidence 2=Promising Evidence 1=Weak Evidence 0=No Evidence

C. Appropriate Statistical Analysis

Analysis 1 Hierarchical Linear Modeling

- Appropriate unit of analysis
- Familywise/experimenter wise error rate controlled when applicable – N.A.
- Sufficiently large N – A power analysis using G*Power (power = 0.80, alpha = 0.05) reflected a minimum sample size of 43.

Analysis 2 _____

- Appropriate unit of analysis
- Familywise/experimenter wise error rate controlled when applicable
- Sufficiently large N

Analysis 3 _____

- Appropriate unit of analysis
- Familywise/experimenter wise error rate controlled when applicable
- Sufficiently large N

Overall rating for Statistical Analysis 3

3= Strong Evidence 2=Promising Evidence 1=Weak Evidence 0=No Evidence

I. Follow Up Assessment

Timing of follow up assessment: Further follow-up assessments underway

Number of participants included in the follow up assessment: specify: - _____

Consistency of assessment method used: specify: Not specified

Rating for Follow Up Assessment _1_

3= Strong Evidence

2=Promising Evidence

1=Weak Evidence

0=No Evidence