

■ What works in the psychological treatment of child conduct problems? An umbrella review of meta-analytic studies

Estrella Romero, María Álvarez-Voces, Beatriz Díaz-Vázquez, & Laura López-Romero
University of Santiago de Compostela

Abstract

Conduct problems (CP) constitute a major field for child clinical psychology, in terms of not only prevalence, but also developmental, social, and clinical relevance. This study carried out an umbrella review of meta-analyses on the efficacy of treatment/indicated prevention of child CP (mean age range up to 12 years). Following a registered protocol, several databases (i.e., Web of Science, PsycINFO, PROSPERO, and The Cochrane Library) were searched for meta-analyses published from January 2002 to March 2022. Nine meta-analytic publications met the required criteria, and the main characteristics and findings of the studies were systematically described. Weighted effect sizes (ESs) were calculated through RStudio program. Analyses of heterogeneity, publication bias, quality (AMSTAR-2), and credibility were also conducted. Results indicate that parent training has been the most studied intervention, with a weighted ES of $d = -0.49$ (95% CI -0.67 to -0.32). The reviewed studies identified several moderators for the efficacy of parent training, including individual (severity of CP), family (financial disadvantage), and intervention (delivery format) characteristics. Results from child-centered play therapy were also analyzed, $d = -.34$ (95% CI -0.40 to -0.28), but the reduced number of meta-analyses and the weakness detected by quality assessment suggest the need of cautiously considering the pooled effects. In general, evidence seems to be suggestive of the efficacy of treatment of CP, particularly for parent training. Nevertheless, efficacy seems to be moderate, heterogeneity indexes are high, and quality assessments of meta-analyses are often suboptimal. This study suggests several avenues to strengthen knowledge in this field.

Keywords: Conduct problems; Intervention; Children; Meta-analysis; Umbrella review.

Resumen

¿Qué funciona en el tratamiento psicológico de los problemas de conducta en niños? Revisión umbrella de estudios meta-analíticos. Los problemas de conducta (PC) son un foco prioritario de atención en la psicología clínica infantil. Este estudio presenta una revisión umbrella de meta-análisis sobre la eficacia del tratamiento/prevenición indicada de los PC infantiles (rango de edad media hasta 12 años). Siguiendo un protocolo registrado, se buscaron meta-análisis (enero 2002 a marzo 2022) en Web of Science, PsycINFO, PROSPERO y The Cochrane Library. Nueve meta-análisis cumplieron con los criterios requeridos y sus características y hallazgos fueron descritos sistemáticamente; además, con RStudio se calcularon los tamaños del efecto (TEs) ponderados. Se realizaron análisis de heterogeneidad, sesgo de publicación, calidad (AMSTAR-2) y credibilidad. El entrenamiento parental fue la intervención más estudiada, con un TE medio ponderado de $d = -0.49$ (IC del 95%: -0.67 a -0.32), y se identificaron diversos moderadores de eficacia, incluyendo características individuales (gravedad de los PC), familiares (desventaja económica) y de intervención (formato de administración). También se analizaron los resultados de la terapia de juego centrada en el niño, $d = -.34$ (IC del 95%: -0.40 a -0.28), pero el reducido número de meta-análisis y la debilidad detectada por la evaluación de la calidad sugieren la necesidad de considerar con cautela los efectos agrupados. En general, la evidencia es sugestiva de eficacia del tratamiento de los PC, específicamente del entrenamiento parental. No obstante, la eficacia parece ser moderada, los índices de heterogeneidad elevados y las evaluaciones de calidad de los meta-análisis no suelen ser óptimas. Este estudio sugiere varias vías para reforzar el conocimiento en este campo.

Palabras clave: Problemas de conducta; Intervención; Niños; Meta-análisis; Umbrella review.

Conduct problems (CP) are among the most relevant disturbances addressed by clinical child and adolescent psychology. Under the concept of CP a variety of behavioural patterns are subsumed,

including oppositionality, vindictiveness, aggression, temper tantrums, noncompliance and rule breaking. These patterns may reach clinical significance, conforming disorders enlisted by the diagnos-

Corresponding author:

Estrella Romero.

Facultad de Psicología, University of Santiago de Compostela.

Campus Sur, s/n, 15781 Santiago de Compostela.

E.mail: estrella.romero@usc.es

tic classifications, such as oppositional-defiant disorder (ODD) and conduct disorder (CD). Both at clinical and subclinical levels, CP represent a significant public health concern, as they are one of the most prevalent types of problems in children and adolescents (e.g., Merikangas & Kessler, 2022), and, additionally, CP seem to have experienced a sharp increase in the context of the pandemic times (Lebrun-Harris et al., 2022; Romero et al., 2020). During the last years, large-scale surveys have shown that CP are the primary cause for referring children to mental health services (Encuesta Nacional de Salud de España [ENSE], 2017; Ghandour et al., 2019). The implications of CP for behavioural, emotional, and social development are also well described in the literature, including an increased risk for academic failure, depression, drug abuse, social exclusion, and criminality, thus entailing substantial health, educational and social costs (Rissanen et al., 2021).

The existence of diverse developmental profiles among children with CP has been one major tenet in developmental psychopathology for the last decades (e.g., Cicchetti, 2016). Even when the specific number and shape of the different trajectories is still a debated issue (Girard et al., 2018; Gutman et al., 2019), a big body of research has supported the need to differentiate at least between childhood-onset and adolescence-onset CP (Moffitt, 1993; Patterson, 1982). While adolescent-onset CP tend to be linked to the developmental tasks of the teenaging years and, therefore, they present a higher chance of remission after adolescence, childhood-onset CP tend to be associated to a complex net of interwoven determinants, operating in a cascade mode (Masten & Cicchetti, 2010); this process may lead to chronic CP patterns, which are increasingly difficult to treat as the child grows up. Based on this evidence, the intervention on childhood-onset CP becomes a major challenge for mental health providers; intervening on children's CP is not only important in terms of treatment needs, but it is also considered a major opportunity for prevention of future clinical and psychosocial dysfunctions (Harley et al., 2008).

For many years, researchers have tried to design, implement, and evaluate treatment strategies for children's CP. Psychological treatments have aimed to change children's CP through the intervention in different settings and domains: e.g., parenting/family relations (Brinkmeyer & Eyberg, 2003), teacher/classroom interactions (Leflot et al., 2010), and children's cognitions and skills (Lochman & Wells, 2002). To apprehend and summarize evidence on this topic, several reviews have been conducted so far, with results that have conveyed both optimism (Furlong et al., 2012) and pessimism (Bakker et al., 2017) on the efficacy of available treatments.

Among the published reviews, meta-analyses occupy a central space, as they provide quantitative estimations of treatment effects, and they allow delineation of possible individual, social or treatment-related moderators. Meta-analytic reviews have been performed in the field of children's CP (e.g., Comer et al., 2013; Ray et al., 2015), yet there is a paucity of overarching studies synthesizing what we have learned from such quantitative reviews. This study attempts to contribute to this line, by systematically collecting and analyzing results from meta-analyses on the efficacy of psychological treatments for children with CP. In the view of the developmental implications of CP (with or without the ODD/CD diagnostic labels), our review includes meta-analyses dealing with strict-sense *treatment* and with *indicated prevention* (Gordon, 1987). Therefore, we considered studies on interventions for children that were referred for clinical assistance, but also interventions for children whose CP were identified through screening procedures in non-clinical populations. With an *umbrella*

perspective (Aromataris et al., 2015; López-López et al., 2022), the present study portrays the meta-analytic results on this topic, with the aim of identifying the types of programs that have been most studied during the last two decades, the efficacy associated to the different therapies, the strength of such evidence, and the limitations to be addressed by future research.

Method

Registration and guidelines

The protocol for this study was registered on PROSPERO (registration number CRD42022322438). This study was conducted according to the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) (Page et al., 2021), as well as recommendations for umbrella reviews (Fusar-Poli & Radua, 2018) and the Assessing the Methodological Quality of Systematic Reviews (AMSTAR-2) (Shea et al., 2017). The PRISMA checklist is provided in Table A1 (see Appendix, <https://osf.io/6wvb5/>).

Search strategy

A comprehensive search was performed by two researchers (M.A.V., B.D.V.) in the following data bases: Web of Science, PROSPERO, PsycINFO and The Cochrane Library. The descriptors used for the search are presented in Table 1. Language restrictions were applied (English and Spanish). Other restrictions were publication date (01/01/2002 to 15/03/2022), age range of participants (mean age up to 12 years old) and article type: meta-analysis and network meta-analysis. The final literature search was conducted on 18 March 2022. The complete search strategy is described in PROSPERO and in the Appendix (Table A2).

Table 1. Search Strategy for the Meta-Analyses

| Search items | Descriptors |
|--------------------------------------|--|
| 1. Condition or domain being studied | (conduct problem*) OR (conduct disorder*) OR (CD) OR (disruptive behaviour*) OR (externaliz*) OR (oppositional defiant disorder*) OR (ODD) |
| 2. Intervention | (treatment*) OR (invervent*) OR (therap*) OR (psychotherap*) |
| 3. Age | (child*) |
| Combination | #1 AND #2 AND #3 |

Studies retrieved from the databases were encoded in a bibliographic manager software (Refworks). These studies were systematically reviewed by two researchers independently of each other (M.A.V., B.D.V.). Moreover, the reference list of all studies was checked for additional relevant records. Interrater reliability was Kappa = .89, which can be rated as strong (McHugh, 2012). Divergences in study selection were resolved by two independent researchers (E.R.T., L.L.R.).

Inclusion and exclusion criteria

The eligibility criteria for this umbrella review were based on the Population, Intervention, Comparator, Outcome, Study design (PICOS) statement (Higgins et al., 2019), as shown in Table 2.

Table 2. Eligibility Criteria for Inclusion of Meta-Analyses

| | | Inclusion criteria | Exclusion criteria |
|---|--------------|--|---|
| P | Participate | Children whose mean age is up to 12 years old with CP defined on the basis of standardized tools. CP were referred to both clinical (e.g., conduct disorder, oppositional defiant disorder) and subclinical (e.g., disruptive, externalizing) levels | Children with neurodevelopmental problems (e.g., autism spectrum disorder) or intellectual disabilities |
| I | Intervention | Psychological interventions specifically designed to treat CP | Pharmacological interventions, prevention programs (universal and selective prevention), interventions not specifically designed to reduce child CP, interventions specifically designed to treat internalizing disorders or ADHD |
| C | Comparison | Comparison group: no treatment, treatment as usual, another type of treatment and waiting-list control group | With individuals without CP |
| O | Outcome | Effects on CP, moderating treatment variables | - |
| S | Study | Metanalysis and network meta-analysis | Qualitative systematic reviews |

Data extraction

Two researchers (B.D.V., M.A.V.) extracted data independently using a previously defined protocol. Divergences were reviewed by a third (L.L.R.) and fourth (E.R.T.) researcher. Qualitative data extracted included: (a) first author and year of publication, (b) countries of the articles included in the meta-analysis, (c) participant data (sample size, gender distribution, mean age range), (d) study objectives, (e) description of the control group, (f) instruments used to measure CP, (g) informants for the outcome measures, (h) type of intervention studied, (i) intervention target, (j) intervention characteristics and (k) main outcomes (see Table A3 in Appendix). Quantitative data included: (a) total number of individual studies included in each meta-analysis (k), (b) reported ES in relation to the decrease in externalizing problems after the application of the intervention (e.g., Cohen's *d*, Hedges' *g*), (c) 95% confidence interval of the ES, and (d) heterogeneity (I^2 and Q) (see Table 3). In addition, Table A4 (see Appendix), shows the search strategy (database, search period) and the number of WOS citations for each meta-analysis.

When it was not possible to extract some relevant data (e.g., gender distributions, individual studies entered to calculate each ES, country of origin of the studies), we tried to contact the authors. From

a total of seven authors who were contacted, one of them provided the necessary missing data.

Correction of primary study overlap

The overlap of individual studies among the included meta-analyses can lead to a significant distortion of the results (Munder et al., 2013). Therefore, following the instructions of Fusar-Poli & Radua (2018), when two or more meta-analyses overlapped more than 50% of their individual studies, we included the one with the largest database and the most recent one.

After elimination, we observed that the highest overlap was 15 studies (Leijten et al., 2018; Lundahl et al., 2006), followed by 11 (Leijten et al., 2008; Menting et al., 2013). Out of the total of 278 studies we observed 30 overlaps (10.79%). Additional information on overlapping articles is included in Table A5 in Appendix.

Data analysis

Quantitative data analysis was carried out with the RStudio program for Windows. As an umbrella pooling approach, bare bones meta-analysis of standardized differences was performed using the psychometric method of Hunter & Schmidt (2004). As input data, the ESs reported for each study were considered. The index used to obtain the overall ES was Cohen's *d*; for the studies that did not report this index, the ES was transformed into it. Then, the ES weighted by the sample size (pooled Cohen's *d*), the standard error (SE) and the 95% confidence intervals (CI) were calculated. Heterogeneity was calculated with the I^2 and Q statistics. In relation to the I^2 statistic, heterogeneity was considered low if it was 25%, moderate if it was 50% and high if it was 75% (Higgins et al., 2003).

Publication bias

Publication bias was analysed using the File-safe N with the Orwin's approximation to detect the number of studies with an ES of zero that would reduce the observed average ES to a specific value (e.g., half or quarter). Funnel plots and asymmetry indexes with Egger's Test could not be performed, as the number of resulting meta-analyses was deemed too low to provide reliable results, according to the guidelines by Guyatt et al. (2011).

Quality assessment and credibility of evidence

The methodological quality of each meta-analysis was assessed by two researchers (B.D.V. and M.A.V.) who worked independently with the Assessment of Multiple Systematic Reviews tool (AMSTAR-2); resulting minor discrepancies were discussed to reach a complete agreement. AMSTAR-2 classifies reviews into: (a) critically low-quality review, (b) low quality review, (c) moderate quality review and (d) high quality review.

To assess credibility, studies were classified into 4 different classes (Belbasis et al., 2015; Bellou et al., 2017; Fusar-Poli & Radua, 2018):

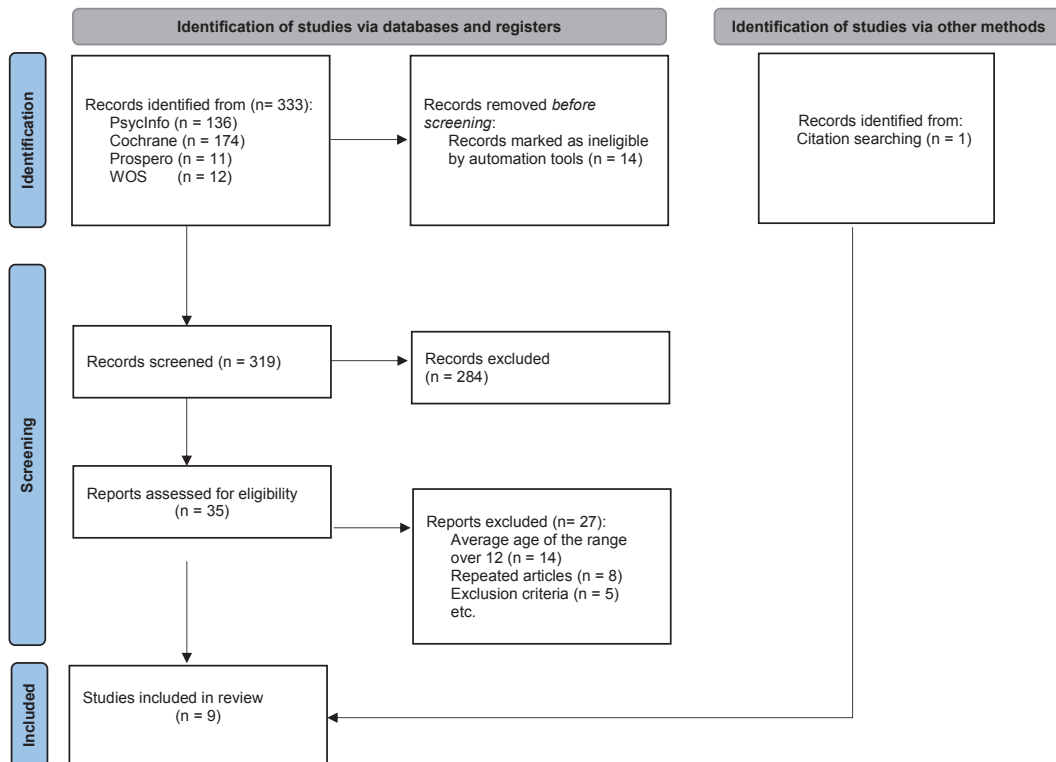
- Class I – Convincing: number of cases >1000, $p < 10^{-6}$, $I^2 < 50\%$, prediction interval excluding the null, no small-study effects and no excess significance bias.
- Class II – Highly suggestive: number of cases >1000, $p < 10^{-6}$, largest study with a statistically significant effect and class I criteria not met.
- Class III – Suggestive: number of cases >1000, $p < 10^{-3}$ and class I–II criteria not met.

Table 3. Results of all Meta-Analysis Included in the Review (K = 9)

| Reference | Participants (N) | Studies (k) | ES ^a | ES (Cohen's d) | 95 % CI (p) | Heterogeneity |
|------------------------------|------------------|--|-------------------------|----------------|-----------------------------------|--|
| Burkey et al. (2018) | 4,441 | 11 | SMD = -0.56 | -.56 | -0.79 to -0.32 | I ² = 73.6% |
| Comer et al. (2013) | 3,042 | 36 | Hedges' g = .82 | -.82 | -1.01 to -.63 (p<.001) | Q = 299.3 |
| Leijten et al. (2018) | 4,892 3,327 | IP = 50 T = 45 | d = -0.55 d = - 0.69 | -.55 -.69 | -0.70 to -0.39 -0.84 to - 0.54 | I ² = 66.66% - |
| Lundahl et al. (2006) | 1,838 648 | 29 (five years or less) 10 (5-10 years) | d = .44 d = .31 | -.44 -.31 | - - | Q = 49.24 p < .05 Q = 19.65 p < .05 |
| Menting et al. (2013) | 4,745 | 50 | d = .27 | -.27 | .21 to .34 (p < .001) | - |
| Parker et al. (2021) | 504 | 14 | Hedges' g = -.34 | -.34 | -.52 to -.17 (p<.001) | Q=5.08 p = .97 I ² = 0% |
| Ray et al. (2015) | 407 | 7 | d =.34 | -.34 | .10 to .58 (p <.05) | Q = 2.16 (p >.05) |
| Van IJzendoorn et al. (2022) | 1,03 | 13 | r = .07 | -.14 | -0.42 to -0.09 (p <.01) | I ² = 63.83% |
| Veenman et al. (2018) | 1,081 828 | At risk: 6 Clinical: 7 | d = -.26 d = -.19 | -.26 -.19 | -0.35 to -0.04 (p =.01) | Q = 18.58 I ² = 73.00% (p < .001) Q = 6.78 I ² = 11.47% (p = .34) |

Note. (-) = data not reported in the article. ES (Cohen's d): ES calculated from the ESs reported by the articles.
^aThe signs used by each author are respected.

Figure 1. PRISMA Flow Diagram Illustrating Literature Research and Selection Process



Note. Extracted from "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews" (Page et al., 2021).

- Class IV – Weak: $p < .05$ and class I–III criteria not met.
- Non-significant: $p > .05$.

Results

Literature search results / Study selection

The literature search identified 333 relevant articles and 319 unique citations across all sources. Of these, 35 full papers were assessed for eligibility and nine reviews were selected for synthesis (see Figure 1 for the PRISMA flow chart). The list of articles excluded following full-text review, together with the reasons for exclusion, is also available in Appendix (Table A6).

Study characteristics

A total of nine meta-analyses evaluating the efficacy of the treatment of CP were included. A list of the individual studies included in each meta-analysis can be found in Table A7 in the Appendix. It should be noted that three meta-analyses did not provide such information. Description of the study characteristics are presented in Table A3 in Appendix. The oldest publication was by Lundahl et al. (2006) and the most recent by Van Ijzendoorn et al. (2022). The most cited article in WOS was Lundahl et al. (2006) with a total of 608 citations so far (see Table A4 in Appendix). Three meta-analyses did not provide data on the country of origin of the individual studies. Of those meta-analyses that did provide data, the most frequent country of origin was USA (57.14%), Canada (42.86%) and the Netherlands (42.86%).

The total sample consisted of 26,763 boys and girls with detected clinical or subclinical CP. Percentages by gender could not be reported because five of the nine meta-analyses included did not report such data. The type of intervention most studied by the meta-analyses was parent training (6/9); a relevant number of studies were focused, particularly, on Incredible Years (3/9). Other interventions addressed were play therapy (3/9), classroom programs (1/9) and counselling (1/9). The target of the interventions was parents (6/9), children (6/9), and teachers (2/9). A few studies were specifically focused on randomized controlled trials (RCTs; 3/9), while some others included both randomized and non-randomized designs (6/9).

Intervention characteristics varied widely across studies due to the different nature of the therapies studied. Interventions included both individual and group formats, with an average duration ranging from seven days to four years. As for the control groups, the most frequent types were waiting list [WL] (5/9) and no treatment [NT] (4/9). As outcome measures, 25 different instruments for measuring CP were used, the most frequent being: Child Behaviour Checklist [CBC] (7/9), Eyberg Child Behaviour Inventory [ECBI] (4/9), Strengths and Difficulties Questionnaire [SDQ] (3/9) and Behavioural Assessment System for Children [BASC] (3/9).

Main findings from the individual meta-analyses

Efficacy of the interventions

Most of the treatments assessed showed to be effective in reducing CP; only one study analyzing videofeedback-based interventions showed no substantial treatment effects for externalizing problems (Van Ijzendoorn et al., 2022). Parent training and play therapy were the most investigated interventions across all the reviews, with significant short- and long-term effects on the reduction of behavioural problems. Timing of post-treatment assessment significantly moderated the effects, with immediate effects being larger than effects at subsequent

follow-ups (Comer et al., 2013). However, some studies reported that the benefits of the intervention were maintained up to six-month (Burkey et al., 2018) and one-year follow-ups (Leijten et al., 2018).

Studies also found that the largest reductions in CP were found in those groups with more severe problems at the start of treatment (Lundahl et al., 2006; Menting et al., 2013).

In terms of theoretical orientation, behavioural and non-behavioural interventions were compared (Comer et al., 2013); results showed that behavioural interventions offered better results in reducing externalising behaviours than non-behavioural treatments (e.g., family systems approaches, nondirective counselling).

The efficacy of specific therapeutic techniques was evaluated by the examined studies. Specifically, three techniques were associated with stronger program effects; positive reinforcement as a general technique, praise as a specific operationalization of positive reinforcement, and the use of natural or logical consequences as a discipline technique (Leijten et al., 2018).

Finally, considering the level of intervention, Leijten et al. (2018) found that indicated prevention and treatment were the most effective interventions in reducing behavioural problems compared to the other types of prevention. Some other studies also noted that treatment was more powerful than prevention approaches in reducing externalizing problems (Burkey et al., 2018; Menting et al., 2013).

Moderator variables

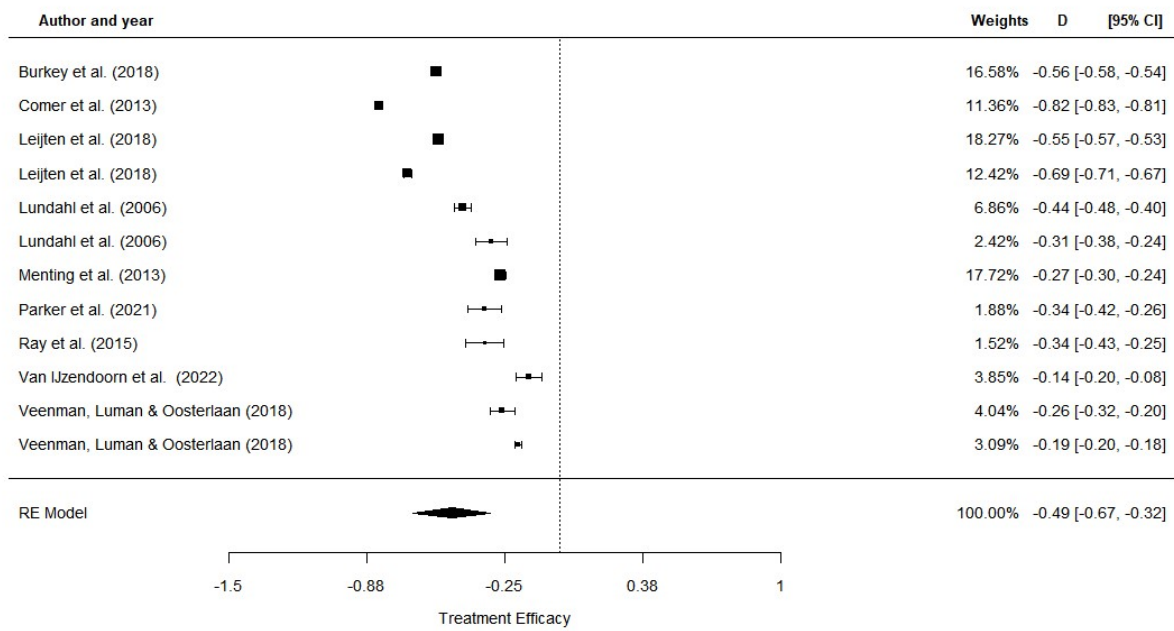
The examined studies highlighted some relevant variables as potential moderators of intervention efficacy.

With respect to age of children, the studies found that programmes were effective at all ages (Burkey et al., 2018) or that there is no significant difference considering age (Lundhal et al., 2016). However, some authors reported larger effects in samples of older youth and with higher percentages of males (Comer et al., 2013).

Some family characteristics were also analysed. As for ethnic characteristics, treatment effects were consistent across samples of varying ethnic compositions (Comer et al., 2013; Parker et al., 2012). Some studies considered the marital status of parents: while Lundahl et al. (2006) found that being single did not moderate the effectiveness of parent training, Menting et al. (2013) corroborated these findings, but pointed to a certain tendency towards downward moderation of being single. In terms of socio-economic status, some authors found that economically disadvantaged families had fewer benefits than those who were not economically disadvantaged (Lundahl et al., 2006). Additionally, the delivery mode (individual vs. group) was found to affect the strength of the effects in specific social groups, i.e., parents who participated only in individually delivered parent training changed significantly more than those who participated only in group-delivered parent training among economically disadvantaged families (Lundhal et al., 2006). With regards to the delivery mode, no significant difference was found between self – directed parent training and face – to – face parent training (Lundhal et al., 2006).

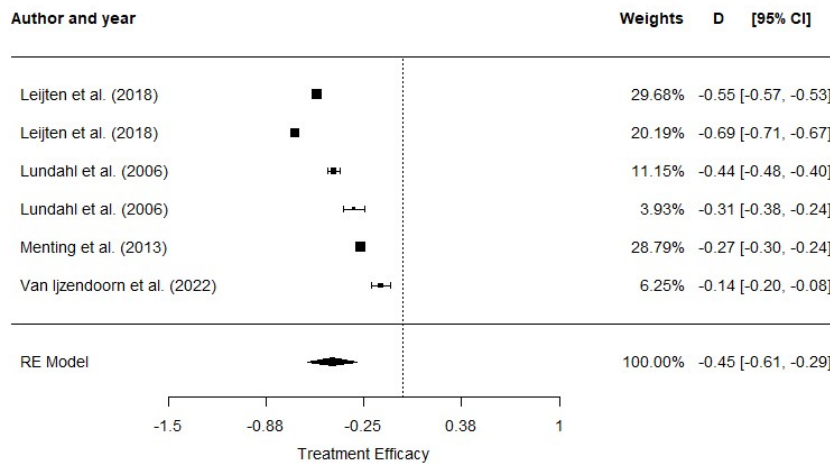
With respect to the type of control group used as comparison, it was reported that effects of parent training varied depending on the type of control group, with Treatment as Usual (TAU) comparisons yielding even stronger effects than comparisons with no active treatment (Comer et al., 2013). On the other hand, the number of sessions attended by parents was positively related to the intervention effects (Menting et al., 2013). Additionally, with regards to parent training, effects measured through teacher reports were smaller than the ones reported by parents (Menting et al., 2013).

Figure 2. Forest Plot of the Results of the Meta-Analyses of All Interventions



Note. Black boxes represent the ESs of the studies, the lines through them correspond to the 95% CI.

Figure 3. Forest Plot of The Results of The Meta-Analyses of Parent Training Interventions



Note. Black boxes represent the ESs of the studies, the lines through them correspond to the 95% CI.

Quantitative results: Umbrella pooling of ESs

Table 3 displays the results of each meta-analysis. Different types of indicators for ES were found; Cohen’s *d* was used in eight of the 12 associations, Hedges’ *g* in two, the correlation coefficient *r* in one, and SMD in one. Therefore, all measures were transformed to Cohen’s *d*. Some studies used the plus sign (+) in ES (e.g., Comer et al., 2013) to indicate that the treatment was effective; however, other authors used the minus sign (-) (e.g., Veenman et al., 2018) to indicate a decrease in behavioural problems. For our study, we decided to homogenise the results so that a minus sign indicates that CP decreased, and a

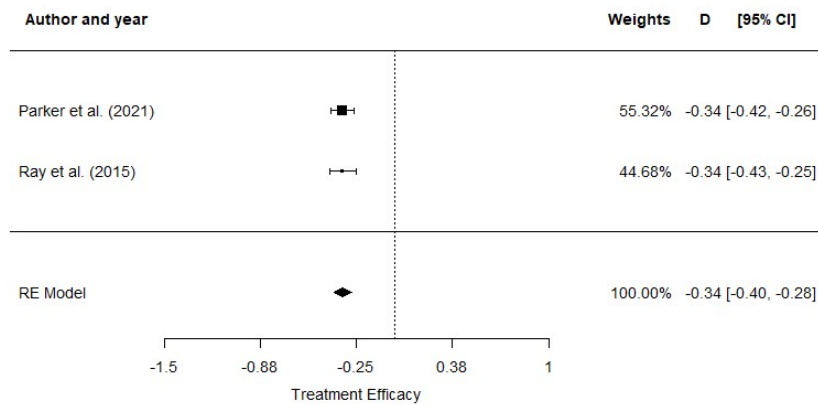
plus sign indicates that they increased. Results expressed in Cohen’s *d* ranged from -0.82 (95% CI = -1.01; -.63) (Comer et al., 2013) to -.14 (95% CI = -.04; -.19) (Van Ijzendoorn et al., 2022).

The *I*² statistic ranged from 0% (Parker et al., 2021) to 73.6% (Burkey et al., 2018), and was classified as moderate heterogeneity (Higgins & Thomson, 2002) in four studies (Burkey et al., 2018; Leijten et al., 2018; Van Ijzendoorn et al., 2022; Veenman et al., 2018).

Pooling results from all types of psychological interventions

Considering all nine meta-analyses, 278 studies (median *k*= 13.5) and 26,783 individuals (median *n* = 1,460), the average ES weighted

Figure 4. Forest Plot of the Results of The Meta-Analyses of Play Therapy Interventions



Note. Black boxes represent the ESs of the studies, the lines through them correspond to the 95% CI.

by the sample size was $d = -0.49$ (95% CI -0.67 to -0.32). The corresponding Z value of -5.47 was statistically significant ($p < .0001$). Evidence of significant heterogeneity was supported by the significance test for the observed value of Q and the I^2 statistical ($Q^{11} = 5688.38$, $p < .001$, $I^2 = 99.76\%$). It was also found that 87.5% of the meta-analyses had a significant ES. Figure 2 provides a forest plot of the ESs, the weights, and the 95% intervals for all included studies.

Pooling results from parent training interventions

Six meta-analyses grouped 167 individual studies (median $k = 37$) and 16,480 individuals (median $n = 2,582$) to evaluate the specific efficacy of parent training interventions. The mean ES weighted by sample size was $d = -.45$ (95% CI $-.61$ to $-.29$). The corresponding Z -value of -5.60 was statistically significant ($p < .001$). Evidence of significant heterogeneity was found by Q -test and I^2 statistic ($Q = 910.78$, $p < .001$, $I^2 = 99.24\%$). The highest ES was $d = -.69$ (Leijten et al., 2018) and 91.67% of the associations had a significant ES, i.e., all but one of the meta-analyses (Van Ijzendoorn et al., 2022). Figure 3 shows a forest plot of these associations.

Pooling results from play therapy interventions

The two meta-analyses who evaluated the specific efficacy of play therapy included 21 studies (median $k = 10.5$) and 911 individuals (median = 455.5). The mean ES weighted by sample size was $d = -.34$ (95% CI $-.40$ to $-.28$); the corresponding Z value of -11.59 was statistically significant ($p < .001$). No evidence of significant heterogeneity was found by Q test and statistical I^2 ($Q = 0$, $p = 1$, $I^2 = 0\%$). The ES in the two meta-analyses was $d = -.34$, significant in both cases. Figure 4 shows a forest plot of these meta-analyses.

Risk of bias in included studies and quality and credibility assessment

Assessment of potential risk of bias was reported in three of the nine meta-analyses using the Cochrane Collaboration Tool (Higgins et al., 2011). The meta-analysis by Burkey et al. (2018) indicated that 32% of their individual studies were at high risk of bias. Their main problem was blinding of outcome informants. The meta-analysis by Leijten et al. (2018) indicated that, for most of the individual studies, the risk of bias was low, with appropriate blinding of outcome informants, treatment of incomplete data, analysis of dropouts, and selective outcome reporting. As for the meta-analysis by Van Ijzendoorn et al.

(2022), the risk of bias raised some concern in more than half of the studies (56%).

The AMSTAR-2 rating of overall confidence in the reviews was considered critically low in six studies, moderate in two studies and high in one study. At first, the interobserver agreement for all items was 77.7%. Differences in criteria were discussed and finally the agreement obtained was 100%. The items that were most frequently unmet or that omitted information were registering the protocol beforehand (item 2), providing a list of excluded studies and justifying the exclusions (item 7), items related to the risk of bias (items 9, 12, 13), and reporting on the sources of funding of the studies included in the review (item 10). The AMSTAR-2 assessment is presented in Table A8 in the Appendix.

Regarding credibility, 10 associations obtained suggestive evidence (meeting the following criteria: number of cases $> 1,000$, $p < 10^{-3}$ and class I-II criteria not met), one obtained weak evidence (meeting the criteria $p < .05$ and class I-III not met) (Ray et al., 2015) and one was not significant because it did not meet the criterion of statistical significance (Van Ijzendoorn et al. 2022).

Publication bias across studies

Orwin's method estimated that 12 (for all treatments), six (for parent training) and two (for play therapy) missing studies with a null effect would be necessary to reduce the pooled ES by half in each association. Thus, the number of meta-analyses required with zero ES is unlikely to be met for the "all treatments" analysis; robustness of results for specific types of interventions (and, especially, play therapy) appear to be weaker in terms of publication bias.

Discussion

The development of effective and feasible treatments for children's CP is a major aspiration for clinical child psychology. Many experiences and trials have been reported in the scientific literature; in parallel, a variety of meta-analytic studies have been published, providing material for higher-level umbrella reviews (Fusar-Poli & Radua, 2018). This study aimed to organize and summarize the evidence provided by meta-analyses performed during the last 20 years, so that we can achieve a more comprehensive view of what has been studied and with which results.

Our review shows that family programs, and, specifically, parent

training programs are, by far, the most studied interventions in this field. Parent training programs are grounded on the evidence proving that parenting practices are a major component in the pathways to CP (Patterson, 1982). They are typically based on Operant Conditioning Learning and Social Learning Theory, and they aim to induce changes in the parent-child interactions, by breaking coercive relationships and by promoting consistent and effective discipline. Parenting programs have gained recognition over the years, and, in fact, they are often considered as first-line treatment for CP at childhood (NICE, 2013). Our review estimates a pooled weighted ES of $d = -.45$, which is significant, but still not high according to the usual thresholds (Cohen, 1988).

In this line, identifying the factors that boost or reduce the effect of parent training emerges as a meaningful question, and the meta-analyses reviewed by this umbrella provide some relevant insights. For example, some results distil which ingredients may be more active within parent training programs (Leitjen et al., 2018): positive reinforcement (particularly, social praise) and natural/logical consequences seem to be the associated to stronger effects in the intervention on CP. Such evidence may help researchers and practitioners in selecting potentially effective programs, and also in designing new parenting-based interventions.

The reviewed studies could also identify some moderators of the effects of parent training, considering either participants' or intervention's characteristics. In this regard, results show that parent training may not be equally useful for all the intervention levels (Leitjen et al., 2018; Menting et al., 2013): effects are stronger for treatment than for indicated prevention (Burkey et al., 2018; Menting et al., 2013), and they are stronger for indicated than for universal or selective prevention levels (Leitjen et al., 2018). This pattern of results may respond to the different motivations for enrolling in treatment and prevention conditions; parenting programs are often perceived as quite demanding by participants (Nock & Ferriter, 2005), and it may be plausible that treatment (i.e., help-seeking) parents are more willing to comply with attendance and tasks required by the programs, in comparison to parents who have been recruited in the community after screening processes. Alternatively, the difference in efficacy between treatment and prevention conditions may be related to CP severity. Previous research has already shown that intervention effects may be stronger for more severe CP cases (e.g., Leitjen et al., 2013), as there is a wider range of potential improvement; this pattern was also found in the meta-analyses reviewed by the present study (e.g., Lundahl et al., 2006; Menting et al., 2013).

Differential effects related to delivery modalities were also examined (Lundahl et al., 2006). Group formats were found to be as efficacious as individually administered programs (e.g., Comer et al., 2013), a finding that favors the group delivery when cost-effectiveness is a priority. Nonetheless, this finding seems to be moderated by socioeconomic conditions: disadvantaged families tend to obtain less benefit from parent training, and this holds particularly true when the programs are administered in groups (Lundahl et al., 2006). The stronger efficacy of individual vs. group-oriented interventions for disadvantaged families may be due to multi-problematic circumstances (Tausenfreund et al., 2016), including financial strain, lower literacy, psychosocial stress, and lack of social support, which may raise personal and situational barriers to the participation in group interventions. This may be an important take-home message, as disadvantaged families may need a more flexible, individualized approach that nourishes motivation and engagement, considering the net of family problems coexisting with CP.

In fact, the degree of participation in the programs was also identified as another efficacy moderator (Menting et al., 2013); this finding suggests a dose-effect relationship, and reinforces the need to prioritize motivation and adherence, because even the best quality programs work just if parents are appropriately involved (Nix et al., 2009). In this sense, along with good delimitation of effective contents and training techniques, research in this field should go deeper into the factors and processes that influence implication and permanence in the interventions (Hackworth et al., 2018).

In terms of effect maintenance, the reviewed meta-analyses prove that the effects of parent training may be significant up to one-year follow-up (Leitjen et al., 2018; Lundahl et al., 2006); however, results also show that effects are attenuated over time (Lundahl et al., 2006). It has been suggested that a continued care model may be appropriate for families with CP, so that the effective parenting can be supported and adapted to developmental and family changes (Lundahl et al., 2006). It has also been suggested that effect maintenance could be bolstered by including non-behavioural components (e.g., communication styles, improving emotional family climate) that may contribute to the internalization of the values and principles of positive parenting (Lundahl et al., 2006). Yet, the usefulness of including such relational components is still controversial (Leitjen et al., 2018), and more long-term evaluations of parenting programs need to be conducted (see, for example, Romero et al., 2017). The need of long-term studies in this field is also implied by Van Ijzendoorn et al. (2022), who suggest that attachment-oriented components, potentially effective to sustain the parenting changes, may take longer to show their effects on child behaviour.

Despite the gaps that still must be covered (also addressed below), in general, the meta-analyses based on more than 160 individual studies and 16,000 participants support parent training as an effective intervention for CP problems. Interestingly, Comer et al. (2013) found that the effects are stronger when parent training is compared to TAU than to non-active treatment, thus revealing that the usual service given to CP children is inadequate, and that the evidence-supported treatments should be better disseminated. As other authors have claimed (Lebrun-Harris et al., 2022), attention to child mental health should be strengthened, especially considering the increasing rates of behavioural and emotional problems in young people. Such reinforcement may involve not only an improvement in quantity of resources, but also in *quality* of training and interventions, with more effective insertion of research results into the clinical practice.

Apart from parent training, not many eligible meta-analyses were focused on other specific types of treatments. An exception comes from child-centered play therapy, with two studies entirely devoted to this approach (Parker et al., 2021; Ray et al., 2015). Child-centered play therapy, with roots in humanistic theory, is a modality of person-centered counselling based on the use of toys and plays to meet the developmental needs of children (Landreth, 2012); with a non-directive approach, child-centered play therapy emphasizes the communication with children, with a supportive, safe, and caring environment, which promotes the development of healthy attitudes and skills. Our study, pooling results from meta-analyses on child-centered play therapy, shows a significant ES ($d = -.34$); nevertheless, a cautious consideration of this ES is sensible, as it stems from a low number of meta-analyses and individual studies, with relatively low sample sizes, and with weak AMSTAR-2 and credibility assessments. Research on child-centered play therapy should be encouraged, as preliminary results may be promising; however, more investigation

is needed not only to make this evidence more robust, but also to describe longer-term effects and to identify efficacy moderators.

Limitations and future directions

Results of this umbrella review need to be seen in the light of some limitations. First, although high-quality studies (e.g., based on RCTs) are increasingly conducted in this field, research quality is quite uneven, often including small sample sizes, inadequate reporting of important data, and blurry definition of treatments. Likewise, the AMSTAR-2 quality of the reviewed meta-analyses was often compromised. Of note, most meta-analyses were in fact conducted before the dissemination of the AMSTAR-2 chart in the scientific community (Shea et al., 2017); thus, it is expected that future reviews will meet the stringent requirements of current protocols. Overall, a proactive consideration of potential reviews should be already adopted by primary studies, as the lack of information in the individual studies drives a *domino effect* on the quality of future secondary (meta-analyses) and tertiary (umbrella) reviews. Second, this umbrella review could just pool results from two treatment types (i.e., parent training and child-centered play therapy). Other types, like child-centered cognitive-behavioural therapy (CBT) could not be analyzed; although meta-analyses have been conducted on child-centered CBT (e.g., Battagliese et al., 2015), they typically cluster together results from children and adolescents, thus precluding the specific analyses required by this review. Expectedly, more research will be accumulated during the following years, separately addressing treatment effects on individuals in different developmental stages; in this line, the effect of age on treatment effects deserves more consideration; results on this topic are far from being clarified, as it is suggested by the present review. Third, this study provided pooled ESs for all treatments, and for two specific types of treatments, but significant heterogeneity was found across studies; to the extent that more meta-analyses can be further conducted, the sources of such variability should be a target of future umbrella reviews. Relatedly, the study of moderators of treatment effects should be a priority for research (see also McMahon et al., 2021). Although the reviewed studies show some interesting results, much more is needed to be known, for example, in terms of comorbidities, outcome informants, intervention length, combined treatments, or modality of delivery; specifically, the effects of digitalized interventions as compared to traditional formats deserve a closer view, given the current bourgeon of e-Health interventions, and the need of rendering programs accessible to the communities (Florea et al., 2020). Finally, the diversity of children with CP should be considered. For the last years, the phenotypic, etiological, and developmental heterogeneity of CP has been emphasized (Waller et al., 2013). Accordingly, the need of tailored interventions has been highlighted (Glenn et al., 2019; Romero et al., 2019), for a better tuning of programs to the specific needs of children and families; the effects of such interventions should come at the forefront for further research.

Concluding remarks

Our umbrella review indicates that there is suggestive evidence of efficacy for treatment of childhood CP. Evidence is more robust for parent training, with a number of moderators suggested in terms of individual, family or implementation factors. Further research on effective and sustainable treatments emerges as a major practical need nowadays, in a socio-historical context of recognized vulnerability for children's mental health; at the same time, the inquiry into effective

treatments will provide valuable feedback for theory building on etiology and developmental pathways of CP. So far, results point at the efficacy of specific treatment approaches, but results also indicate that there is much room for improvement. As in other health disciplines (e.g., Allen et al., 2009), rather than wondering *what works*, research may need to be pushed into more refined questions: *what works, for whom, and under which circumstances*.

Funding

This study was supported by grant PID2019-107897RB-I00, with funds from MCIN/AEI/10.13039/501100011033

References

- Allen, D., Gillen, E., & Rixson, L. (2009). Systematic review of the effectiveness of integrated care pathways: what works, for whom, in which circumstances? *International Journal of Evidence-Based Healthcare*, 7(2), 61–74. <https://doi.org/10.1111/j.1744-1609.2009.00127.x>
- Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *International Journal of Evidence-Based Healthcare*, 13(3), 132–140. <https://doi.org/10.1097/XEB.0000000000000055>
- Bakker, M. J., Greven, C. U., Buitelaar, J. K., & Glennon, J. C. (2017). Practitioner Review: Psychological treatments for children and adolescents with conduct disorder problems—a systematic review and meta-analysis. *Journal of Child Psychology and Psychiatry*, 58(1), 4–18. <https://doi.org/10.1111/jcpp.12590>
- Battagliese, G., Caccetta, M., Luppino, O. I., Baglioni, C., Cardi, V., Mancini, F., & Buonanno, C. (2015). Cognitive-behavioural therapy for externalizing disorders: A meta-analysis of treatment effectiveness. *Behaviour Research and Therapy*, 75, 60–71. <https://doi.org/10.1016/j.brat.2015.10.008>
- Belbasis, L., Bellou, V., Evangelou, E., Ioannidis, J. P., & Tzoulaki, I. (2015). Environmental risk factors and multiple sclerosis: an umbrella review of systematic reviews and meta-analyses. *The Lancet Neurology*, 14(3), 263–273. [https://doi.org/10.1016/S1474-4422\(14\)70267-4](https://doi.org/10.1016/S1474-4422(14)70267-4)
- Bellou, V., Belbasis, L., Tzoulaki, I., Evangelou, E., & Ioannidis, J. P. (2016). Environmental risk factors and Parkinson's disease: an umbrella review of meta-analyses. *Parkinsonism & Related Disorders*, 23, 1–9. <https://doi.org/10.1016/j.parkreldis.2015.12.008>
- Brinkmeyer, M. Y., & Eyberg, S. M. (2003). Parent-child interaction therapy for oppositional children. In A. E. Kazdin & J. R. Weisz (Eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 204–223). The Guilford Press.
- Broidy, L. M., Nagin, D. S., Tremblay, R. E., Bates, J. E., Brame, B., Dodge, K. A., Fergusson, D., Horwood, J. L., Loeber, R., Laird, R., Lynam, D. R., Moffitt, T. E., Pettit, G. S. & Vitaro, F. (2003). Developmental trajectories of childhood disruptive behaviours and adolescent delinquency: A six-site, cross-national study. *Developmental Psychology*, 39(2), 222–245. <https://doi.org/10.1037/0012-1649.39.2.222>
- *Burkey, M. D., Hosein, M., Morton, I., Purgato, M., Adi, A., Kurzrok, M., Kohrt, B. A., & Tol, W. A. (2018). Psychosocial interventions for disruptive behaviour problems in children in low- and middle-income countries: A systematic review and meta-analysis. *Journal of Child Psychology and Psychiatry*, 59(9), 982–993. <http://dx.doi.org/10.1111/jcpp.12894>
- Cicchetti, D. (Ed.) (2016). *Developmental psychopathology* (Vol. 3, *Maladaptation and Psychopathology*, 3rd ed.). Wiley.

- Cicchetti, D., Rogosch, F. A., Lynch, M., & Holt, K. D. (1993). Resilience in maltreated children: Processes leading to adaptive outcome. *Development and Psychopathology*, 5(4), 629–647. <https://doi.org/10.1017/S0954579400006209>
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*, 2nd ed. Erlbaum.
- *Comer, J. S., Chow, C., Chan, P. T., Cooper-Vince, C., & Wilson, L. A. S. (2013). Psychosocial treatment efficacy for disruptive behaviour problems in very young children: A meta-analytic examination. *Journal of the American Academy of Child & Adolescent Psychiatry*, 52(1), 26–36. <http://dx.doi.org/10.1016/j.jaac.2012.10.001>
- Florean, I. S., Dobrean, A., Păsărelu, C. R., Georgescu, R. D., & Milea, I. (2020). The Efficacy of internet-based parenting programs for children and adolescents with behavior problems: A meta-analysis of randomized clinical trials. *Clinical Child and Family Psychology Review* 23, 510–528. <https://doi.org/10.1007/s10567-020-00326-0>
- Furlong, M., & McGilloway, S. (2012). The Incredible Years Parenting program in Ireland: A qualitative analysis of the experience of disadvantaged parents. *Clinical Child Psychology and Psychiatry*, 17(4), 616–630. <https://doi.org/10.1177/1359104511426406>
- Fusar-Poli, P., & Radua, J. (2018). Ten simple rules for conducting umbrella reviews. *Evidence-Based Mental Health*, 21(3), 95–100. <https://doi.org/10.1136/ebmental-2018-300014>
- Ghandour, R. M., Sherman, L. J., Vladutiu, C. J., Ali, M. M., Lynch, S. E., Bitsko, R. H., & Blumberg, S. J. (2019). Prevalence and treatment of depression, anxiety, and conduct problems in US children. *The Journal of Pediatrics*, 206, 256–267. <https://doi.org/10.1016/j.jpeds.2018.09.021>
- Girard, L. C., Tremblay, R. E., Nagin, D., & Côte, S. M. (2019). Development of aggression subtypes from childhood to adolescence: A group-based multi-trajectory modelling perspective. *Journal of Abnormal Child Psychology*, 47(5), 825–838. <https://doi.org/10.1007/s10802-018-0488-5>
- Glenn, A. L., Lochman, J. E., Dihson, T., Powell, N. P., Boxmeyer, C., Kassing, F., Qu, L., & Romero, D. (2019). Toward tailored interventions: Sympathetic and parasympathetic functioning predicts responses to an intervention for conduct problems delivered in two formats. *Prevention Science*, 20, 30–40. <https://doi.org/10.1007/s11121-017-0859-0>
- Gordon, R. (1987). An operational classification of disease prevention. In J. A. Steinberg & M. M. Silverman (Eds.), *Preventing mental disorders* (pp. 20–26). Department of Health and Human Services.
- Gutman, L. M., Joshi, H., & Schoon, I. (2019). Developmental trajectories of conduct problems and cumulative risk from early childhood to adolescence. *Journal of Youth and Adolescence*, 48(2), 181–198. <https://doi.org/10.1007/s10964-018-0971-x>
- Guyatt, G. H., Oxman, A. D., Schünemann, H. J., Tugwell, P., & Knottnerus, A. (2011). GRADE guidelines: A new series of articles in the Journal of Clinical Epidemiology. *Journal of Clinical Epidemiology*, 64(4), 380–382. <https://doi.org/10.1016/j.jclinepi.2010.09.011>
- Hackworth, N. J., Matthews, J., Westrupp, E. M., Nguyen, C., Phan, T., Scicluna, A., Cann, W., Bethelsen, D., Bennetts, S. K., & Nicholson, J. M. (2018). What influences parental engagement in early intervention? Parent, program and community predictors of enrolment, retention and involvement. *Prevention Science* 19, 880–893. <https://doi.org/10.1007/s11121-018-0897-2>
- Harley, M., Murtagh, A., & Cannon, M. (2008). Conduct disorder—psychiatry’s greatest opportunity for prevention: A commentary on ‘From conduct disorder to severe mental illness: associations with aggressive behaviour, crime and victimization’ by Hodgins et al. (2007). *Psychological Medicine*, 38(7), 929–931. <https://doi.org/10.1017/S0033291708002699>
- Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11), 1539–1558. <https://doi.org/10.1002/sim.1186>
- Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L., & Sterne, J. A. (2011). The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials. *Bmj*, 343. <https://doi.org/10.1136/bmj.d5928>
- Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (eds). (2019) *Cochrane handbook for systematic reviews of interventions*. 2nd Edition. John Wiley & Sons.
- Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *Bmj*, 327(7414), 557–560. <https://doi.org/10.1136/bmj.327.7414.557>
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage.
- Lahey, B. B., Van Hulle, C. A., Waldman, I. D., Rodgers, J. L., D’Onofrio, B. M., Pedlow, S., Rathouz, P. & Keenan, K. (2006). Testing descriptive hypotheses regarding sex differences in the development of conduct problems and delinquency. *Journal of Abnormal Child Psychology*, 34, 737–755. <https://doi.org/10.1007/s10802-006-9064-5>
- Landreth, G. L. (2012). *Play Therapy. The Art of the Relationship*. Routledge.
- Lebrun-Harris, L. A., Sappenfield, O. R., & Warren, M. D. (2022). Missed and delayed preventive health care visits among US children due to the COVID-19 pandemic. *Public Health Reports*, 137(2), 336–343. <https://doi.org/10.1177/00333549211061322>
- Leflot, G., van Lier, P. A. C., Onghena, P., & Colpin, H. (2010). The role of teacher behavior management in the development of disruptive behaviors: An intervention study with the good behavior game. *Journal of Abnormal Child Psychology* 38, 869–882. <https://doi.org/10.1007/s10802-010-9411-4>
- *Leijten, P., Gardner, F., Melendez-Torres, G. J., Van Aar, J., Hutchings, J., Schultz, S., Knerr, W., & Overbeek, G. (2018). What to teach parents to reduce disruptive child behaviour: two meta-analyses of parenting program components. *Journal of the American Academy of Child and Adolescent Psychiatry*, 58(2). <https://doi.org/10.1016/j.jaac.2018.07.900>
- Leijten, P., Raaijmakers, M. A., de Castro, B. O., & Matthys, W. (2013). Does socioeconomic status matter? A meta-analysis on parent training effectiveness for disruptive child behaviour. *Journal of Clinical Child & Adolescent Psychology*, 42(3), 384–392. <https://doi.org/10.1080/15374416.2013.769169>
- Lochman, J. E., & Wells, K. C. (2002). Contextual social-cognitive mediators and child outcome: A test of the theoretical model in the Coping Power program. *Development and Psychopathology*, 14(4), 945–967.
- López-López, J. A., Rubio-Aparicio, M., & Sánchez-Meca, J. (2022). Overviews of reviews: Concept and Development. *Psicothema*, 34(2), 175–181. <https://doi.org/10.7334/psicothema2021.586>
- *Lundahl, B., Risser, H. J., & Lovejoy, M. C. (2006). A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review*, 26(1), 86–104. <http://dx.doi.org/10.1016/j.cpr.2005.07.004>
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3), 276–282.
- Masten, A., & Cicchetti, D. (2010). Developmental cascades. *Development and Psychopathology*, 22(3), 491–495. <https://doi.org/10.1017/S0954579410000222>
- McMahon, R. J., Goulter, N., & Frick, P. J. (2021). Moderators of psychosocial intervention response for children and adolescents with conduct problems. *Journal of Clinical Child & Adolescent Psychology*, 50(4), 525–533. <https://doi.org/10.1080/15374416.2021.1894566>
- *Menting, A. T. A., de Castro, B. O., & Matthys, W. (2013). Effectiveness of the incredible years parent training to modify disruptive and prosocial child behaviour: A meta-analytic review. *Clinical Psychology Review*, 33(8), 90–913. <http://dx.doi.org/10.1016/j.cpr.2013.07.006>

- Merikangas, K. R., Nakamura, E. F., & Kessler, R. C. (2022). Epidemiology of mental disorders in children and adolescents. *Dialogues in Clinical Neuroscience*, 11(1), 7–20. <https://doi.org/10.31887/DCNS.2009.11.1/krmerikangas>
- Ministerio de Sanidad, Consumo y Bienestar Social del Gobierno de España (2017). Encuesta Nacional de Salud [ENSE] https://www.sanidad.gob.es/estadEstudios/estadisticas/encuestaNacional/encuestaNac2017/ENSE2017_notatecnica.pdf
- Moffitt, T. E. (2003). Life-course-persistent and adolescence-limited antisocial behaviour: A 10-year research review and a research agenda. En B. B. Lahey, T. E. Moffitt y A. Caspi (Eds.), *Causes of conduct disorder and juvenile delinquency* (pp. 49–75). Guilford Press.
- Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behaviour: A developmental taxonomy. *Psychological Review*, 100, 674–701. <https://doi.org/10.1037/0033-295X.100.4.674>
- Munder, T., Brüttsch, O., Leonhart, R., Gerger, H., & Barth, J. (2013). Researcher allegiance in psychotherapy outcome research: An overview of reviews. *Clinical Psychology Review*, 33(4), 501–511. <https://doi.org/10.1016/j.cpr.2013.02.002>
- National Collaborating Centre for Mental Health (2013). *Antisocial behaviour and conduct disorders in children and young people: The NICE guideline on recognition, intervention and management*. London: The British Psychological Society and The Royal College of Psychiatrists.
- Nix, R. L., Bierman, K. L., McMahon, R. J., & The Conduct Problems Prevention Research Group. (2009). How attendance and quality of participation affect treatment response to parent management training. *Journal of Consulting and Clinical Psychology*, 77(3), 429–438. <https://doi.org/10.1037/a0015028>
- Nock, M. K., & Ferriter, C. (2005). Parent management of attendance and adherence in child and adolescent therapy: A conceptual and empirical review. *Clinical Child and Family Psychology Review*, 8(2), 149–166. <https://doi.org/10.1007/s10567-005-4753-0>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Bmj*, 372, 71. <https://doi.org/10.1136/bmj.n71>
- *Parker, M. M., Hunnicutt Hollenbaugh, K. M., & Kelly, C. T. (2021). Exploring the impact of child-centered play therapy for children exhibiting behavioural problems: A meta-analysis. *International Journal of Play Therapy*, 30(4), 259–271. <http://dx.doi.org/10.1037/pla0000128>
- Patterson, G. R. (1982). *Coercitive family process: A social learning approach to family intervention* (vol. 3). Castalia Publishing Company.
- *Ray, D. C., Armstrong, S. A., Balkin, R. S., & Jayne, K. M. (2015). Child-centered play therapy in the schools: review and meta-analysis. *Psychology in the Schools*, 52(2), 107–123. <https://doi.org/10.1002/pits.21798>
- Rissanen, E., Kuvaja-Köllner, V., Elonheimo, H., Sillanmäki, L., Sourander, A., & Kankaanpää, E. (2022). The long-term cost of childhood conduct problems: Finnish Nationwide 1981 Birth Cohort Study. *Journal of Child Psychology and Psychiatry*, 63(6), 683–692. <https://doi.org/10.1111/jcpp.13506>
- Romero, E., Gómez-Fraguela, X. A., Villar, P., & Rodríguez, C. (2019). Prevención indicada de los problemas de conducta: entrenamiento de habilidades socioemocionales en el contexto escolar. *Revista de Psicología con Niños y Adolescentes*, 6(3), 39–47. <https://doi.org/10.21134/rpcna.2019.06.2.1>
- Romero, E., López-Romero, L., Domínguez-Álvarez, B., Villar, P., & Gómez-Fraguela, J. A. (2020). Testing the effects of COVID-19 confinement in Spanish children: The role of parents' distress, emotional problems and specific parenting. *International Journal of Environmental Research and Public Health*, 17(19), 6975. <https://doi.org/10.3390/ijerph17196975>
- Romero, E., Rodríguez, C., Villar, P., & Gómez-Fraguela, X. A. (2017). Intervention on early-onset conduct problems as indicated prevention for substance use: A seven-year follow up. *Adicciones*, 29(3), 150–162.
- Schaeffer, C. M., Petras, H., Ialongo, N., Masyn, K. E., Hubbard, S., Poduska, J., & Kellam, S. (2006). A comparison of girls' and boys' aggressive-disruptive behaviour trajectories across elementary school: Prediction to young adult antisocial outcomes. *Journal of Consulting and Clinical Psychology*, 74(3), 500–510. <https://doi.org/10.1037/0022-006X.74.3.500>
- Schardt, C., Adams, M. B., Owens, T., Keitz, S., & Fontelo, P. (2007). Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Medical Informatics and Decision Making*, 7, 16. <https://doi.org/10.1186/1472-6947-7-16>
- Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., & Henry, D. A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions. *Bmj*, 358. <https://doi.org/10.1136/bmj.j4008>
- Tausendfreund, T., Knot-Dickscheit, J., Schulze, G. C., Knorth, E. J., & Grietens, H. (2016). Families in multi-problem situations: Backgrounds, characteristics, and care services. *Child & Youth Services*, 37(1), 4–22. <https://doi.org/10.1080/0145935X.2015.1052133>
- *van IJzendoorn, M. H., Schuengel, C., Wang, Q., & Bakermans-Kranenburg, M. J. (2022). Improving parenting, child attachment, and externalizing behaviours: Meta-analysis of the first 25 randomized controlled trials on the effects of Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline. *Development and Psychopathology*, 1–16. <https://doi.org/10.1017/S0954579421001462>
- *Veenman, B., Luman, M., & Oosterlaan, J. (2018). Efficacy of behavioural classroom programs in primary school. A meta-analysis focusing on randomized controlled trials. *PLoS ONE*, 13(10), 23. <http://dx.doi.org/10.1371/journal.pone.0201779>
- Waller, R., Gardner, F., & Hyde, L. W. (2013). What are the associations between parenting, callous–unemotional traits, and antisocial behaviour in youth? A systematic review of evidence. *Clinical Psychology Review*, 33, 593–608. <http://dx.doi.org/10.1016/j.cpr.2013.03.001>

*Articles included in the umbrella review

Supplementary materials

Additional supporting information (Appendix) can be found online at <https://osf.io/6wvb5/>

Table A1. PRISMA Checklist

Table A2. Search Strategy Used

Table A3. Main Results of each Study Included in the Umbrella Review

Table A4. Databases Searched, Search Period and Citations in WOS for Each Meta-Analysis

Table A5. Overlapping of Individual Studies Included in the Meta-Analyses

Table A6. Excluded Items After Abstract and Full Text Review and Reasons (N=27)

Table A7. Individual Studies Included in the Selected Meta-Analyses

Table A8. AMSTAR-2 Quality Assessment